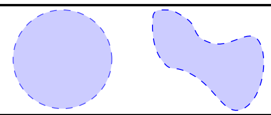
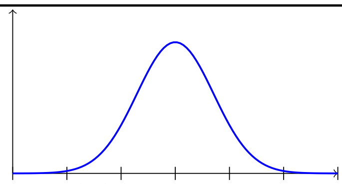
Motivation for calculus



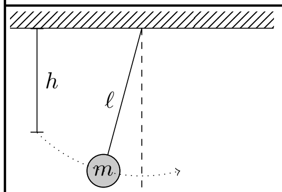
Given a fence of a fixed perimeter, what is the maximum possible area inscribed by that fence?

Amol drives on average 45 mph from NYC to Boston and gets a speeding ticket on I-91, where the speed limit is 65 mph. How is this possible?

A very large class of students takes an exam and their normalized grade distribution is shown. What is the probability that a student chosen at random obtained a grade between 85 and 95?



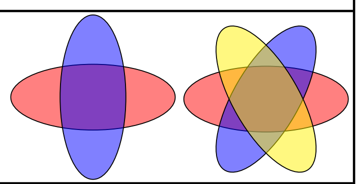
If $[0,100] \ni x \mapsto p(x)$ describes the probability density associated with a grade of x, then the integral $\int_{85}^{95} p(x)\,dx$ gives this probability.

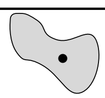


A pendulum of length ℓ and mass m is released from a fixed height h. What is the period of oscillation of the pendulum? Assume air resistance, friction, and other forces are absent. In particular, does the period depend on the mass m, the initial height h, or the length ℓ?

If the initial height is h, then the initial angle is $\theta_0 = \arccos\left(\frac{h}{\ell}\right)$ with θ_0 measured clockwise from the vertical axis. The period is given by the integral $T = 4\sqrt{\frac{\ell}{2g}}\int_0^{\theta_0}\frac{d\theta}{\sqrt{\cos(\theta)-\cos(\theta_0)}}$. [Reference: Wiki Pendulum (mathematics)]

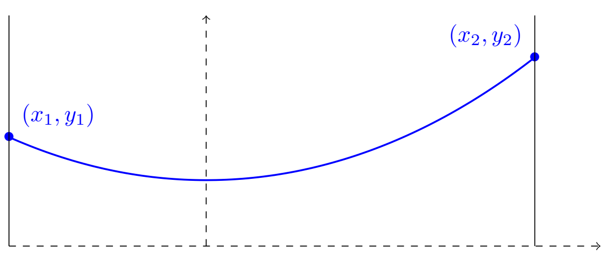
What is the area of the common region shared by two ellipses whose semi-major axes are at 90° from each other? How about three ellipses at 60°?





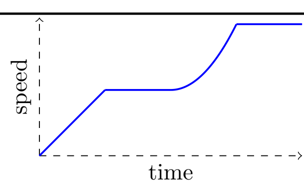
What is the center of mass of a rigid body whose mass distribution is known?

What is the shape of a hanging rope whose two ends are fixed when it is exposed to a constant gravitational field? Assume the rope cannot stretch.



The general curve y is of the form $y(x) = a\cosh\left(\frac{x}{a}\right)$ and the relationship between a and the length ℓ and the horizontal and vertical distances x_2-x_1 and y_2-y_1 is given by $\ell^2-(y_2-y_1)^2 = 4a^2\sinh^2\left(\frac{x_2-x_1}{2a}\right)$. [Reference: Michael Fowler (UVa)]

What is the distance covered by a car if we know the speed at all instances of time and the duration of the trip?

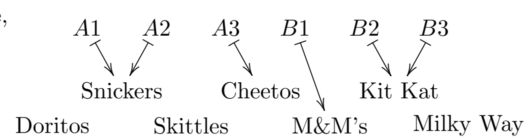


Functions

If A is a set, the notation $a \in A$ denotes "a is an element of A." If A and B are sets, $A \subseteq B$ denotes "A is a subset of B" and means every element of A is an element of B. $b \notin A$ means b is not an element of A. $B \setminus A := \{b \in B \ : \ b \notin A\}$, which reads "$B \setminus A$ is the set of elements b in B such that b is not an element of A." $A \times B := \{(a,b) \ : \ a \in A, b \in B\}$ is the set of ordered pairs of elements in A and B. $A \times B$ is called the *cartesian product* of A and B. [Reference: Herb Gross I.2. Functions]
$\mathbb{N} := \{1,2,3,4,\dots\}$ is the set of *natural numbers*,
$\mathbb{Z} := \{\dots,-3,-2,-1,01,2,3,\dots\}$ is the set of *integers*,
$\mathbb{Q} := \{p/q \ : \ p,q \in \mathbb{Z}\}$ is the set of *rational numbers*, and
$\mathbb{R}$ is the set of all *real numbers*.
Let $a, b \in \mathbb{R}$. The following intervals appear often:
$[a,b] := \{r \in \mathbb{R} \ : \ a \le r \le b\}$, $[a,b) := \{r \in \mathbb{R} \ : \ a \le r < b\}$
$(a,b] := \{r \in \mathbb{R} \ : \ a < r \le b\}$, $(a,b) := \{r \in \mathbb{R} \ : \ a < r < b\}$.
It is an unfortunate face that the notation (a,b) is used to denote both the ordered pair and the open interval. Context should clarify.

Let A and B be sets. A *function* f from A to B, written $f : A \to B$, $A \xrightarrow{f} B$, or $B \xleftarrow{f} A$, is an assignment sending each element $a \in A$ to some element $f(a) \in B$. A is called the *source* (*domain*) of f and B is called the *target* (*codomain*) of f. $A \ni a \mapsto f(a) \in B$ is used to denote the assignment of f on elements of A. $f(a)$ is called the *image* of a under f. The collection of all images of elements, $\mathrm{image}(f) := \{f(a) \ : \ a \in A\}$, is called the *image* of f and is a subset of B. The image of f is sometimes denoted by $f(A)$. The latter notation is abusive because it does not make sense to apply f to A, which may or may not be an element of A.

A vending machine has several buttons on it such as $X := \{A1, A2, A3, B1, B2, B3\}$ and each button outputs a particular snack (upon payment and selection). Let Y be some fixed set of popular snacks such as $Y = \{\text{Doritos, Snickers, Skittles, Cheetos, M\&M's, Kit Kat, Milky Way}\}$. A properly working vending machine defines a function $f : X \to Y$. For example,

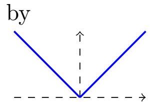


Notice that not all of the snacks were used. It is permissible for a function to not "hit" all of the elements in the target. Furthermore, it is possible for a snack to be "hit" more than once. The reason might be that a particular snack is so popular that it has two rows stocked.

$\mathbb{R} \setminus \{0\} \to \mathbb{R}$ defined by $\mathbb{R} \setminus \{0\} \ni x \mapsto \frac{1}{x}$ defines a function. Notice that this example of a function is quite different from the previous one in that we have an algebraic *formula* that describes the function.

We can break up parts of any set into subsets and define functions on these subsets to define a function on the whole. Such functions are called *piecewise* functions. An example of a piecewise function is the absolute value function $|\cdot| : \mathbb{R} \to \mathbb{R}$ given by

$$\mathbb{R} \ni x \mapsto \begin{cases} x & \text{if } x \ge 0 \\ -x & \text{if } x < 0 \end{cases}$$



Let A be any set and define $\mathrm{id}_A : A \to A$ to be the function sending $a \in A$ to a. This function is called the *identity* function on A.

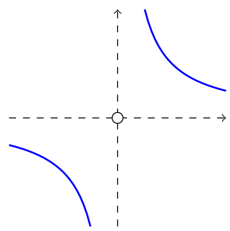
$\mathbb{R} \to \mathbb{R}$ given by $\mathbb{R} \ni x \mapsto \frac{1}{x}$ does not define a function because the assignment is not defined for $x = 0$.

$\mathbb{N} \to \mathbb{N}$ given by $\mathbb{N} \ni x \mapsto \frac{1}{x}$ also does not define a function because $\frac{1}{x}$ is not a natural number in general.

Let $A \xrightarrow{f} B$ be a function. The *graph* of f is the subset of $A \times B$ given by $\Gamma(f) := \{(a, f(a)) \ : \ a \in A\}$. Γ is the Greek letter "Gamma."

For the vending machine function from above, the graph is the list of elements in $X \times Y$ given by $\{(A1, \text{Snickers}), (A2, \text{Snickers}), (A3, \text{Cheetos}), (B1, \text{M\&M's}), \dots\}$.

The graph of $\mathbb{R} \setminus \{0\} \ni x \mapsto \frac{1}{x}$ can be visualized by drawing two coordinate axes by viewing the first element x in the ordered pair $\left(x, \frac{1}{x}\right)$ along the horizontal axis and viewing the second element along the vertical axis. The point $(0,0)$ has been circled to indicate that the function is not defined at $0 \in \mathbb{R}$. We often conflate a function with its graph but please be aware that they are different.



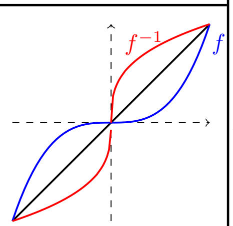
We often use the horizontal axis to indicate the source (domain) of the function and the vertical axis for the target (codomain).

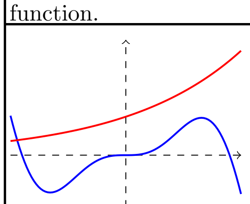
Let A, B, and C be sets and let $A \xrightarrow{f} B$ and $B \xrightarrow{g} C$ be two functions. The *composition* of f followed by g, written as $g \circ f$, is the function $A \xrightarrow{g \circ f} C$ defined by $A \ni a \mapsto g(f(a))$. You might wonder why the composition is written as $g \circ f$ instead of $f \circ g$. The reason is because for any function $h : X \to Y$ when we plug in an element $x \in X$ to the function h, we write $h(x)$ instead of $(x)h$—we place the input x on the right. For this reason, it is preferable to write the arrows from right to left, i.e. write $Y \xleftarrow{h} X$ so that $Y \ni h(x) \leftarrow\!\shortmid x \in X$. Another reason is for composition. The composition $C \xleftarrow{g} B \xleftarrow{f} A$ is $C \xleftarrow{g \circ f} A$. Notice that when we write our arrows from right to left, composition is just concatenation of symbols.

Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by $\mathbb{R} \ni x \mapsto f(x) := x^2$ and let $g : \mathbb{R} \to \mathbb{R}$ be given by $\mathbb{R} \ni y \mapsto g(y) := \cos(y)$. Both the compositions $g \circ f$ and $f \circ g$ are defined, but they are different. These two functions are given by $\mathbb{R} \ni x \mapsto g(f(x)) = \cos(x^2)$ and $\mathbb{R} \ni y \mapsto f(g(y)) = \cos^2(y)$, respectively.

A function $A \xrightarrow{f} B$ is *invertible* (or "reversible" in English) iff there exists a function $A \xleftarrow{g} B$ such that $f \circ g = \mathrm{id}_B$ and $g \circ f = \mathrm{id}_A$. Due to this and the fact that if such a g exists, it is unique, we often denote g by $g = f^{-1}$. f^{-1} is called the *inverse* of f. The meaning of an invertible function is more easily understood if f is thought of as an operation. f is invertible means there is another operation that can undo f, and this means that there is an operation g such that applying f and then g gets you back to where you started (and similarly if you apply g first and then f). It is a theorem that a function f is invertible if and only if it is one-to-one and onto. Recall, a function $A \xrightarrow{f} B$ is *onto* iff for every $b \in B$ there exists an $a \in A$ such that $f(a) = b$ and f is *one-to-one* iff $f(a_1) = f(a_2)$ implies $a_1 = a_2$. [Reference: Herb Gross I.3. Inverse Functions].

The function $\mathbb{R} \ni x \xmapsto{f} x^3$ is invertible and its inverse is $\mathbb{R} \ni x \xmapsto{f^{-1}} x^{1/3}$, the cube root of x. Notice that x is a variable and it is used for two different functions. We could have denoted it by anything else. Also notice that the *graph* of the inverse function f^{-1} is the reflection of the *graph* of f through the graph of the identity function.

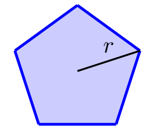




The function $\mathbb{R} \ni x \mapsto x^2 \sin(200x) \in \mathbb{R}$ is surjective but not invertible because it is not one-to-one (injective). The function $\mathbb{R} \ni x \mapsto \frac{1}{3}e^x \in \mathbb{R}$ is one-to-one (injective) but not invertible because it is not onto. However, the function $\mathbb{R} \ni x \mapsto \frac{1}{3}e^x \in (0, \infty)$ is invertible (because it is onto its codomain).

Exercises

Exercise 1. Fix a positive number r and a positive integer $n \geq 3$. Find the area of a regular n-gon whose distance from the center to any one of its vertices is r. For example, for $n = 3$, the regular 3-gon is an equilateral triangle and its area is $\frac{3\sqrt{3}}{4}r^2$. For $n = 4$, the shape is a square and its area is $2r^2$. What do you notice as n increases?



Exercise 2. Is the absolute value function $| \cdot | : \mathbb{R} \to \mathbb{R}$ onto? Is it one-to-one? Is it invertible? How about the absolute value function restricted to the domain and codomain given by $[0, \infty)$, i.e. $| \cdot | : [0, \infty) \to [0, \infty)$?

Exercise 3. What is the largest subset $A \subseteq \mathbb{R}$ such that the function $A \ni x \mapsto \frac{1}{1-x^2}$ is well-defined?

Exercise 4. What is the image of $\mathbb{R} \ni x \mapsto x + 5$? What is the inverse of this function? Be sure to identify the source of the inverse as well as the target. How about for the function $\mathbb{R} \ni x \mapsto 5x$?

Exercise 5. For which natural numbers $n \in \mathbb{N}$ is the function $\mathbb{R} \ni x \xmapsto{p_n} x^n \in \mathbb{R}$ invertible? For these values of n, find the inverse of p_n. For the values $m \in \mathbb{N}$ for which p_m is *not* invertible, explain why. For example, is p_m not one-to-one? Is p_m not onto? Sketch the graph of p_m for various values of m on the domain $[-1, 1]$. What do you notice happens as m increases? Try to sketch a graph of the limit of these functions $f := \lim\limits_{m \to \infty} p_m$ as m increases on the domain $[-1, 1]$.

Exercise 6. Using the notation from the previous exercise, let $m \in \mathbb{N}$ be a value for which p_m is not invertible. Let $q_m : [0, \infty) \to [0, \infty)$ be defined by the same formula as p_m (but with its domain and codomain slightly alterred). Is q_m invertible now? Explain.

Exercise 7. Let $C \xleftarrow{f} B$ and $B \xleftarrow{g} A$ be invertible functions with inverses given by $C \xrightarrow{f^{-1}} B$ and $B \xrightarrow{g^{-1}} A$, respectively. Is the composition $C \xleftarrow{f \circ g} A$ invertible? If so, find the inverse of $f \circ g$. If not, find an example of two functions f and g that are both invertible but $f \circ g$ is not invertible (such an example is called a *counter example* because it would provide an example of something contradicting the claim).

Exercise 8. Let $f : \mathbb{R} \to \mathbb{R}$ be an invertible function. Explain why the graph of f^{-1} is the reflection of the graph of f through the graph of the identity function (this is the line $\{(x, x) \ : \ x \in \mathbb{R}\} \subseteq \mathbb{R} \times \mathbb{R}$). Your explanation should be *independent* of the function f you choose!

Exercise 9. Let $D \xleftarrow{f} C \xleftarrow{g} B \xleftarrow{h} A$ be a triple of composable functions between sets. The composition, as defined in these notes, has only been defined for a *pair* of functions. Therefore, one can first form the composition $D \xleftarrow{f \circ g} B \xleftarrow{h} A$ and then $D \xleftarrow{(f \circ g) \circ h} A$ to obtain a function from A to D. Another option is to first form the composition $D \xleftarrow{f} C \xleftarrow{g \circ h} A$ and then $D \xleftarrow{f \circ (g \circ h)} A$ to get another, a-priori different, function from A to D. Are these two functions the same, i.e. does $f \circ (g \circ h) = (f \circ g) \circ h$ for *all* triples of composable functions? When this is true, the composition is said to be *associative*.

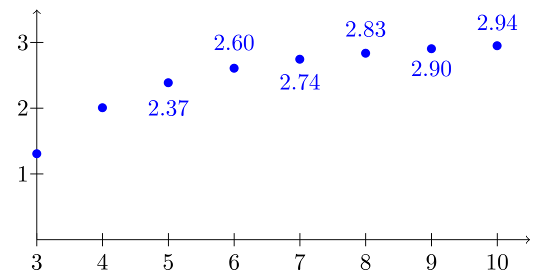
Exercise 10. Prove that $\sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x}$ for all $x \in \mathbb{R} \setminus \{1\}$. [Hint: compute $(1 + x + x^2 + \cdots + x^n)(1 - x)$.] What's wrong with this formula when $x = 1$?

Exercise 11. Let $a, b, c \in \mathbb{R}$ with $a \neq 0$. Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by $\mathbb{R} \ni x \mapsto ax^2 + bx + c$. Find the largest subsets $A, B \subseteq \mathbb{R}$ such that the functions $g : \left(-\infty, -\frac{b}{2a}\right) \to A$ and $h : \left(-\frac{b}{2a}, \infty\right) \to B$ defined by $g(x) := f(x)$ and $h(x) := f(x)$ on their respective domains are invertible. Find g^{-1} and h^{-1}. [Hint: $-\frac{b}{2a}$ is the inflection point for the function f—plot the case $a = 1, b = -2, c = 1$.]

Sequences

A *sequence* of real numbers is a function $\mathbb{N} \to \mathbb{R}$. More generally, if A is any set, a *sequence* in A is a function $B \to A$ where $B \subseteq \mathbb{N}$ is a subset of natural numbers that is not finite. [Reference: Herb Gross VII.1 Many Versus Infinite]

For any natural number $n \geq 3$ the area of a regular n-sided polygon whose distance from the center to any vertex is $r \in (0, \infty)$ is given by $a_n := \frac{nr^2}{2} \sin\left(\frac{2\pi}{n}\right)$. Then $\mathbb{N} \setminus \{1, 2\} \ni n \mapsto a_n := \frac{nr^2}{2} \sin\left(\frac{2\pi}{n}\right)$ defines a sequence of real numbers. When r is a specified value, such as $r = 1$, we can plot the graph of (a part of) this sequence.



These numbers seem to tend to a particular value. Can you guess what it is? [Hint: compute a_{1000}.]

Let $x \in \mathbb{R}$. Then $\mathbb{N} \ni n \mapsto x^n$ is a sequence of real numbers. What is the graph of this sequence when $x > 1$? What happens when $x = 1$? What happens when $x \in (0, 1)$?

A special type of sequence occurs when summing numbers. Fix $x \in \mathbb{R}$. In an earlier exercise, it was claimed that $\sum_{k=0}^{n} x^k = \frac{1-x^n}{1-x}$. This gives a number for each $n \in \mathbb{N}$. In other words, $\mathbb{N} \ni n \mapsto \sum_{k=0}^{n} x^k$ defines a sequence. Such a sequence of sums of numbers is called a *partial sum*. The limit of a partial sum as n tends to infinity is called an *infinite series*, something you will learn more about in the second semester of calculus. This particular series is called the *geometric series*.

The definition of a sequence that tends to a limit is as follows. A sequence $\mathbb{N} \ni n \mapsto a_n \in \mathbb{R}$ *converges* to a number, $a \in \mathbb{R}$, iff for any real number $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $|a - a_n| < \epsilon$ for all $n \geq N$. The number a is called the *limit* of the sequence and is denoted by $a = \lim\limits_{n\to\infty} a_n$. The ϵ is a measure our precision. This definition says that eventually the terms in the sequence are so close to a particular number that they are at most within an ϵ-distance from this number. A sequence that converges to some real number is said to be *convergent* and a sequence that does not converge to any real number is said to be a *divergent* sequence. Note that if a sequence converges, the number it converges to is unique.

The sequence $\mathbb{N} \ni n \mapsto x^n$ is convergent if $x \in (-1, 1]$. Furthermore, it converges to 0 when $x \in (-1, 1)$ and converges to 1 when $x = 1$.

Sketching out an idea for a proof. Fix $\epsilon > 0$. The goal is to show there exists an $N \in \mathbb{N}$ such that $|x^n| \leq \epsilon$ for all $n \geq N$. This is automatically true for $x = 0$ for all $n \geq 1$. Therefore, suppose that $x \neq 0$. The desired inequality is equivalent to $|x|^n \leq \epsilon$. Taking the logarithm (base 10, say) gives $n \log(|x|) \leq \log(\epsilon)$. Since $|x| < 10$, $\log(|x|)$ is negative so dividing by it reverses the inequality to $n \geq \frac{\log(\epsilon)}{\log(|x|)} = \log_{|x|}(\epsilon)$, the last equality holding since $x \neq 0$. Therefore, pick N to be any integer greater than $\log_{|x|}(\epsilon)$. ■

Proof. Fix $\epsilon > 0$ and suppose that $x \neq 0$ (the case $x = 0$ is true since $|x|^n = 0 < \epsilon$ for all $n \geq 1$). Set N to be any integer such that $N > \log_{|x|}(\epsilon)$. Then $|x^n| = |x|^n \leq |x|^{\log_{|x|}(\epsilon)} = \epsilon$ for all $n \geq N$. ■

To see how big N can be, consider the special case where $x = 1/2$ and $\epsilon = 1/1000$. Then $\log_{1/2}(1/1000) \approx 9.97$ so that $N = 10$ works.

The *algebraic limit theorem* for sequences is useful when proving that certain sequences converge.

Theorem 1. *Let $a, b : \mathbb{N} \to \mathbb{R}$ be two sequences (the value of a at n is denoted by a_n instead of $a(n)$). Then*

(a) $\lim\limits_{n\to\infty} (c a_n) = x \lim\limits_{n\to\infty} a_n$ *for all* $c \in \mathbb{R}$,

(b) $\lim\limits_{n\to\infty} (a_n + b_n) = \lim\limits_{n\to\infty} a_n + \lim\limits_{n\to\infty} b_n$,

(c) $\lim\limits_{n\to\infty} (a_n b_n) = \left(\lim\limits_{n\to\infty} a_n\right)\left(\lim\limits_{n\to\infty} b_n\right)$, *and*

(d) $\lim\limits_{n\to\infty} \left(\frac{a_n}{b_n}\right) = \frac{\lim\limits_{n\to\infty} a_n}{\lim\limits_{n\to\infty} b_n}$, *provided that* $\lim\limits_{n\to\infty} b_n \neq 0$.

If $b_n = 0$ for some n in part (d), the meaning of $\frac{a_n}{b_n}$ is meant only for sufficiently large n, where $b_n \neq 0$.

Suppose that $x \in (-1, 1)$. Then $\lim\limits_{n\to\infty} \sum_{k=0}^{n} x^k = \lim\limits_{n\to\infty} \frac{1 - x^n}{1 - x} = \frac{1}{1-x} \lim\limits_{n\to\infty} (1 - x^n) = \frac{1}{1-x}\left(1 - \lim\limits_{n\to\infty} x^n\right) = \frac{1}{1-x}$ by the previous observation, part (a), and part (b) of the algebraic limit theorem.

A sequence $a : \mathbb{N} \to \mathbb{R}$ is *bounded* if there exists a real number $M > 0$ such that $|a_n| \leq M$ for all $n \in \mathbb{N}$ A sequence $a : \mathbb{N} \to \mathbb{R}$ is *non-decreasing* if $a_{n+1} \geq a_n$ for all $n \in \mathbb{N}$ and *non-increasing* if $a_{n+1} \leq a_n$ for all $n \in \mathbb{N}$. A sequence $a : \mathbb{N} \to \mathbb{R}$ is *monotone* if it is either non-decreasing or non-increasing. The *monotone-convergence theorem* for sequences is incredibly helpful in identifying convergent sequences.

Theorem 2. *If a sequence $a : \mathbb{N} \to \mathbb{R}$ is monotone and bounded, then it converges.*

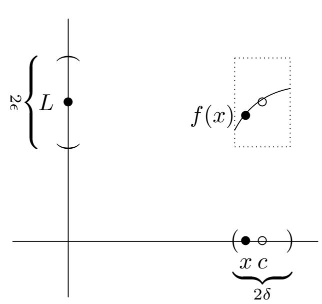
The sequence $\mathbb{N} \ni n \mapsto \sum_{k=0}^{n} \frac{1}{k!}$ converges (here, $k! := 1 \cdot 2 \cdot 3 \cdots k$) to a positive number e called *Euler's constant*.

Proof. $\sum_{k=0}^{n} \frac{1}{k!} = 1 + \frac{1}{1} + \frac{1}{1\cdot 2} + \frac{1}{1\cdot 2\cdot 3} + \frac{1}{1\cdot 2\cdot 3\cdot 4} + \cdots + \frac{1}{1\cdot 2\cdots n}$
$< 1 + \frac{1}{2^0} + \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots + \frac{1}{2^{n-1}}$
$< 1 + \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k$ since $\sum_{k=0}^{n} \left(\frac{1}{2}\right)^k < \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k$
$= 1 + 2$ since this is the value of the geometric series
$= 3$ showing that the sequence is bounded. By the monotone convergence theorem, $\mathbb{N} \ni n \mapsto \sum_{k=0}^{n} \frac{1}{k!}$ converges. ■

Limits of functions

The limit of a sequence $\mathbb{N} \to \mathbb{R}$ is obtained by observing what happens to an infinite ordered list of numbers $(x_1, x_2, x_3, x_4, \ldots)$ as you progress forward in this list (the numbers x_k need not increase or decrease as k increases!). The limit of a sequence is a number that the sequence tends to. The limit of a function is similar, but not quite the same. First, the limit of a function $f : \mathbb{R} \to \mathbb{R}$ is obtained at every point in its domain so saying "the limit of f" alone is meaningless. Instead, one must say "the limit of $f(x)$ as x approaches c" or "the limit of f at c," where c is in the domain of f (we will assume the domain always contains an open interval of points around c). Second, the limit can be defined in more than one way. One can approach a from above (often called "the right") or from below (often called "the left"). If the two limits agree, then the limit can be defined unambiguously. [Reference: Herb Gross I.4. Derivatives and Limits]

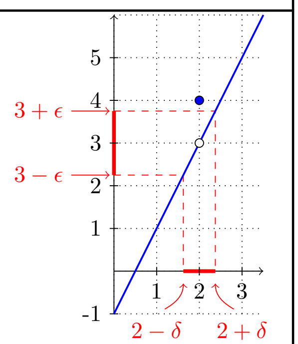
Let $a, b, c \in \mathbb{R}$ with $a < c < b$ and let $f : (a, b) \to \mathbb{R}$ (or $f : (a, c) \cup (c, b) \to \mathbb{R}$). L is said to be the *limit of $f(x)$ as x approaches c*, written $\lim\limits_{x\to c} f(x) = L$, iff for any $\epsilon > 0$, there exists a $\delta > 0$ such that $0 < |x - c| < \delta$ implies $|f(x) - L| < \epsilon$. L is called the *functional limit* of f as x approaches c. One can also define limits from below (the left) $\lim\limits_{x\nearrow c} f(x) = L$, and above (the right) $\lim\limits_{x\searrow c} f(x) = L$, by using "$c - \delta < x < c$ implies $|f(x) - L| < \epsilon$" and "$c < x < c + \delta$ implies $|f(x) - L| < \epsilon$," respectively. **Warning**: The limit of f at c is independent of the value of f at c, even if it is defined.



Rather than memorizing the definition of the limit of a function as its input variable approaches a number in the domain of that function, a visualization in terms of the *graph* of the function is helpful. In fact, it is helpful to visualize computations of limits by using graphs. Notice that the limit of a function is independent of whether or not that function is defined at the limiting point. The limit of f at c is also independent of the value of the function away from a neighborhood of c. When we discuss *continuity* of a function, we will see a relationship between the limit and the value of the function.

Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by

$$\mathbb{R} \ni x \mapsto \begin{cases} 4 & \text{if } x = 2 \\ \frac{2x^2-5x+2}{x-2} & \text{otherwise} \end{cases}.$$



Then $f(x) = \frac{2x^2-5x+2}{x-2} = \frac{(2x-1)(x-2)}{x-2} = 2x - 1$ for all $x \in \mathbb{R} \setminus \{2\}$. Hence, $\lim_{x\to 2} f(x) = 2(2) - 1 = 3$. Notice that $f(2) \neq \lim_{x\to 2} f(x)$ in this example. But why is $\lim_{x\to 2}(2x-1) = 3$? Fix $\epsilon > 0$. The goal is to find a $\delta > 0$ such that $0 < |x-2| < \delta$ implies $|(2x-1)-3| < \epsilon$. The latter inequality is equivalent to $-\epsilon < 2x - 4 < \epsilon$. Adding 4 to all parts of this inequality gives $4 - \epsilon < 2x < 4 + \epsilon$. Dividing by 2 gives $2 + \frac{\epsilon}{2} < x < 2 + \frac{\epsilon}{2}$. Hence, we can set $\delta := \frac{\epsilon}{2}$. In this case, the distance to the left and to the right of 2 are the same. More generally, one must choose the smaller of the two widths. [Reference: Herb Gross I.4. Derivatives and Limits—understand his example!]

Just as there was an algebraic limit theorem for limits of sequences, there is an <u>algebraic limit theorem</u> for functional limits. [Reference: Herb Gross I.5. A More Rigorous Approach to Limits]

Theorem 3. *Let $a, b, c \in \mathbb{R}$ with $a < c < b$ and let $f, g : (a, b) \to \mathbb{R}$. Furthermore, suppose that $\lim_{x\to c} f(x) = L$ and $\lim_{x\to c} g(x) = M$. Then the following facts hold.*

(a) $\lim_{x\to c} \big[(kf)(x)\big] = kL$ *for all* $k \in \mathbb{R}$.

(b) $\lim_{x\to c} \big[(f+g)(x)\big] = L + M$.

(c) $\lim_{x\to c} \big[(fg)(x)\big] = LM$.

(d) Suppose g is nonzero in a neighborhood of c. Then $\lim_{x\to c} \left[\left(\frac{f}{g}\right)(x)\right] = \frac{L}{M}$, *provided that* $M \neq 0$.

[Herb Gross gives wonderful intuition behind assumptions for theorems and their proofs at 27:05 of I.5.]

Theorem 4. *Let $a, b, c \in \mathbb{R}$ with $a < c < b$ and let $f, g, h : (a, b) \to \mathbb{R}$ be functions satisfying $f(x) \leq g(x) \leq h(x)$ for all $x \in (a, c) \cup (c, b)$. If $\lim_{x\to c} f(x) = \lim_{x\to c} h(x) =: v$, then $\lim_{x\to c} g(x) = v$.*

This is called the <u>squeezing principle</u>. Exercise 10 explores a useful application of this theorem.

One can also compute limits "at infinity" such as $\lim_{x\to\infty} \left(\frac{8x^2-2x+1}{(2x+2)^2}\right) = \lim_{x\to\infty} \left(\frac{8-2/x+1/x^2}{4+8/x+4/x^2}\right) = 2$. In this case, 2 is called a <u>horizontal asymptote</u>. As another example, $\lim_{x\to\infty} \arctan(x) = 1$ while $\lim_{x\to-\infty} \arctan(x) = -1$. A <u>vertical asymptote</u> occurs when a function becomes unbounded on a finite interval. For example, $\lim_{x\nearrow 1} \frac{1}{1-x} = \infty$ while $\lim_{x\searrow 1} \frac{1}{1-x} = -\infty$.

Exercises

Exercise 1. Graph the sequences below and verify the claims.

i. $\mathbb{N} \ni n \mapsto x^n$ is convergent if and only if $x \in (-1, 1]$. Furthermore, it converges to 0 when $x \in (-1, 1)$ and converges to 1 when $x = 1$. In particular, the sequence $\mathbb{N} \ni n \mapsto (-1)^n$ is divergent.

ii. $\mathbb{N} \ni n \mapsto \sin(n)$ is divergent.

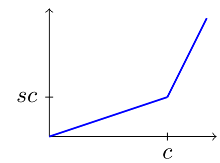
iii. $\mathbb{N} \ni n \mapsto \frac{1}{\sqrt{n}}$ converges to 0.

iv. $\mathbb{N} \ni n \mapsto \sin\left(\frac{1}{n}\right)$ converges to 0.

v. $\mathbb{N} \ni n \mapsto n\sin\left(\frac{1}{n}\right)$ converges to 1 (then use this to prove $\lim_{n\to\infty} \left(\frac{nr^2}{2} \sin\left(\frac{2\pi}{n}\right)\right) = \pi$).

Exercise 2. Fix $c \in \mathbb{R}$ and $s, r > 0$. Let $f : \mathbb{R} \to \mathbb{R}$ be the function

$$\mathbb{R} \ni x \mapsto \begin{cases} sx & \text{if } x \leq c \\ rx + c(s-r) & \text{if } x > c \end{cases}$$



Fix $\epsilon > 0$. Find $\delta > 0$ such that $0 < |x - c| < \delta$ implies $|f(x) - sc| < \epsilon$. Include a graph and visualization of ϵ and δ with your work.

Exercise 3. Using the ϵ, δ definition of the limit, prove $\lim_{x\to\frac{\pi}{2}} \cos(x) = 0$. To do this, first fix $\epsilon > 0$. Then write down what needs to be shown, namely, there exists a δ such that what? Using the definition of the absolute value and the required inequality, derive an expression similar to your ϵ-inequality. Finally, compute δ. Draw an accompanying diagram to help you.

Exercise 4. Fix $a, c \in \mathbb{R}$. Using the ϵ, δ definition of the limit, prove $\lim_{x\to c} a = a$ and prove $\lim_{x\to c} x = c$.

Exercise 5. Using the ϵ, δ definition of the limit, prove $\lim_{x\to 1} \frac{1}{x} = 1$.

Exercise 6. Using the algebraic limit theorem for functional limits and the previous exercise, prove the following limits are valid.

i. $\lim_{x\to 1} \frac{1}{x} = 1$. ii. $\lim_{x\to 2} x^3 = 8$. iii. $\lim_{x\to 1} \frac{3x^2-4x+1}{x^4+1} = 0$.

Exercise 7. If possible, give examples of functions $f, g : \mathbb{R} \to \mathbb{R}$ and a point $c \in \mathbb{R}$ such that $\lim_{x\to c} f(x) = P$, $\lim_{x\to c} g(x) = Q$, and $\lim_{x\to c} \big(f(x)g(x)\big) = R$, where $P, Q, R \in \{\text{exists}, \text{does not exist}\}$. If it is not possible for a given choice of P, Q, and R, explain why. [Hint: There are four different possibilities (excluding symmetry by replacing $f \leftrightarrow g$).] Do the same for $f + g$.

Exercise 8. Let $f, g : \mathbb{R} \to \mathbb{R}$ be functions and $a \in \mathbb{R}$ with $\lim_{x\to a} f(x) = b$ and $\lim_{y\to b} g(y) = c$. Is it true that $\lim_{x\to a} \big(g(f(x))\big) = c$? Explain.

Exercise 9. Evaluate the following limits.

i. $\lim_{x\nearrow 2} \frac{|x-2|}{x-2}$, $\lim_{x\searrow 2} \frac{|x-2|}{x-2}$, and $\lim_{x\to 2} \frac{|x-2|}{x-2}$.

ii. $\lim_{x\to 0} \frac{\sqrt{1+x}-\sqrt{1-x}}{x}$.

iii. $\lim_{x\nearrow 3} \left(\frac{1}{x-3} - \frac{1}{|x-3|}\right)$ and $\lim_{x\searrow 3} \left(\frac{1}{x-3} - \frac{1}{|x-3|}\right)$.

iv. $\lim_{x\to 1} \left(\frac{(x+7)^{1/3}+2}{x+1}\right)$.

v. $\lim_{x\to 3} \left(\frac{x-3}{2x^2-7x+3}\right)$.

vi. $\lim_{x\to\infty} \frac{1+\sin(x)+|x|}{x}$ and $\lim_{x\to-\infty} \frac{1+\sin(x)+|x|}{x}$.

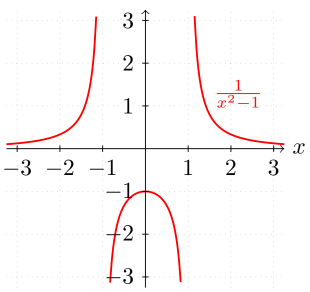
Exercise 10. Using the squeezing principle, prove $\lim_{x\to 0} \frac{\sin(x)}{x} = 1$. Hint: use the inequality $0 < \cos(x) < \frac{\sin(x)}{x} < \frac{1}{\cos(x)}$ valid for all $x \in \left(-\frac{\pi}{2}, 0\right) \cup \left(0, \frac{\pi}{2}\right)$.

Continuity

One of the main assumptions about calculus, as formulated by Newton, Leibniz, and others, is that there are quantities that one can break down into smaller pieces indefinitely, i.e. infinitesimally. Whether this is actually possible in our word has been debated over millennia. Based on our understanding of quantum mechanics, time, and space, this seems like an invalid assumption. Nevertheless, it holds to a good enough accuracy for many practical purposes.

Let $a, b, c \in \mathbb{R}$ with $a < c < b$ and let $f : (a, b) \to \mathbb{R}$ be a function. f is *continuous at c* iff for every $\epsilon > 0$, there exists a $\delta > 0$ such that $f(x) \in \big(f(c) - \epsilon, f(c) + \epsilon\big)$ for all $x \in (c - \delta, c + \delta) \cap (a, b)$, i.e. iff $f(c) = \lim_{x\to c} f(x)$. f is *discontinuous at c* iff f is not continuous at c. f is *continuous on (a, b)* iff f is continuous at c for all $c \in (a, b)$. [Reference: Herb Gross II.6 Continuity]

A real *polynomial* is a function $p : \mathbb{R} \to \mathbb{R}$ of the form $\mathbb{R} \ni x \mapsto p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$, where $a_0, a_1, a_2, \ldots, a_m$ are real numbers and where $m \in \mathbb{N}$. The largest m for which the coefficient in front of x^m is not zero is called the *degree* of the polynomial. A degree zero polynomial is a constant, a degree one polynomial is a linear function, and a degree two polynomial is a quadratic function. For example, $\mathbb{R} \ni x \mapsto p(x) = 5x^4 - 3x^3 + 2x - 4$ is a degree 4 polynomial. As another example, and to set some notation, the polynomial p_m will denote $\mathbb{R} \ni x \mapsto p_m(x) := x^m$ and is called *the standard homogeneous degree m polynomial*. The *roots* of a polynomial p are the set of numbers $\mathcal{R}(p) := \{z \in \mathbb{R} : p(z) = 0\}$.



A *rational function* r is the ratio (aka quotient) of two polynomials, i.e. there exist polynomials p and q such that the domain of r is $\mathbb{R} \setminus \mathcal{R}(q)$ and $r(x) = \frac{p(x)}{q(x)}$ for all $x \in \mathbb{R} \setminus \mathcal{R}(q)$. For example, $r(x) = \frac{1}{x^2-1}$ defines a rational function r and its domain is $\mathbb{R}\setminus\{-1, 1\} = (-\infty, -1) \cup (-1, 1) \cup (1, \infty)$.

Other common functions include trigonometric functions such as cos and sin, the exponential function exp, and the logarithm log. Most common functions are compositions of all of the previous ones. For example, the formula $x \mapsto e^{-x^2}$ is obtained from the composition of the functions $x \xmapsto{-p_2} -x^2 \xmapsto{\exp} e^{-x^2}$. As another more complicated example, the formula $x \mapsto \cot(2\pi x)$ is the composition $\mathbb{R} \setminus \mathbb{Z} \ni x \xmapsto{\pi p_1} \pi x \xmapsto{(\cos,\sin)} \big(\cos(\pi x), \sin(\pi x)\big) \xmapsto{\text{divide}} \frac{\cos(\pi x)}{\sin(\pi x)} = \cot(\pi x)$. Here, $(\cos, \sin)$ denotes the function whose domain is $\mathbb{R}$ and whose codomain is $\mathbb{R} \times \mathbb{R}$ and whose value at $x \in \mathbb{R}$ is $(\cos(x), \sin(x))$.

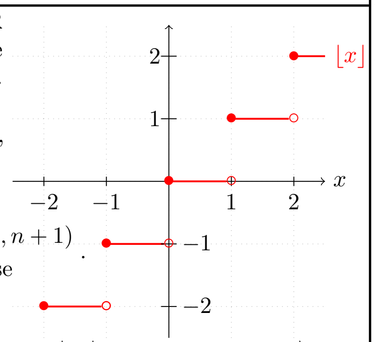
Theorem 1. *Polynomials, rational functions, trigonometric functions, root functions, exponentials, and logarithms are continuous on their domain of definition.*

The floor function $\lfloor\,\cdot\,\rfloor : \mathbb{R} \to \mathbb{R}$ is defined by $\mathbb{R} \ni x \mapsto \lfloor x \rfloor$ is the largest integer n such that $n \le x$.

$$\mathbb{R} \ni x \xmapsto{\lfloor\,\cdot\,\rfloor} \sum_{n=-\infty}^{+\infty} \Theta\big([n, n+1)\big)(x),$$

where

$$\Theta\big([n, n+1)\big)(x) := \begin{cases} n & \text{if } x \in [n, n+1) \\ 0 & \text{otherwise} \end{cases}.$$



$\lfloor\,\cdot\,\rfloor$ is discontinuous on $\mathbb{Z} \subseteq \mathbb{R}$ and is continuous everywhere else, i.e. $\lfloor\,\cdot\,\rfloor$ is continuous on $\mathbb{R} \setminus \mathbb{Z}$.

Theorem 2. *Let f be a function continuous at $b \in \mathbb{R}$ and let g be a function with $\lim_{x\to a} g(x) = b$, then $\lim_{x\to a} f(g(x)) = f(b)$, i.e. $\lim_{x\to a} f(g(x)) = f\left(\lim_{x\to a} g(x)\right)$.*

This is one of many theorems in calculus where the order of certain operations does not matter. Here the operations are limits and applying functions. Note, however, that the assumptions are important.

Let $f, g : \mathbb{R} \to \mathbb{R}$ be defined by $\mathbb{R} \ni x \mapsto \sqrt{|x|}$ and $\mathbb{R} \ni x \mapsto g(x) := \begin{cases} 1 & \text{if } x = 0 \\ x^2 & \text{otherwise} \end{cases}$. Then $\lim_{x\to 0} g(x) = 0$ and f is continuous everywhere. The composition of g followed by f is given by $\mathbb{R} \ni x \mapsto f(g(x)) = \begin{cases} 1 & \text{if } x = 0 \\ |x| & \text{otherwise} \end{cases}$. Hence, $\lim_{x\to 0} f(g(x)) = 0$ and $f\left(\lim_{x\to 0} g(x)\right) = f(0) = 0$. This example shows that it is not necessary for g to be continuous at a. What is required, however, is that it must have a well-defined limit at a.

Theorem 3. *If g is continuous at c and f is continuous at $g(c)$, then $f \circ g$ is continuous at c.*

We can use this theorem to prove that functions such as the ones given before Theorem 1 are continuous (after we have proven that the basic functions are continuous). For example, if we show exp is continuous and any polynomial p is continuous, then $\exp \circ p \equiv e^p$ is continuous.

The largest subset $A \subseteq \mathbb{R}$ for which $A \ni x \mapsto \ln\big(1 + \cos(x)\big)$ is continuous is $A = \bigcup_{n\in\mathbb{Z}} \big((2n-1)\pi, (2n+1)\pi\big)$.

Theorem 4. *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then for any L satisfying either $f(a) < L < f(b)$ or $f(a) > L > f(b)$, there exists a $c \in (a, b)$ with $f(c) = L$.*

This is called the *Intermediate Value Theorem* (IVT) and can be used to prove that continuous functions must take on certain values.

The polynomial $p : \mathbb{R} \ni x \mapsto x^4 + 3x^3 - 3x^2 - 7x + 3$ has at least four (real) roots.

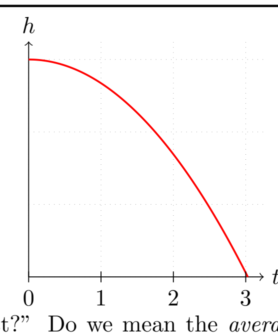
Proof. Note that $p(-4), p(0), p(2) > 0$ and $p(-3), p(1) < 0$. Since p is continuous by Theorem 1, the IVT applies and there are roots in the intervals $(-4, -3)$, $(-3, 0)$, $(0, 1)$, and $(1, 2)$. ■

There exists a $z \in \mathbb{R}$ such that $\arctan(z) = 1 - z$.

Proof. Consider the function $f : \mathbb{R} \ni x \mapsto \arctan(x) - 1 + x$. f is continuous on its domain and $f(0) = -1 < 0$ while $f(1) = \frac{\pi}{4} > 0$. By the IVT, there exists a $z \in (0, 1)$ such that $\arctan(z) - 1 + z = 0$. ■

Rates of change and derivatives

When dropping a bowling ball off of a medium-sized cliff, the vertical position h (in meters) of the ball as a function of time t (in seconds) is given roughly by $h(t) = -9.8t^2 + H$, where H is the initial height at which the ball is dropped. The ball appears to move slowly initially and it picks up speed the farther it falls from the top of the cliff. How fast is it going after t seconds?



In fact, what do we mean by "how fast?" Do we mean the *average* speed of the ball? This would not distinguish the earlier slower speed or the later faster speed. What we are looking for is some average speed over a small interval of time, such as $[t_0, t_1]$ for some $t_0, t_1 \in \mathbb{R}$ with $t_1 > t_0$. The *average velocity* over this time period is defined to be $\frac{h(t_1) - h(t_0)}{t_1 - t_0}$. As this interval gets smaller, we expect a better approximation to the *instantaneous velocity*, which is defined as the limit of average velocities. More precisely, the (instantaneous) velocity at $t \in \mathbb{R}$ is given by $v(t) := \lim_{\epsilon\to 0} \dfrac{h(t+\epsilon) - h(t-\epsilon)}{2\epsilon}$ if it exists.

Setting $g := 9.81$, the limit of this quotient is given by $\lim_{\epsilon\to 0}\frac{h(t+\epsilon)-h(t-\epsilon)}{2\epsilon} = \lim_{\epsilon\to 0}\frac{-g(t+\epsilon)^2+H+g(t-\epsilon)^2-H}{2\epsilon} = \lim_{\epsilon\to 0}\frac{-4gt\epsilon}{2\epsilon} = -2gt$. This result has the property when t is small, the velocity is small. Conversely, when t increases, the speed increases. This is consistent with our observation that the ball picks up speed as it falls. Notice this limit is the same as $\lim_{\epsilon\to 0}\frac{h(t+\epsilon)-h(t)}{\epsilon}$. Since the latter is more often used as the definition of the derivative, we will use this one as well. [Reference: Herb Gross I.4 Derivatives and Limits (first 16 minutes and 30 seconds)] Let $f:(a,b)\to\mathbb{R}$ be a function with $b>a$. The *derivative of f at $c\in(a,b)$*, if it exists, is the limit $\lim_{\epsilon\to 0}\frac{f(x+\epsilon)-f(x)}{\epsilon}$ and is denoted by $f'(c)$ or $(D_cf)(1)$, where $D_cf:\mathbb{R}\to\mathbb{R}$ is the linear function $\mathbb{R}\ni v\mapsto (D_cf)(v):=(f'(c))v$. If x is used as the variable input for the function f, $f'(c)$ might also be written as $\frac{df}{dx}(c)$ or $\left(\frac{d}{dx}f\right)(c)$. The *derivative of f* is the function $f':(a,b)\ni x\mapsto f'(x)$, provided that it exists for all $x\in(a,b)$. f is *differentiable at c* iff $f'(c)$ exists. f is *differentiable* on (a,b) iff f is differentiable at c for all $c\in(a,b)$. Higher derivatives are defined inductively: f is *n times differentiable* iff f' is $(n-1)$ times differentiable. The n-th derivative of f is denoted by $f^{(n)}$ or $\frac{d^nf}{dx^n}$. When n is small, such as $n=3$, the n-th derivative may also be written as f'''. [Reference: Herb Gross II.1 Derivatives of Some Simple Functions]

Let $f:(0,\infty)\to\mathbb{R}$ be the function defined by $(0,\infty)\ni x\mapsto f(x):=\frac{1}{x}$. Then $f'(x)=-\frac{1}{x^2}$ for all $x\in(0,\infty)$.

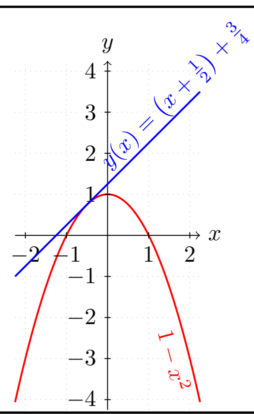
Proof. $\lim_{\epsilon\to 0}\frac{\frac{1}{x+\epsilon}-\frac{1}{x}}{\epsilon} = \lim_{\epsilon\to 0}\frac{x-(x+\epsilon)}{\epsilon(x+\epsilon)x} = \lim_{\epsilon\to 0}\frac{-1}{x(x+\epsilon)} = -\frac{1}{x^2}$. ∎

$\sin' = \cos$. [Reference: Herb Gross III.1 Circular Functions]

Proof. Since $\sin(y)-\sin(x) = 2\sin\left(\frac{y-x}{2}\right)\cos\left(\frac{y+x}{2}\right)$ the difference quotient becomes $\lim_{\epsilon\to 0}\frac{\sin(x+\epsilon)-\sin(x)}{\epsilon} = \lim_{\epsilon\to 0}\left(\frac{\sin(\epsilon)}{\epsilon}\cos\left(\frac{2x+\epsilon}{2}\right)\right) = \left(\lim_{\epsilon\to 0}\frac{\sin(\epsilon)}{\epsilon}\right)\lim_{\epsilon\to 0}\cos\left(x+\frac{\epsilon}{2}\right) = 1\cdot\cos(x)=\cos(x)$. In this calculation, we used a result from a previous exercise and continuity of $\cos$. ∎

The absolute value function $f := |\ \cdot\ |:\mathbb{R}\to\mathbb{R}$ is differentiable everywhere except at 0. Its derivative is broken up into two cases. If $x<0$, then $f'(x) = \lim_{\epsilon\to 0}\frac{|x+\epsilon|-|x|}{\epsilon} = \lim_{\epsilon\to 0}\frac{-x-\epsilon+x}{\epsilon} = -1$. If $x>0$, then $f'(x)=\lim_{\epsilon\to 0}\frac{|x+\epsilon|-|x|}{\epsilon} = \lim_{\epsilon\to 0}\frac{x+\epsilon-x}{\epsilon} = 1$. For $x=0$, the derivative does not exist because $\lim_{\epsilon\nearrow 0}\frac{|\epsilon|}{\epsilon} = \lim_{\epsilon\nearrow 0}\frac{-\epsilon}{\epsilon} = -1$ while $\lim_{\epsilon\searrow 0}\frac{|\epsilon|}{\epsilon} = \lim_{\epsilon\searrow 0}\frac{\epsilon}{\epsilon} = 1$.

The geometric meaning of the derivative of a function $f:\mathbb{R}\to\mathbb{R}$ is visualized by looking at the graph of f. The *tangent line* to (the graph of) f at the point $(z,f(z))\in\Gamma(f)$ is the straight line through $(z,f(z))$ with slope $f'(z)$, provided that $f'(z)$ exists. Explicitly, if the second coordinate is described by the variable y, this line is given by $y(x) = (f'(z))x + f(z) - (f'(z))z = (f'(z))(x-z)+f(z)$. For example, if $f(x)=1-x^2$, then $f'(x)=-2x$, $f(-\frac{1}{2})=\frac{3}{4}$, and $f'(-\frac{1}{2})=1$. Hence the tangent line to f is given by $y(x) = \left(x+\frac{1}{2}\right)+\frac{3}{4}$.



Theorem 5. *Let $f:(a,b)\to\mathbb{R}$ be a differentiable function with $f'(x)>0$ ($f'(x)<0$) for all $x\in(a,b)$. Then $f(y)>f(x)$ ($f(y)<f(x)$) for all $y,x\in(a,b)$ with $y>x$.*

[Reference: Herb Gross II.7 Curve Plotting]

Exercises

For these exercises, for each $n\in\mathbb{N}$, recall p_n denotes the homogeneous degree n polynomial given by $\mathbb{R}\ni x\mapsto p_n(x):=x^n$.

Exercise 1. Let $a,b,c\in\mathbb{R}$. Give an example of a function $f:\mathbb{R}\to\mathbb{R}$ with $\lim_{x\to b}f(x)=c$ and an example of a function $g:\mathbb{R}\to\mathbb{R}$ with $\lim_{x\to a}g(x)=b$ but for which $\lim_{x\to a}f(g(x))\ne f\left(\lim_{x\to a}g(x)\right)$.

Exercise 2. Plot p_n on the domain $(-1,1]$ for $n=1,2,\dots,10$. For each $x\in(-1,1]$, compute $\lim_{n\to\infty}p_n(x)$ (be careful about what limit you're taking—the answer depends on x). What happens if $x=-1$? What about $x\in(-\infty,-1)\cup(1,\infty)$? Hint: use Desmos, the online graphing calculator!

Exercise 3. What are the largest domains $A,B\subseteq\mathbb{R}$ such that the functions $A\ni x\mapsto\lfloor x^2\rfloor$ and $B\ni x\mapsto\tan(\sqrt{x})$ are continuous?

Exercise 4. What are the roots of the polynomial $p:\mathbb{R}\ni x\mapsto(x+3)(x+2)x(x-1)(x-2)$? Over what subset of $\mathbb{R}$ is p positive? Over what subset of $\mathbb{R}$ is p negative? Provide a rough sketch of the graph of p without using a calculator. If $q'=p$ for some $q:\mathbb{R}\to\mathbb{R}$, indicate the subsets over which q is increasing and the subsets over which q is decreasing. On the same plot, sketch a possible graph of q.

Exercise 5. Let $a,b,c\in\mathbb{R}$. From the limit definition of the derivative, prove the derivative of $\mathbb{R}\ni x\mapsto ax^2+bx+c$ is $\mathbb{R}\ni x\mapsto 2ax+b$.

Exercise 6. From the limit definition, prove that $p_3'=3p_2$ and $p_4'=4p_3$. Then, try to prove $p_n'=np_{n-1}$ for all $n\in\mathbb{N}$. Hint: obtain and prove a formula for factoring x^n-y^n into $(x-y)$ times something.

Exercise 7. Let $f:[0,\infty)\to\mathbb{R}$ be the n-th root function, i.e. $f(x)=x^{1/n}$. Prove $f'(x)=\frac{1}{n}x^{(1-n)/n}$ for all $x\in(0,\infty)$. Hint: in the difference quotient, set $u:=(x+\epsilon)^{1/n}$ and $v:=x^{1/n}$ and use the formula for factoring u^n-v^n into $(u-v)$ times something.

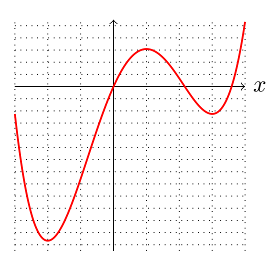
Exercise 8. Fix $n\in\mathbb{N}$. Let $g:(0,\infty)\to\mathbb{R}$ be the function $(0,\infty)\ni x\mapsto\frac{1}{x^n}$. Using the definition of the derivative, prove $g'(x)=-\frac{n}{x^{n+1}}$. Hint: use ideas similar to the previous two exercises.

Exercise 9. Following a similar method to the one in this handout, prove that $\cos'=-\sin$ from the definition of the derivative.

Exercise 10. Prove that the function $\mathbb{R}\ni x\mapsto\begin{cases}1 & \text{if } x=0\\ \frac{\sin(x)}{x} & \text{otherwise}\end{cases}$ is continuous. Compute its derivative for all points in its domain for which it is differentiable. Is it differentiable everywhere?

Exercise 11. Prove that $f:\mathbb{R}\ni x\mapsto\sqrt{|x|}$ is continuous. Prove that f is differentiable on $\mathbb{R}\setminus\{0\}$ and compute its derivative. What are $\lim_{x\nearrow 0}f'(x)$ and $\lim_{x\searrow 0}f'(x)$? Draw a graph of f on the domain $(-4,4)$, compute the equation for the tangent line to f at the point $(2,g(2))$, and draw this tangent line on the same plot.

Exercise 12. Let $f:\mathbb{R}\to\mathbb{R}$ be the function defined by $\mathbb{R}\ni x\mapsto\frac{1}{4}x^4-\frac{2}{3}x^3-\frac{5}{2}x^2+6x$. Indicate the domain over which f' is strictly positive. What is the domain over which f' is negative? Over what domain is $f'=0$? Compute $f'(x)$ for all $x\in(-3,4)$ and graph as much of this function as is possible on the grid to the right.



Exercise 13. Fix $a>0$ and $N\in\mathbb{N}$. Set $L:=Na$. For every ordered $(N+1)$-tuple of numbers $\vec{f}:=(f_0,f_1,f_2,\dots,f_N)$, define the *discrete derivative* of $\vec{f}$ to be the ordered $(N+1)$-tuple of numbers $(f_0',f_1',f_2',\dots,f_N')$ with $f_n':=\frac{f_{n+1}-f_{n-1}}{2a}$ (set $f_{-1}:=f_{N-1}$ and $f_{N+1}=f_0$). Setting N to be large (such as $N=100$) and a to be small (such as $a=0.1$), write a program that computes the discrete derivative given a list of numbers. To verify your program works, set $L:=2\pi$ and $N:=100$ and apply it to the list of numbers given by $f_n:=\sin\left(\frac{2\pi n}{100}\right)$.

Properties of differentiation

Theorem 1. *Let $f,g:\mathbb{R}\to\mathbb{R}$ be differentiable functions and let $c\in\mathbb{R}$. Then $f+g$ and cf are differentiable and their derivatives are given by $(f+g)'=f'+g'$ and $(cf)'=cf'$, respectively. In other words, differentiating is a linear operation.*

This result lets us compute derivatives of arbitrary polynomials given that we know $p_n'=np_{n-1}$ for all $n\in\mathbb{N}$. This is because every polynomial p is of the form $\mathbb{R}\ni x\mapsto p(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$ for some $n\in\mathbb{N}$ with $a_k\in\mathbb{R}$ for all $k\in\{0,1,2,\dots,n\}$. The derivative is $p'=(a_0+a_1p_1+a_2p_2+\cdots+a_np_n)'=a_0'+a_1p_1'+a_2p_2'+\cdots+a_np_n'=0+a_1+2a_2p_1+\cdots+na_np_{n-1}$.

Theorem 2. *Let $f,g:\mathbb{R}\to\mathbb{R}$ be differentiable functions. Then fg is differentiable and $(fg)'(x)=f'(x)g(x)+f(x)g'(x)$ for all $x\in\mathbb{R}$.*

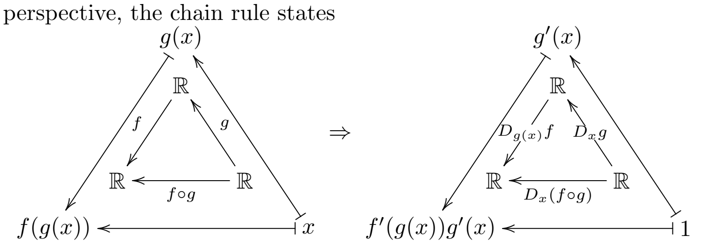
In this notation $fg:\mathbb{R}\to\mathbb{R}$ denotes the *product* of f and g (not the composition) and its value at $x\in\mathbb{R}$ is given by $(fg)(x):=f(x)g(x)$. This theorem is called the <u>product rule</u>.

$(p_3\sin)'=3p_2\sin+p_3\cos$, i.e. $\frac{d}{dx}\big(x^3\sin(x)\big)=3x^2\sin(x)+x^3\cos(x)$ for all $x\in\mathbb{R}$.

Given three differentiable functions $f,g,h:\mathbb{R}\to\mathbb{R}$ the derivative of the product is $(fgh)'=f'(gh)+f(gh)'=f'gh+fg'h+fgh'$.

Theorem 3. *Let $\mathbb{R}\xleftarrow{f}\mathbb{R}\xleftarrow{g}\mathbb{R}$ be two differentiable functions. Then $\mathbb{R}\xleftarrow{f\circ g}\mathbb{R}$ is differentiable and $(f\circ g)'(x)=f'\big(g(x)\big)g'(x)$ for all $x\in\mathbb{R}$.*

This theorem is called the <u>chain rule</u>. [Reference: Herb Gross: II.3 Composite Functions and the Chain Rule] This result is less mysterious if we use the alternative perspective on the derivative of f at x as a linear function $D_xf:\mathbb{R}\to\mathbb{R}$ whose value at 1 is $(D_xf)(1)=f'(x)$ and whose value at $v\in\mathbb{R}$ is $(D_xf)(v)=(f'(x))v$. This linear operator D_xf is just multiplication by the slope of function f at x. From this perspective, the chain rule states



i.e. $D_x(f\circ g)=(D_{g(x)}f)\circ(D_xg)$ for all $x\in\mathbb{R}$. In Leibniz notation, if y represents f as a function of u and u represents g as a function of x, the chain rule says $\frac{dy}{dx}=\frac{dy}{du}\frac{du}{dx}$ though care should be taken to make sure this is all expressed in terms of x.

$(\sin^3)'=(p_3\circ\sin)'=3(p_2\circ\sin)\cos$, i.e. $\frac{d}{dx}\sin^3(x)=3\sin^2(x)\cos(x)$ for all $x\in\mathbb{R}$. Notice the difference between $p_3\circ\sin$ and $p_3\sin$ above. As another example, $(p_3\circ(p_2+7))'=(3p_2\circ(p_2+7))2p_1=6p_1(p_2\circ(p_2+7))$, i.e. $\frac{d}{dx}(x^2+7)^3=3(x^2+7)^2 2x=6x(x^2+7)^2$ for all $x\in\mathbb{R}$.

Theorem 4. *Let $f,g:\mathbb{R}\to\mathbb{R}$ be differentiable functions with $g(x)\neq0$ for any $x\in\mathbb{R}$. Then $\frac{f}{g}$ is differentiable and its derivative is given by $\left(\frac{f}{g}\right)'(x)=\frac{f'(x)g(x)-g'(x)f(x)}{(g(x))^2}$ for all $x\in\mathbb{R}$.*

This theorem is called the <u>quotient rule</u>. The following proof is an illuminating application of the chain rule.

Proof. Let $h:\mathbb{R}\setminus\{0\}\to\mathbb{R}$ be given by $\mathbb{R}\setminus\{0\}\ni x\mapsto h(x):=\frac{1}{x}$. Since g is never zero, $\mathbb{R}\setminus\{0\}$ can be taken to be the target of g. Hence, the function $\frac{1}{g}:\mathbb{R}\ni x\mapsto 1/g(x)$ is the composition $\mathbb{R}\xleftarrow{h}\mathbb{R}\setminus\{0\}\xleftarrow{g}\mathbb{R}$. Using $h'(x)=-\frac{1}{x^2}$ and the chain rule, $\left(\frac{1}{g}\right)'(x)=(h\circ g)'(x)=g'(x)h'(g(x))=-\frac{g'(x)}{(g(x))^2}$. Finally, applying this result and the product rule gives $\left(\frac{f}{g}\right)'(x)=\left(f\frac{1}{g}\right)'(x)=f'(x)\frac{1}{g(x)}+f(x)\left(\frac{1}{g}\right)'(x)=\frac{f'(x)}{g(x)}-\frac{f(x)g'(x)}{(g(x))^2}=\frac{f'(x)g(x)-g'(x)f(x)}{(g(x))^2}$ for all $x\in\mathbb{R}$. ∎

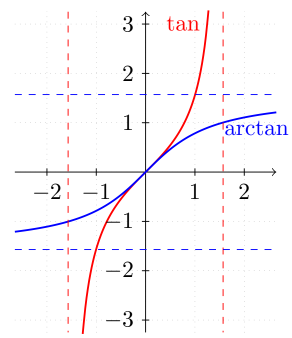
If $f:A\to B$ is one-to-one and onto, the <u>inverse</u> of f is the function $f^{-1}:B\to A$ such that $f^{-1}\circ f=\mathrm{id}_A$ and $f\circ f^{-1}=\mathrm{id}_B$. Do not conflate f^{-1} with the function $\frac{1}{f}$ (which might not even make sense if A and B are not subsets of real numbers).

Theorem 5. *Let $f:[a,b]\to\mathbb{R}$ be a differentiable one-to-one function with $f'(x)\neq0$ for all $x\in[a,b]$. Then $f([a,b])$ is a closed interval $[c,d]$ in $\mathbb{R}$ and $f^{-1}:[c,d]\to[a,b]$ is differentiable with derivative given by $(f^{-1})'(y)=\frac{1}{f'(f^{-1}(y))}$ for all $y\in[c,d]$.*

This is called the <u>inverse function theorem</u>. The fact that $f([a,b])$ is a closed interval only relies on the continuity of f. The hard part of the theorem is proving that f^{-1} is differentiable. Rather than proving this, let us take this for granted to derive the formula for $(f^{-1})'$. Since $f\circ f^{-1}=\mathrm{id}$, the chain rule gives $f'(f^{-1}(y))(f^{-1})'(y)=1$, which gives the desired result. [Reference: Herb Gross II.4 Differentiation of Inverse Functions]

$\mathbb{R}\ni x\mapsto p_3^{-1}(x)=x^{1/3}$ is the inverse of $p_3:\mathbb{R}\ni x\mapsto x^3$. Notice that the inverse function is defined on all of $\mathbb{R}$ but $p_3'(0)=0$ so the inverse function theorem does not apply at this point. Nevertheless, it applies everywhere else on $\mathbb{R}\setminus\{0\}$. Hence, the derivative of p_3^{-1} on $\mathbb{R}\setminus\{0\}$, using the inverse function theorem, is given by $(p_3^{-1})'(x)=\frac{1}{p_3'(p_3^{-1}(x))}=\frac{1}{3p_2(p_3^{-1}(x))}=\frac{1}{3(x^{1/3})^2}=\frac{1}{3}x^{-2/3}$ for all $x\in\mathbb{R}\setminus\{0\}$.

$\tan:\left(-\frac{\pi}{2},\frac{\pi}{2}\right)\to\mathbb{R}$ is an invertible differentiable function whose derivative vanishes nowhere on the domain $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ as it is given by $\tan'=\frac{1}{\cos^2}$ (see Exercise 6). Hence, the inverse function theorem applies. The inverse of tan is denoted by arctan and is a function with domain $\mathbb{R}$. Its derivative is therefore given by $\arctan'(y)=\frac{1}{\tan'(\arctan(y))}=\cos^2(\arctan(y))=\left(\frac{1}{\sqrt{1+y^2}}\right)^2=\frac{1}{1+y^2}$ for all $y\in\mathbb{R}$.



[Reference: Herb Gross III.2 Inverse Circular Functions]

Exercises

For these exercises, for each $n\in\mathbb{N}$, recall p_n denotes the homogeneous degree n polynomial given by $\mathbb{R}\ni x\mapsto p_n(x):=x^n$.

Exercise 1. Let $a,b\in\mathbb{R}$ with $a<b$. Prove that if a function $f:(a,b)\to\mathbb{R}$ is differentiable at $c\in(a,b)$, then f is continuous at c.

Exercise 2. Using the product rule, linearity of differentiation, and $p_1'=1$, prove that $p_n'=np_{n-1}$ for all $n\in\mathbb{N}$.

Exercise 3. Using the quotient rule and the result from Exercise 2, prove that $\mathbb{R}\setminus\{0\}\ni x\mapsto\frac{1}{x^n}$ has derivative $\mathbb{R}\setminus\{0\}\ni x\mapsto -n\frac{1}{x^{n+1}}$.

Exercise 4. Using the inverse function theorem and the result from Exercise 2, prove that $(0,\infty)\ni x\mapsto x^{1/n}$ has derivative $(0,\infty)\ni x\mapsto\frac{1}{n}x^{(1-n)/n}$.

Exercise 5. Using the chain rule and the results from Exercises 2 and 4, prove that if $m,n\in\mathbb{N}$, the function $f:(0,\infty)\ni x\mapsto x^{m/n}$ is differentiable and its derivative is given by $f'(x)=\frac{m}{n}x^{(m-n)/n}$ for all $x\in(0,\infty)$.

Exercise 6. Using the quotient rule, prove $\tan'=\frac{1}{\cos^2}$. Be sure to include the domain of $\tan'$.

Exercise 7. Draw a graph of $\cos:[0,\pi]\to[-1,1]$ to convince yourself that it is invertible. The inverse of cos is denoted by arccos and is a function $\arccos:[-1,1]\to[0,\pi]$. What is the largest subset $A\subseteq[0,\pi]$ such that $\cos:A\to[-1,1]$ satisfies the assumptions in the inverse function theorem? What is the domain $B\subseteq[-1,1]$ such that $\arccos:B\to A$ is differentiable? Compute the derivative of arccos over this domain.

The exponential and logarithm

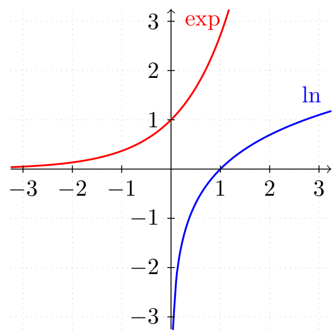
The *exponential function* $\exp : \mathbb{R} \to \mathbb{R}$ can be defined in several ways:

i. the power series $\exp(x) = \displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!}$ for all $x \in \mathbb{R}$

ii. the (unique) differentiable function satisfying $\exp' = \exp$ and $\exp(0) = 1$

iii. the limit $\exp(x) = \displaystyle\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$ for all $x \in \mathbb{R}$.

Proving either of these is actually well-defined, much less proving the equivalence between any of these definitions, is a little beyond our current capabilities and would require us to discuss series, integrals, and interchanging limits. Instead, we will take these for granted and point out that naively differentiating the power series definition shows that i. implies ii. Indeed, $\frac{d}{dx}\left(\sum_{k=0}^{\infty} \frac{x^k}{k!}\right) = \sum_{k=0}^{\infty} \frac{d}{dx}\left(\frac{x^k}{k!}\right) = \sum_{k=1}^{\infty} \frac{kx^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.



It is a fact that the exponential function is always positive and one-to-one. In fact, $\exp : \mathbb{R} \to (0,\infty)$ is a differentiable bijection whose derivative vanishes nowhere. Therefore, it has a differentiable inverse $\ln : (0,\infty) \to \mathbb{R}$ called the *(natural) logarithm* and its derivative, by the inverse function theorem, is given by $\ln'(y) = \frac{1}{\exp'(\ln(y))} = \frac{1}{\exp(\ln(y))} = \frac{1}{y}$ for all $y \in (0,\infty)$.

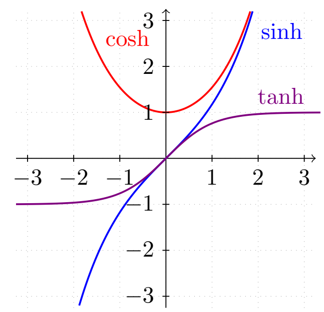
Warning: because exp is never zero, it has a *multiplicative* inverse, i.e. there is a function h such that $h\exp = 1$, where 1 denotes the constant function whose value is 1 for all values of $x \in \mathbb{R}$. Of course, this function is $\frac{1}{\exp}$ and is not to be confused with ln. For shorthand, we often write $e^x := \exp(x)$ and $e^{-x} := \frac{1}{\exp(x)}$.

The exponential and logarithm functions have many important properties. This includes $e^{x+y} = e^x e^y$ for all $x, y \in \mathbb{R}$. From this single crucial property, the usual laws for logarithms follow. For example, this equality implies $\ln(e^{x+y}) = \ln(e^x e^y)$ for all $x, y \in \mathbb{R}$. But $\ln(e^{x+y}) = x + y = \ln(e^x) + \ln(e^y)$ by definition of the logarithm as the inverse of exp. Hence, $\ln(e^x e^y) = \ln(e^x) + \ln(e^y)$ for all $x, y \in \mathbb{R}$. Since $\exp : \mathbb{R} \to (0,\infty)$ is a bijection, this shows that $\ln(ab) = \ln(a) + \ln(b)$ for all $a, b \in (0,\infty)$. By similar, but careful, analysis, one can also show $\ln(a^b) = b\ln(a)$ for all $a \in (0,\infty)$ and $b \in \mathbb{R}$.

In a previous exercise, you proved that $\frac{d}{dx}x^r = rx^{r-1}$ for any positive rational number r and for any $x \in (0,\infty)$. By using $\frac{d}{dx}\frac{1}{x} = \frac{-1}{x^2}$ and the chain rule, this can be extended to any rational number (exercise!). Although it might not seem to make sense to take the π-th power of another number, the functions $(0,\infty) \ni x \xmapsto{f} x^y \xleftarrow{g} y \in \mathbb{R}$ can be defined. Furthermore, f and g are both differentiable on their domains. As you can imagine, due to the asymmetry, these derivatives are probably different. Using logarithms and the chain rule, we can compute $f'(x)$ and $g'(y)$. Applying the logarithm to both sides gives $(\ln\circ f)(x) = y\ln(x) = (\ln\circ g)(y)$. Taking the derivatives with respect to x and y give two separate equations, namely $(\ln\circ f)'(x) = \frac{y}{x}$ and $(\ln\circ g)'(y) = \ln(x)$, respectively. Applying the chain rule to the first and second gives $f'(x) = yx^{y-1}$ and $g'(y) = \ln(x)x^y$, respectively. In short, $\frac{d}{dx}x^y = yx^{y-1}$ and $\frac{d}{dy}x^y = \ln(x)x^y$.

A crude model for the number of *E. coli* bacteria in a petri dish states that the quantity doubles every hour, i.e. is given by the function $N : [0,\infty) \ni t \mapsto N_0 2^t$ where $N_0 \in \mathbb{N}$ is the initial concentration and $t \in [0,\infty)$ is given in hours. The rate of growth of the bacteria is given by the derivative, which is $N'(t) = N_0 \ln(2)2^t$ for all $t \in [0,\infty)$.

Hyperbolic functions



The *hyperbolic sine and cosine* functions $\sinh, \cosh : \mathbb{R} \to \mathbb{R}$ are defined by $\sinh(x) := \frac{e^x - e^{-x}}{2}$ and $\cosh(x) := \frac{e^x + e^{-x}}{2}$ for all $x \in \mathbb{R}$. *Hyperbolic tangent* $\tanh : \mathbb{R} \to \mathbb{R}$ is defined by the ratio $\tanh := \frac{\sinh}{\cosh}$. The other hyperbolic functions can be defined analogously to how the trigonometric functions are defined as appropriate ratios. Although the graph may be misleading, there are no vertical asymptotes for $\sinh, \cosh, \tanh$.

A quick calculation shows that $\cosh' = \sinh$ and $\sinh' = \cosh$. Hence, $\tanh' = \left(\frac{\sinh}{\cosh}\right)' = \frac{\cosh^2 - \sinh^2}{\cosh^2}$ by the quotient rule. While $\cos^2 + \sin^2 = 1$, the hyperbolic version of this identity states (exercise!) $\cosh^2 - \sinh^2 = 1$. Therefore, $\tanh' = \frac{1}{\cosh^2}$. By plugging formulas into the definitions of the hyperbolic functions and using the properties of the logarithm and the exponential, one can show that the inverses of these functions are given by

i. $\sinh^{-1}(x) = \ln(x + \sqrt{x^2+1})$ for all $x \in \mathbb{R}$,

ii. $\cosh^{-1}(x) = \ln(x + \sqrt{x^2-1})$ for all $x \in [1,\infty)$,

iii. $\tanh^{-1}(x) = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right)$.

Hyperbolic functions can be used to describe the curve of a hanging rope or chain affected by a uniform gravitational field (see Week #01 handout).

Exercises

Exercise 1. A set of functions $\{f_1, f_2, \ldots, f_n\}$ is *linearly independent* iff the only numbers $a_1, a_2, \ldots, a_n \in \mathbb{R}$ satisfying $a_1f_1 + a_2f_2 + \cdots + a_nf_n = 0$ (the zero function) is $a_1 = a_2 = \cdots = a_n = 0$. Find a set of four linearly independent functions $\{f_1, f_2, f_3, f_4\}$ whose fourth derivative equal themselves, i.e. $f_k^{(4)} = f_k$ for all $k \in \{1,2,3,4\}$.

Exercise 2. Compute the derivatives of all six hyperbolic functions. Then compute their inverses by choosing suitable domains and prove that they are in fact the inverses. Compute the derivatives of the inverses. Graph all of your results. Be sure to indicate which theorems you are using.

Exercise 3. Find the domains of the following functions and their derivatives. Find explicit formulas for their derivatives indicating which theorems you use.

i. $x \mapsto \ln\left(\frac{1}{8\pi} + \cos^2(x)\right)$

ii. $x \mapsto \tan\left(\frac{\pi}{2}\tanh(x)\right)$

iii. $x \mapsto \tan\left(\frac{\pi-\epsilon}{2}\tanh(x)\right)$, where $0 < \epsilon < 1$

iv. $x \mapsto \tan\left(\frac{\pi+\epsilon}{2}\tanh(x)\right)$, where $0 < \epsilon < 1$

v. $x \mapsto \tanh(\tan(x))$

vi. $x \mapsto e^{\sin^3(x)}$

vii. $x \mapsto \ln(\cosh(x) - 1)$

viii. $x \mapsto (\sin(x))^{\cos(3)}$

ix. $x \mapsto x^{e^x}$

x. $x \mapsto \tanh^{-1}(2^{-3})$

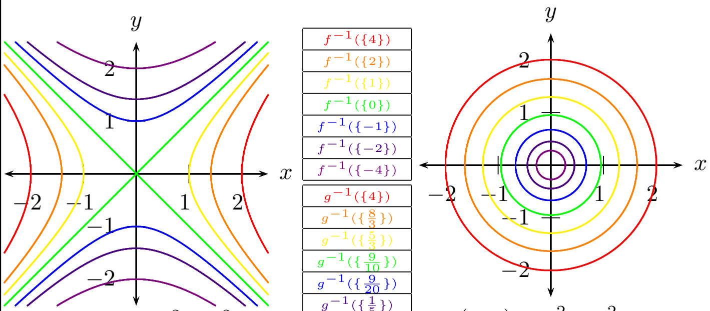
xi. $x \mapsto \tanh^{-1}(\cos(x))$

xii. $x \mapsto \tanh\left(\ln\left(\frac{1}{|x|}\right)\right)$

xiii. $x \mapsto 2e^{-x^2}$

Level sets and implicit functions

The *level set* of $f : \mathbb{R}^2 \to \mathbb{R}$ at $c \in \mathbb{R}$ is $f^{-1}(\{c\}) = \{(x, y) \in \mathbb{R}^2 : f(x, y) = c\}$. The notation $f^{-1}(\{c\})$ will be used instead of the more common $f^{-1}(c)$ because the latter conflicts with the notation used for the value of f^{-1}, the inverse of f, at the point c if it exists.



$f(x, y) = x^2 - y^2$ $\quad$ $g(x, y) = x^2 + y^2$

The *Implicit Function Theorem* (IFT) provides a way of expressing part of a level set of $f : \mathbb{R}^2 \to \mathbb{R}$ as the graph of a function. In what follows, let $(a, b) \in \mathbb{R}^2$ and let $\partial_1 f(a, b) := \frac{d}{dx} f(x, b)\Big|_{x=a}$ denote the *partial derivative* of f with respect to the first variable evaluated at (a, b). Similarly, let $\partial_2 f(a, b) := \frac{d}{dy} f(a, y)\Big|_{y=b}$ denote the partial derivative of f with respect to the second variable evaluated at (a, b).

Theorem 1. *Fix $c \in \mathbb{R}$, let $f : \mathbb{R}^2 \to \mathbb{R}$ be a continuously differentiable function, and let $(a, b) \in \mathbb{R}^2$ be a point satisfying $f(a, b) = c$ and $\partial_1 f(a, b) \neq 0$ (or $\partial_2 f(a, b) \neq 0$). Then there exists an open interval (α, β), a point $\gamma \in (\alpha, \beta)$, and a function $g : (\alpha, \beta) \to \mathbb{R}$ (or $h : (\alpha, \beta) \to \mathbb{R}$) with the following properties:*

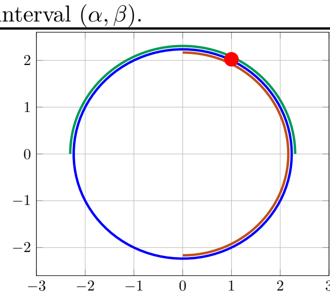
i. g (or h) is continuously differentiable,

ii. $g(\gamma) = c$ (or $h(\gamma) = c$), and

iii. $f(x, g(x)) = c \;\forall\, x \in (\alpha, \beta)$ (or $f(h(y), y) = c \;\forall\, y \in (\alpha, \beta)$).

The function $(\alpha, \beta) \ni x \mapsto g(x)$ is obtained by solving for y as a function of x in the expression $f(x, y) = c$ (for the second situation, $(\alpha, \beta) \ni y \mapsto h(y)$ is obtained by solving for x as a function of y in the expression $f(x, y) = c$). This final condition in this theorem says that the composition $\mathbb{R} \xleftarrow{f} (\alpha, \beta) \times \mathbb{R} \xleftarrow{(\mathrm{id}, g)} (\alpha, \beta)$ is the constant function whose value is c at all points in the interval (α, β).

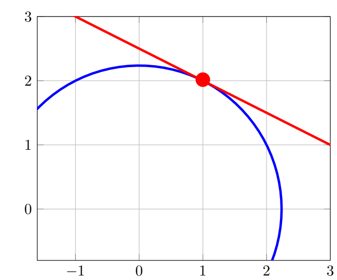
Let $\mathbb{R}^2 \ni (x, y) \overset{f}{\mapsto} x^2 + y^2$ and let $c := 5$ so that $f^{-1}(\{c\})$ is the circle of radius $\sqrt{5}$ centered at $(0, 0)$ in $\mathbb{R}^2$. Let $(a, b) = (1, 2) \in f^{-1}(\{c\})$. Then $\frac{d}{dy} f(1, y)\Big|_{y=2} = \frac{d}{dy}(1 + y^2)\Big|_{y=2} = 2y\Big|_{y=2} = 4 \neq 0$. Solving for y as a function of x in the expression $x^2 + y^2 = 5$ gives two solutions



$g(x) = \sqrt{5 - x^2}$ and $g(x) = -\sqrt{5 - x^2}$ for $x \in (\alpha, \beta) := (-\sqrt{5}, \sqrt{5})$. Since we require some $\gamma \in (-\sqrt{5}, \sqrt{5})$ such that $g(\gamma) = 2$, this forces us to choose the first option where $\gamma = \pm 1$ so that $\sqrt{5 - 1^2} = \sqrt{4} = 2$ as needed. If our point had been $(1, -2)$, we would have chosen $(-\sqrt{5}, \sqrt{5}) \ni x \mapsto g(x) = -\sqrt{5 - x^2}$ instead. Because $\frac{d}{dx} f(x, 2)\Big|_{x=1} = \frac{d}{dx}(x^2 + 4)\Big|_{x=1} = 2x\Big|_{x=1} = 2 \neq 0$, we could have also solved for x in terms of y and we'd get the function $(-\sqrt{5}, \sqrt{5}) \ni y \mapsto h(y) = \sqrt{5 - y^2}$ and $\gamma = \pm 2$ so that $h(\gamma) = 1$. Finally, if our point had been $(a, b) = (\sqrt{5}, 0)$, $\frac{d}{dy} f(\sqrt{5}, y)\Big|_{y=0} = 0$ so that we would not be able to solve for y in terms of x. However, $\frac{d}{dx} f(x, 0)\Big|_{x=\sqrt{5}} = 2\sqrt{5} \neq 0$ so that we can solve for x in terms of y. This gives $(-\sqrt{5}, \sqrt{5}) \ni y \mapsto h(y) = \sqrt{5 - y^2}$.

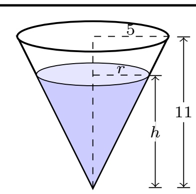
The IMT combined with the chain rule allows us to obtain $g'(x)$ (or $h'(y)$). Since the composition $\mathbb{R} \xleftarrow{f} (\alpha, \beta) \times \mathbb{R} \xleftarrow{(\mathrm{id}, g)} (\alpha, \beta)$ (or $\mathbb{R} \xleftarrow{f} \mathbb{R} \times (\alpha, \beta) \xleftarrow{(h, \mathrm{id})} (\alpha, \beta)$) is constant, the derivative of this composition is 0. The chain rule (when generalized to functions of several variables) can be applied to each function so that we can solve for the derivative of g (or h). These functions provides us with Jacobian matrices, which we multiply to obtain $(D_{(x, g(x))} f) \circ (D_x(\mathrm{id}, g)) = \begin{bmatrix} \partial_1 f(x, g(x)) & \partial_2 f(x, g(x)) \end{bmatrix} \begin{bmatrix} 1 \\ g'(x) \end{bmatrix} = \begin{bmatrix} \partial_1 f(x, g(x)) + g'(x) \partial_2 f(x, g(x)) \end{bmatrix}$ (or $(D_{(h(y), y)} f) \circ (D_y(h, \mathrm{id})) = \begin{bmatrix} \partial_1 f(h(y), y) & \partial_2 f(h(y), y) \end{bmatrix} \begin{bmatrix} h'(y) \\ 1 \end{bmatrix} = \begin{bmatrix} \partial_1 f(h(y), y) h'(y) + \partial_2 f(h(y), y) \end{bmatrix}$). Because $(\alpha, \beta) \ni x \mapsto f(x, g(x)) = c$ is constant, this derivative equals zero and we have the equation $\partial_1 f(x, g(x)) + g'(x) \partial_2 f(x, g(x)) = 0$. This allows us to solve for $g'(x)$ provided that $\partial_2 f(x, g(x))$ does not vanish and it is given by $g'(x) = -\frac{\partial_1 f(x, g(x))}{\partial_2 f(x, g(x))}$ for all $x \in (\alpha, \beta)$ (or $h'(y) = -\frac{\partial_2 f(h(y), y)}{\partial_1 f(h(y), y)}$ for all $y \in (\alpha, \beta)$) The usefulness of this result is that it is not always easy to obtain an explicit formula for g (or h) and it is therefore not always easy to explicitly take its derivative. For this reason, computing the derivative using the chain rule as via this method is called *implicit differentiation*. [Reference: Herb Gross: II.5 Implicit Dufferentiation]

The tangent line to the level set $f^{-1}(\{5\})$ at the point $(1, 2)$ for the function $\mathbb{R}^2 \ni (x, y) \overset{f}{\mapsto} x^2 + y^2$ is obtained by differentiating $x^2 + (g(x))^2 = 5$ with respect to x and using the chain rule to obtain $2x + 2g(x)g'(x) = 0$. Solving for $g'(x)$ gives $g'(x) = -\frac{x}{g(x)}$. Plugging in $x = 1$ (so that $g(1) = 2$) gives $g'(1) = -\frac{1}{2}$, which is the slope of



the tangent line. Because this line must intersect the point $(1, 2)$, the equation describing this line is $\mathbb{R} \ni x \mapsto y(x) = -\frac{1}{2}x + \frac{5}{2}$ (this formula was obtained by finding b from $y = mx + b$ since we know y, x, and m). One could have also obtained this same result by differentiating $(h(y))^2 + y^2 = 5$ with respect to y to get $h'(y) = -\frac{y}{h(y)}$ and plugging in $y = 2$ to get $h'(y) = -2$ which gives the equation $\mathbb{R} \ni y \mapsto x(y) = -2y + 5$ which is equivalent to our formula for y in terms of x. What happens if we try to find the tangent line at the point $(\sqrt{5}, 0)$? As we noticed before, we cannot solve for y as a function of x so we must instead solve for x as a function of y. This gives us $h'(0) = -\frac{0}{\sqrt{5}} = 0$ (if we had tried to implicitly differentiate with respect to x, we would have obtained $g'(\sqrt{5}) = -\frac{\sqrt{5}}{0}$, which is undefined) and therefore gives the tangent line to be $\mathbb{R} \ni y \mapsto x(y) = \sqrt{5}$.

Related rates



A cone-shaped water tank is being filled with water at a rate of 2 cubic meters per minute. The base of the tank is 5 meters in radius and the height of the tank is 11 meters. Find the rate at which the water level is rising when the water is 7 meters deep. What is the rate when the water level is 3 meters deep?

The relationship between the volume V, radius r, and height h of a cone is $V = \frac{1}{3}\pi r^2 h$. For a cone, the radius and height are related via $\frac{r}{h} = \frac{5}{11}$ due to similarity of triangles. Hence, $r = \frac{5}{11}h$. Therefore, $V = \frac{25}{363}\pi h^3$. Each of these quantities are changing as a function of time. Hence, the derivative is given by $V' = \frac{25}{121}\pi h^2 h'$. Since the flow of water is constant, this gives $2 = \frac{25}{121}\pi (3)^2 h'$ so that $h' = \frac{242}{225\pi} \approx 0.34$ meters per minute when the water level is 3 meters deep. When the water level is 7 meters deep, it is $\left(\frac{3}{7}\right)^2$ times this amount giving $h' = \frac{242}{1225\pi} \approx 0.063$ meters per minute.

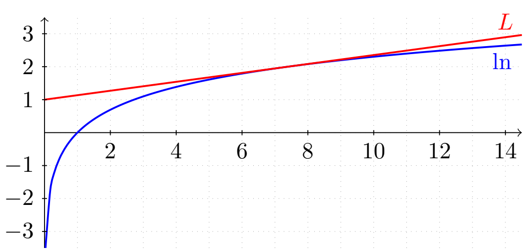
A record is 25 centimeters in diameter and rotates at 78 rotations per minute. There is a chip at the edge of the record. At what rate (in cm per minute) is the chip moving vertically when it is 5 centimeters above the horizontal axis?

Answer. The height h as a function of the angle is given by $h(\theta) = R\sin(\theta)$, where $R = 25$ cm. The angle is a function of time and we wish to compute $\frac{dh}{dt}$. Differentiating with respect to t gives $\frac{dh}{dt} = R\cos(\theta)\frac{d\theta}{dt}$. Note that $R\cos(\theta) = \sqrt{R^2 - h^2}$ for small angles so that this becomes $\frac{dh}{dt} = \sqrt{R^2 - h^2}\frac{d\theta}{dt}$. Since $\frac{d\theta}{dt} = 78(2\pi) = 156\pi$ radians per minute, this gives $\frac{dh}{dt} = 156\pi\sqrt{600} = 1560\pi\sqrt{6} \approx 12000$ cm per minute, i.e. 12 meters per minute.

Linear approximation

The derivative of a differentiable function at a point can be used to learn a lot about the function in a neighborhood of that point. In all of the examples we have seen, the tangent line to a curve seems to be a reasonable linear approximation to the curve provided that one is close enough to that curve. Let $f : (a, b) \to \mathbb{R}$ be a differentiable function on an open interval and let $c \in (a, b)$. The linear function $L : (a, b) \to \mathbb{R}$ given by $(a, b) \ni x \mapsto f'(c)x + f(c) - f'(c)c = f(c) + f'(c)(x - c)$ is called the <u>*linearization of f at c*</u>. The difference $(a, b) \ni x \mapsto |L(x) - f(x)|$ is called the <u>*deviation*</u> and $\sup\limits_{x\in(a,b)} |L(x) - f(x)|$ is called the <u>*accuracy*</u>. Here $\sup\limits_{x\in(a,b)} g(x)$ denotes the lowest upper bound, a number u such that $g(x) \le u$ for all $x \in (a, b)$ and such that for any other number v, the inequality $u \le v$ holds.

Linearizing $\ln : (0, \infty) \to \mathbb{R}$ at $c = e^2$ gives $(0, \infty) \ni x \mapsto L(x) = 2 + e^{-2}(x - e^2) = 1 + e^{-2}x$. Since $e \approx 2.72$, this linear approximation of ln at 9 gives approximately $L(9) \approx 1 + \frac{9}{7.40} \approx 2.22$, which is very close to the real value $\ln(9) \approx 2.20$ as the deviation is only about 0.02.



However, it is apparent from the graph that L is a worse approximation for inputs less than e^2. For example, to obtain an accuracy of about 0.1, we would need to find an interval (a, b) for which $\sup\limits_{x\in(a,b)} |1 + e^{-2}x - \ln(x)| \le 0.1$. It is not possible to obtain a closed form expression for this since one must solve $1 + e^{-2}x - \ln(x) = 0.1$. There are two solutions to this equation (that can be obtained numerically) and they are approximately $x = 4.558$ and $x = 11.203$. Hence, the approximation L is valid to within 0.1 accuracy on the interval $(4.558, 11.203)$. Notice that the difference between 4.558 and e^2 is 2.831 while the difference between e^2 and 11.203 is 3.814 showing that the linearization is a better approximation for values larger than e^2 than it is for values smaller than e^2.

To estimate 1.999^4 note that $2^4 = 16$ and $\frac{d}{dx}x^4 = 4x^3$ so that the linearization of p_4 at $c = 2$ is $\mathbb{R} \ni x \mapsto L(x) = p_4(2) + p_4'(2)(x - 2) = 2^4 + 4(2^3)(x - 2) = 16 + 32(x - 2)$. Hence, $1.999^4 \approx L(1.999) = 16 + 32(-0.001) = 16 - 0.032 = 15.968$ (notice we didn't need a calculator).

To estimate $9^{1.99}$ note that $9^2 = 81$ and $\frac{d}{dx}9^x = \ln(9)9^x$ so that the linearization of $x \mapsto 9^x$ at $c = 2$ is $L(x) = 81 + 81\ln(9)(x - 1.99) \approx 81 + 81(2.22)(-0.01) \approx 81 - 1.78 = 79.22$ which is quite close to the actual value $9^{1.99} \approx 79.24$. Here, we have used our earlier approximation of $\ln(9)$.

Differentials*

Consider the associative unital algebra Ω^* over $\mathbb{R}$ generated by the two symbols dx and dy subject to the relation $(dx)^2 = 0 = (dy)^2$ and $dxdy = -dydx$. A basis for this algebra viewed as a vector space is $\{1, dx, dy, dxdy\}$. Let $C^\infty(\mathbb{R}^2)$ denote the set of <u>*smooth*</u>, i.e. infinitely differentiable, functions $\mathbb{R}^2 \to \mathbb{R}$. The set of <u>*(smooth) differential forms*</u> on $\mathbb{R}^2$ is the tensor product $\Omega(\mathbb{R}^2) := C^\infty(\mathbb{R}^2) \otimes \Omega^*$, i.e. formal sums of elements of the basis with smooth functions as coefficients. In this case, they are all of the form $f + g_1dx + g_2dy + hdxdy$ for some smooth functions $f, g_1, g_2, h : \mathbb{R}^2 \to \mathbb{R}$. This is analogous to linear combinations of elements of a vector space basis except that the coefficients are *functions* instead of *numbers*. Define $d : \Omega(\mathbb{R}^2) \to \Omega(\mathbb{R}^2)$ to be the linear map uniquely determined by

i. $df = \partial_1 f dx + \partial_2 f dy$ for all $f \in C^\infty(\mathbb{R}^2)$ and
ii. $d(f + g_1dx + g_2dy + hdxdy) = df + dg_1dx + dg_2dy + dhdxdy$ for all $f, g_1, g_2, h \in C^\infty(\mathbb{R}^2)$.

$d(df) = d(\partial_1 f dx + \partial_2 f dy) = d(\partial_1 f)dx + d(\partial_2 f)dy = \partial_1^2 f dxdx + \partial_2\partial_1 f dydx + \partial_1\partial_2 f dxdy + \partial_2^2 f dydy = (\partial_1\partial_2 f - \partial_2\partial_1 f)dxdy = 0$ for all $f \in C^\infty(\mathbb{R}^2)$.

Exercises

Exercise 1. A spherical tank with radius r_0 is filled with water at a rate of R gallons per minute. How quickly is the water level rising when the depth of the water is d?

Exercise 2. Provide two qualitatively different counter-examples to the IMT when any of the assumptions fail. Namely, what happens if there is a point $(a, b) \in \mathbb{R}^2$ such that $f(a, b) = c$ but for which $\frac{d}{dx}f(x, b)\Big|_{x=a} = 0$ and $\frac{d}{dy}f(a, y)\Big|_{y=b} = 0$? Can a tangent line be drawn to the level set $f^{-1}(\{c\})$ at the point (a, b) in these cases? Explain.

Exercise 3. Plot the level sets $f^{-1}(\{c\})$ of the function $\mathbb{R}^2 \ni (x, y) \xmapsto{f} 25(x^2 - y^2) - 2(x^2 + y^2)^2$ for the various values of $c \in \{-100, -50, -25, 0, 25, 50, 75\}$. What happens when $c = 80$? Discuss the possibility of being able to draw a line tangent to the level set $f^{-1}(\{0\})$ at the point $(0, 0)$.

Exercise 4. An experiment is done where mixed dough and yeast is placed between glass plates of thickness L and the dough of mass m ferments and expands with the area and density related by $\rho(t) = \frac{m}{LA(t)}$ where t is time. Suppose the area as a function of time is given by $A(t) = a_0(1 + c\tanh(t))$, where $a_0 > 0$ is the initial area and c is a constant determined by how much yeast and dough is initially in the mixture. What is the rate at which the density decreases when the area of the dough is a?

Exercise 5. Prove the function $\mathbb{R} \ni x \xmapsto{f} \begin{cases} 0 & \text{if } x = 0 \\ x^2\sin\left(\frac{1}{x}\right) & \text{otherwise} \end{cases}$ is differentiable. Compute $f'(x)$ for all $x \in \mathbb{R}$. Is f' continuous? Discuss.

Exercise 6. Write down the linearization of $\ln : (0, \infty) \to \mathbb{R}$ at the point $c = e$ (e is Euler's constant) and draw the graph of ln together with the graph of the linearization. Over what interval is the accuracy 0.1? Compare the size of this interval to the size of the interval obtained from linearizing ln at the point e^2. Repeat at the point $c = 1$. Hint: use Desmos, the online graphing calculator!

Exercise 7. Write down the linearization of $\sin : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$ at the point $\theta = 0$ and draw the graph of sin together with the graph of the linearization. Over what interval is the accuracy 0.1? Repeat for the function $\cos : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$. Hint: use Desmos, the online graphing calculator!

Exercise 8. Explain why the functions arctan and tanh are good approximations of each other in a neighborhood of 0. Compute their accuracy over the intervals $(-1, 1)$ and $\left(-\frac{1}{2}, \frac{1}{2}\right)$.

Physics

The infinitesimal change in position x of a particle over an infinitesimal change in time is the velocity of that particle $v(t) := \frac{dx(t)}{dt}$. A particle of mass m is thrown straight down off a cliff. Its change in velocity over an infinitesimal change is its acceleration $a(t) := \frac{dv(t)}{dt}$ and Newton's law says that $F = ma$, the force on the particle of mass m is proportional to its acceleration. Ignoring air resistance, the differential equation describing a mass being thrown down a cliff is $mg = ma = m\frac{dv}{dt}$, where g is the gravitational constant on Earth. The solution to this equation is $v(t) = v_0 + gt$ and the position is $x(t) = -h_0 + v_0 t + \frac{1}{2}gt^2$ valid for all time t. Here h_0 is the initial height and v_0 is the initial velocity. You should verify this by computing the derivatives and checking they solve the differential equation. Notice that the position does not depend on mass. Does the position still not depend on mass if there is air resistance? For a mass dropping off a cliff and being pulled due to gravity but *also* being repelled by air resistance, the differential equation describing this motion is $mg - \gamma v^2 = ma = m\frac{dv}{dt}$, where g is the gravitational constant on Earth and γ is a coefficient describing the strength of the air resistance. One can solve this equation for the velocity and then integrate to find the position. The result for the position is $x(t) = -h_0 + \frac{m}{\gamma}\ln\left(\cosh\left(t\sqrt{\frac{\gamma g}{m}} + \operatorname{arctanh}\left(v_0\sqrt{\frac{\gamma}{mg}}\right)\right)\right) - \frac{m}{\gamma}\ln\left(\cosh\left(\operatorname{arctanh}\left(v_0\sqrt{\frac{\gamma}{mg}}\right)\right)\right)$, where h_0 is the initial height and v_0 is the initial velocity. Computing the derivatives, we can verify this solves the differential equation. The derivative of the position is (using the chain rule twice) $v(t) = \frac{dx(t)}{dt} = \frac{m}{\gamma}\frac{\frac{d}{dt}\cosh\left(t\sqrt{\frac{\gamma g}{m}}+\operatorname{arctanh}\left(v_0\sqrt{\frac{\gamma}{mg}}\right)\right)}{\cosh\left(t\sqrt{\frac{\gamma g}{m}}+\operatorname{arctanh}\left(v_0\sqrt{\frac{\gamma}{mg}}\right)\right)} = \sqrt{\frac{mg}{\gamma}}\tanh\left(t\sqrt{\frac{\gamma g}{m}} + \operatorname{arctanh}\left(v_0\sqrt{\frac{\gamma}{mg}}\right)\right)$ and the derivative of velocity is $a(t) = \frac{dv(t)}{dt} = g\operatorname{sech}^2\left(t\sqrt{\frac{\gamma g}{m}} + \operatorname{arctanh}\left(v_0\sqrt{\frac{\gamma}{mg}}\right)\right)$. By using hyperbolic trigonometric identities, the function $\mathbb{R} \ni t \mapsto x(t)$ solves the differential equation. It is interesting to plot these two solutions and see what happens as γ approaches 0 from above. This is left as an exercise. [Reference: Philosophical Math blog]

Chemistry

Let $\{R_i\}$ and $\{P_j\}$ be finite sets of reactants and products, respectively. Consider a chemical reaction of the form $R_1 + \cdots + R_M \to P_1 + \cdots + P_N$. The *concentration* $[Q]$ of a substance Q is the number of moles (1 mole $= 6.022 \times 10^{23}$ molecules) per liter. During a reaction, the concentration of a reactant or product varies as a function of time because the reaction occurs at a rate proportional to the amount of reactants available at that time. The *instantaneous rate of reaction* of a substance Q is $\frac{d[Q]}{dt}$, the derivative of the concentration $[Q]$ with respect to time. This is negative for products and positive for reactants, i.e. $0 < \frac{d[P_i]}{dt} = -\frac{d[R_j]}{dt}$ for all i,j. In general, the relationship between the concentration $[P]$ of a product and its rate of reaction $\frac{d[P]}{dt}$ is $\frac{d[P]}{dt} = k[P]$, where $k > 0$ is a constant describing the reaction rate.

Biology

The von Bertalanffy growth equation describes the growth of an organism. The mass of an organism in the shape of a cube of length L and mass M that ingests at a rate proportional to its surface area (food is absorbed through the walls) and respires at a rate proportional to its volume (the organism needs to poop) is governed by the differential equation $\frac{dM}{dt} = aL^2 - bL^3$, where $a, b > 0$ are constants. Assuming the organism is composed of water, the mass is L^3 because the density of water is 1 liter per cubic decimeter. Hence, $\frac{dM}{dt} = \frac{d}{dt}(L^3) = 3L^2\frac{dL}{dt}$. This turns the differential equation into $3L^2\frac{dL}{dt} = aL^2 - bL^3$. After rearranging and cancelling terms, it becomes $\frac{dL}{dt} = \frac{a}{3} - \frac{b}{3}L = k(L_\infty - L)$, where $k := \frac{b}{3}$ and $L_\infty := \frac{a}{b}$.

Making fat stacks

Suppose that x_0 dollars are invested at an interest of r ($r = 0.09$ means at 9%) that is compounded annually. The value $x(t)$ of the investment after t years is given by $x(t) = x_0(1+r)^t$. If the interest is compounded n times per year, then the value is $x(t) = x_0\left(1+\frac{r}{n}\right)^{nt}$. As n increases, this tends to $\lim_{n\to\infty} x_0\left(1+\frac{r}{n}\right)^{nt} = \lim_{n/r\to\infty} x_0\left(\left(1+\frac{1}{n/r}\right)^{n/r}\right)^{rt} = x_0 e^{rt}$. Already for $n = 365$, which corresponds roughly to compounding the interest daily, this is a good approximation. Now imagine that x_0/n dollars are invested at the same rate as the interest is compounded (so once per year if $n = 1$). The value after t years is then $\sum_{k=1}^{nt}\frac{x_0}{n}\left(1+\frac{r}{n}\right)^k = \frac{x_0}{n}\left(\frac{\left(1+\frac{r}{n}\right)^{nt+1}-1}{\left(1+\frac{r}{n}\right)-1} - 1\right) = \frac{x_0}{r}\left(1+\frac{r}{n}\right)\left(\left(1+\frac{r}{n}\right)^{nt}-1\right)$. Taking the limit as $n \to \infty$ gives ("inc" stands for increment investments) $x_{\text{inc}}(t) = \lim_{n\to\infty}\sum_{k=1}^{nt}\frac{x_0}{n}\left(1+\frac{r}{n}\right)^k = \frac{x_0}{r}\left(e^{rt}-1\right)$ and gives an approximate for the value after t years. Compare this to somebody who invests x_0 at the beginning of each year but who experiences interest at an annual rate r compounded n times per year. The value of their investment after t years becomes $x_0\sum_{k=1}^{t}\left(1+\frac{r}{n}\right)^{nk} = x_0\left(\frac{\left(1+\frac{r}{n}\right)^{n(t+1)}-1}{\left(1+\frac{r}{n}\right)^n-1} - 1\right) = x_0\left(\frac{\left(1+\frac{r}{n}\right)^n}{\left(1+\frac{r}{n}\right)^n-1}\right)\left(\left(1+\frac{r}{n}\right)^{nt}-1\right)$. The limit of this expression as $n \to \infty$ is ("lum" stands for lump yearly investments) $x_{\text{lum}}(t) = \lim_{n\to\infty} x_0\sum_{k=1}^{t}\left(1+\frac{r}{n}\right)^{nk} = x_0\left(\frac{e^r}{e^r-1}\right)\left(e^{rt}-1\right)$. The person who invests more at the beginning of the year will experience more gains over time. Here is a table for the *profit* (total value minus net invested) of an given investment of $10000 per year after a total of t years comparing frequent (daily) deposits versus yearly deposits assuming an interest rate $r = 0.03$ compounded daily.

Time →	$t=1$	$t=5$	$t=10$	$t=20$	$t=30$	$t=50$
every day	151	3944	16619	74039	186534	660563
per year	304	4757	18377	78170	193868	678058

Making even fatter stacks

An interest rate of $r = 0.03$ is typical of a pretty good CD (certificate of deposit) at your local bank. The total stock market index, say the S&P 500, provides the investor with an average annual return (AAR) of roughly 9.8% when averaged out over its lifetime (which is at least 70 years). To define the AAR, let $r_k := \frac{x_k - x_{k-1}}{x_k}$ be the yearly rate for k-th year, i.e. if x_{k-1} is invested once at the beginning of the k-th year and x_k is the value of the investment at the end of the k-th year, then r_k is the unique real number such that $x_k = (1+r_k)x_{k-1}$ (the difference $x_k - x_{k-1}$ is called the *return* in the k-th year). The *average annual return* (AAR) for an index over N years is defined as $r := \frac{1}{N}\sum_{k=1}^{N} r_k$. Using this AAR, an investor can get a rough estimate (cf. Exercise 4) for the profit earned from a given investment. We cannot apply the $n \to \infty$ limit in this case since we are assuming yearly compounded interest. The table below compares the profits given an investment of $10000 per year assuming that it is compounded yearly based on the formula $x(t) = x_0\left(\frac{1+r}{r}\right)\left((1+r)^t - 1\right)$.

Time →	$t=1$	$t=5$	$t=10$	$t=20$	$t=30$	$t=50$
$r = 0.03$	300	4684	18077	76764	190026	661807
$r = 0.098$	980	16767	73323	414772	1439129	11396583

Maxima and minima of differentiable functions

Let $D \subseteq \mathbb{R}$, let $f : D \to \mathbb{R}$ be a function, and let $c \in D$. $f(c)$ is said to be

i. the *(absolute) maximum* of f iff $f(x) \le f(c)$ for all $x \in D$,

ii. the *(absolute) minimum* of f iff $f(x) \ge f(c)$ for all $x \in D$,

iii. a *local maximum* of f iff $f(c)$ is the maximum on $(c-\epsilon, c+\epsilon) \cap D$ for some $\epsilon > 0$, and

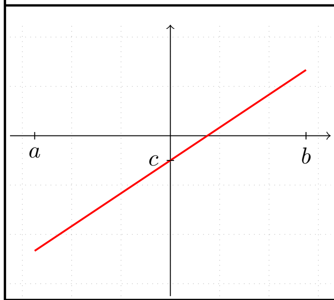
iv. a *local minimum* of f iff $f(c)$ is the minimum on $(c-\epsilon, c+\epsilon) \cap D$ for some $\epsilon > 0$.

[Reference: Herb Gross: II.8 Maxima and Minima]

It is important to know whether a maximum or minimum exists for a given function. The *Extreme Value Theorem* is one such theorem.

Theorem 1. *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a,b] \to \mathbb{R}$ be continuous. Then f attains a maximum and minimum, i.e. there exist $y, z \in [a,b]$ with $f(y)$ a minimum and $f(z)$ a maximum of f.*

The two important assumptions in this theorem is the form of the domain (a closed interval) and the fact that the function must be continuous. If either of these assumptions are removed, the result is false. For example, $(0,1) \ni x \mapsto \frac{1}{x}$ is continuous but is unbounded and obtains neither a maximum nor a minimum. The function $[0,1] \ni x \mapsto \begin{cases} 0 & \text{if } x = 0 \\ \frac{1}{x} & \text{otherwise} \end{cases}$ is defined on a closed interval, is not continuous, and does not attain a maximum.



Consider the function $\mathbb{R} \ni x \mapsto f(x) := mx + c$, where $m, c \in \mathbb{R}$. For concreteness, suppose that $m > 0$. f does not attain its maximum nor its minimum on $\mathbb{R}$ (it is unbounded). However, if f is restricted to a closed interval $[a,b]$, its minimum occurs at a and is given by $f(a) = ma + c$ while its maximum occurs at b and is given by $f(b) = mb + c$.

Fermat's theorem relates local maxima and minima of a function to its derivative.

Theorem 2. *Let $f : D \to \mathbb{R}$ be a function on a domain D with a local maximum or minimum at $c \in D$. If $f'(c)$ exists, then $f'(c) = 0$.*

The converse of this theorem ("if $f'(c) = 0$ then $f(c)$ is a local maximum or minimum") is false. A counter-example is the function $\mathbb{R} \ni x \mapsto x^3$ with $c = 0$. Nevertheless, the points at which the derivative of a function vanish are incredibly important in the study of the behavior of functions. A *critical point* of a function $f : D \to \mathbb{R}$ is a number $c \in D$ such that $f'(c) = 0$ or $f'(c)$ does not exist.

Notice that the assumption that "$f'(c)$ exists" cannot be dropped from Fermat's theorem. The absolute value function $|\cdot| : \mathbb{R} \to \mathbb{R}$ has a local minimum at 0 but is not differentiable at 0.

Consider the function $[-\pi, 2\pi] \ni x \mapsto f(x) := x - 2\sin(x)$. What are the absolute maxima and minima of this function? f is differentiable and its derivative is given by $[-\pi, 2\pi] \ni x \mapsto f'(x) := 1 - 2\cos(x)$ so its local maxima and minima are obtained from setting $0 = f'(x) = 1 - 2\cos(x)$. The solutions to $\cos(x) = \frac{1}{2}$ for $x \in [-\pi, 2\pi]$ are given by $x \in \{-\frac{\pi}{3}, \frac{\pi}{3}, \frac{5\pi}{3}\}$. The values of f at these points and the boundaries of the domain along with which is the max and min are given in the following table.

x	$-\pi$	$-\frac{\pi}{3}$	$\frac{\pi}{3}$	$\frac{5\pi}{3}$	2π
$f(x)$	$-\pi$	$-\frac{\pi}{3} + \sqrt{6}$	$\frac{\pi}{3} - \sqrt{6}$	$\frac{5\pi}{3} + \sqrt{6}$	2π
		local max	local min	local max	
	min			max	

Exercises

Exercise 1. Referring to the physics example of a particle being thrown off of a cliff, let $x(t, \gamma)$ denote the position at time t with air resistance constant γ. Show that $\lim_{\gamma \searrow 0} x(t, \gamma) = -h_0 + v_0 t + \frac{1}{2} g t^2$ for all $t \in \mathbb{R}$.

Exercise 2. Referring to the chemistry example, consider the reaction $4\,NH_3 + 3\,O_2 \longrightarrow 2\,N_2 + 6\,H_2O$, where 4 moles of ammonia reacts with 3 moles of oxygen to form 2 moles of nitrogen and 6 moles of water. If the instantaneous rate of the reaction is $k > 0$ compute the instantaneous rate of change of the concentrations of ammonia, oxygen, nitrogen, and water, in terms of k.

Exercise 3. Referring to the biology example of the von Bertalanffy growth equation, let L_0 be the initial length of an organism. Show that $[0, \infty) \ni t \mapsto L(t) = L_0 e^{-kt} + L_\infty(1 - e^{-kt})$ solves the von Bertalanffy growth differential equation. Sketch a graph of L assuming that $L_0 < L_\infty$. Discuss the significance of this.

Exercise 4. Suppose that Ian invests x_0 dollars in a mutual fund. Let r_k denote the return rate for the k-th year and suppose that Ian leaves his investment alone for N years. The formula for the value of the investment after N years is given by $x_0(1+r_1)(1+r_2)\cdots(1+r_N)$. Let r denote the AAR over this time period. The approximated value of the investment after N years using the AAR is $x_0(1+r)^N$. Assuming that $r_j r_k r_l$ (the product of three rates) is so small that it can be neglected, compute the difference $x_0(1+r_1)(1+r_2)\cdots(1+r_N) - x_0(1+r)^N$. Discuss the significance of this result. Hint: use the binomial expansion theorem.

Exercise 5. Consider the same setup as in Exercise 4. Let $N = 5$ and suppose the yearly rates are given as in the following table.

	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
r_k	0.4	0.4	0.1	0.3	-0.7

Compute the AAR. Assuming that $x_0 = 10000$ is invested every year, compare the net profit after 5 years using the actual rates provided versus the computed net profit using the AAR over the 5 year period. A net profit that is negative indicates a loss. What permutation of the r_k's results in the largest net profit? What permutation of the r_k's results in the lowest net profit (greatest loss)? Discuss the significance of this result and compare this result to Exercise 4. In particular, does it matter *when* the market crashes for the investor who only invests once at the beginning (assume \$50000 is invested so that both parties invest the same amount of money)? What are the benefits/drawbacks to investing annually based on these models?

Exercise 6. Let $r \in \mathbb{R}$. Find all the critical points of the function $[0, \infty) \ni x \mapsto f(x) := x^r$ and indicate whether the derivative vanishes or is undefined at these points.

Exercise 7. Is it possible for a continuous real-valued function f to be defined on *all* of $\mathbb{R}$ and have a critical point $c \in \mathbb{R}$ such that $f'(c)$ is not defined? If so, give an example. Otherwise, prove that every function $f : \mathbb{R} \to \mathbb{R}$ that has a critical point $c \in \mathbb{R}$ satisfies $f'(c) = 0$.

Exercise 8. The Lennard-Jones potential is a model for the potential describing the interaction between a pair of neutral atoms or molecules. Although neutral as a whole, atoms have negatively charged electron clouds surrounding their positively charged nuclei, which create an effective electrostatic potential due to their multipole moments. The potential has the form $(0, \infty) \ni r \mapsto V(r) = \frac{a}{r^{12}} - \frac{b}{r^6}$, where $a, b > 0$ are constants and r is interpreted as the separation between the two atoms. In terms of a and b, find the distance r_m that minimizes V and find the absolute minimum $-\epsilon$ of V. Rewrite the potential in terms of these new constants.

The Mean Value Theorem

Rolle's Theorem is a theorem relating the values of a differentiable function on an interval at its boundary to its derivative on the interior of that interval.

Theorem 1. *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a function satisfying*

i. f is continuous on $[a, b]$,

ii. f is differentiable on (a, b), and

iii. $f(a) = f(b)$.

Then there exists a point $c \in (a, b)$ with $f'(c) = 0$.

[Reference: Herb Gross: II.9 Rolle's Theorem and its Consequences]

Consider the function $\left[0, \frac{\pi}{2}\right] \ni x \mapsto f(x) := \cos(x)\sqrt{x}$. Then f is continuous on its domain. Although f is not differentiable on its domain, it is differentiable on $\left(0, \frac{\pi}{2}\right]$ and its derivative is $f'(x) = \frac{\cos(x)}{2\sqrt{x}} - \sin(x)\sqrt{x}$ for all $x \in \left(0, \frac{\pi}{2}\right]$. Finally, $f(0) = f(\frac{\pi}{2}) = 0$. Hence, there exists a point $c \in \left(0, \frac{\pi}{2}\right)$ such that $f'(c) = 0$. Such a point is a solution to $2c = \frac{\cos(c)}{\sin(c)}$, but it is not obvious a solution to this equation exists. Rolle's theorem guarantees its existence.

What can go wrong if any of the hypotheses of Rolle's theorem fails? The following three examples are ones for which one of the assumptions fails but none of the others do.

i. $[0, 1] \ni x \mapsto f(x) := \begin{cases} 0 & \text{if } x = 0 \\ 1 - x & \text{otherwise} \end{cases}$. Notice that f satisfies ii. and iii. but not i. and $f'(x) = -1$ for all $x \in (0, 1)$.

ii. $[-1, 1] \ni x \mapsto g(x) := |x|$ satisfies i. and iii. but not ii. and $g'(x) \in \{-1, 1\}$ for all $x \in [-1, 0) \cup (0, 1]$. $g'(0)$ is undefined.

iii. $\mathrm{id} : [0, 1] \to [0, 1]$ satisfies i. and ii. but not iii. and $\mathrm{id}' = 1$ everywhere.

A slight generalization of Rolle's theorem is the *Mean Value Theorem* (MVT), which provides some information about a function based on its derivative.

Theorem 2. *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a function satisfying*

i. f is continuous on $[a, b]$ and

ii. f is differentiable on (a, b).

Then there exists a point $c \in (a, b)$ with $f'(c) = \dfrac{f(b) - f(a)}{b - a}$.

Notice that Rolle's Theorem is a special case of the Mean Value Theorem obtained by setting $f(a) = f(b)$.

The MVT can be used to prove our first theorems about anti-derivatives.

Theorem 3. *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a differentiable function such that $f'(x) = 0$ for all $x \in (a, b)$. Then f is a constant function, i.e. there exists a $d \in \mathbb{R}$ such that $f(x) = d$ for all $x \in [a, b]$.*

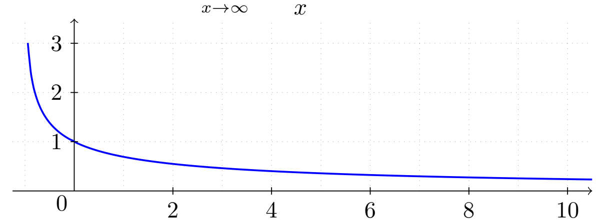
Proof. Let $x, y \in (a, b)$ with $x < y$. Then f restricted to $[x, y]$ satisfies the assumptions of the MVT. Hence, there exists a $c \in (x, y)$ such that $(y - x)f'(c) = f(y) - f(x)$, but the former equals 0 so $f(y) = f(x)$. Since x, y were chosen arbitrarily, we have $f(x) = f(y)$ for all $x, y \in (a, b)$. ■

Theorem 4. *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f, g : [a, b] \to \mathbb{R}$ be differentiable functions such that $f'(x) = g'(x)$ for all $x \in (a, b)$. Then $g = f + c$ for some $c \in \mathbb{R}$.*

Throughout the rest of calculus, you will see that a lot of information about a function can be determined not just by its first derivative, but by its higher order derivatives as well. The culminations of this concept include Lagrange's remainder theorem, which gives an approximation for a function that is differentiable many times in terms of its derivatives and polynomials (see Theorem 7), and the Taylor series.

L'Hospital's rules

Although we were able to prove $\lim\limits_{x\to 0} \dfrac{\sin(x)}{x} = 1$ using trigonometric inequalities, we could not apply the Algebraic Limit Theorem (ALT) because $\lim\limits_{x\to 0} \sin(x) = 0$ and $\lim\limits_{x\to 0} x = 0$ and this ratio is undefined. Occasionally, one is confronted with limits where the ALT fails and also for which trigonometric or geometric identities might not be applicable. For instance, consider the limit $\lim\limits_{x\to 0} \dfrac{\ln(x+1)}{x}$. By graphing the values of the function $(-1, \infty) \ni x \mapsto \frac{\ln(x+1)}{x}$, it is apparent that the limit of this exists as x approaches 0. It is also apparent that the limit $\lim\limits_{x\to\infty} \dfrac{\ln(x+1)}{x}$ should be 0.



The first case mentioned above is covered by *L'Hospital's Rule: 0/0 case*.

Theorem 5. *Let $a < b$ and let $c \in [a, b]$. Let $f, g : [a, b] \to \mathbb{R}$ be continuous functions and suppose that the restrictions of f and g to $[a, c) \cup (c, b]$ are differentiable. Furthermore, suppose that $f(c) = g(c) = 0$ and $g'(x) \neq 0$ for all $x \in [a, c) \cup (c, b]$. Then*

$$\lim_{x\to c} \frac{f'(x)}{g'(x)} = L \qquad \Rightarrow \qquad \lim_{x\to c} \frac{f(x)}{g(x)} = L.$$

Our first example satisfies the assumptions of this theorem so that $\lim\limits_{x\to 0} \dfrac{\ln(x+1)}{x} = \lim\limits_{x\to 0} \dfrac{\frac{1}{x+1}}{1} = \dfrac{\lim\limits_{x\to 0} \frac{1}{x+1}}{\lim\limits_{x\to 0} 1} = 1$. As another example, $\lim\limits_{x\to 0} \dfrac{\tan(x) - x}{x^3} = \lim\limits_{x\to 0} \dfrac{\frac{1}{\cos^2(x)} - 1}{3x^2} = \lim\limits_{x\to 0} \dfrac{\tan^2(x)}{3x^2} = \lim\limits_{x\to 0} \dfrac{2\tan(x)}{6x\cos^2(x)} = \lim\limits_{x\to 0} \left(\dfrac{1}{3\cos^3(x)} \dfrac{\sin(x)}{x}\right) = \dfrac{1}{3}$ by two applications of L'Hospital's rule.

Another instance where the ALT fails is for limits such as $\lim\limits_{x\to 0} \big(x \ln(|x|)\big)$ and yet this limit can be shown to exist.

The second case mentioned above is covered by *L'Hospital's Rule: ∞/∞ case*.

Theorem 6. *Let $a < b$ and let $c \in (a, b)$. Let $f, g : (a, c) \cup (c, b) \to \mathbb{R}$ be differentiable functions. Furthermore, suppose that $g'(x) \neq 0$ for all $x \in (a, c) \cup (c, b)$. If $\lim\limits_{x\to c} f(x) = \pm\infty$ and $\lim\limits_{x\to c} g(x) = \pm\infty$, then*

$$\lim_{x\to c} \frac{f'(x)}{g'(x)} = L \qquad \Rightarrow \qquad \lim_{x\to c} \frac{f(x)}{g(x)} = L.$$

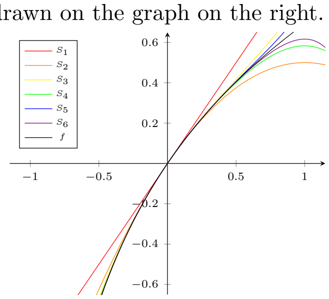
The above limit can be expressed as $\lim\limits_{x\to 0} x \ln(|x|) = \lim\limits_{x\to 0} \dfrac{\ln(|x|)}{1/x}$, where the assumptions of this theorem now apply. The limit is $\lim\limits_{x\to 0} \big(x \ln(|x|)\big) = \lim\limits_{x\to 0} \left(\dfrac{1/x}{-1/x^2}\right) = -\lim\limits_{x\to 0} x = 0$. One can also extend this theorem to compute limits as $x \to +\infty$ or $x \to -\infty$. For example, one could ask does $\ln(x)$ grow as $x^{1/n}$ for any $n \in \mathbb{N}$? Computing $\lim\limits_{x\to\infty} \dfrac{\ln(x)}{x^{1/n}} = \lim\limits_{x\to\infty} \dfrac{1/x}{(1/n)x^{(1/n)-1}} = \lim\limits_{x\to\infty} \dfrac{n}{x^{1/n}} = 0$ shows this is false.

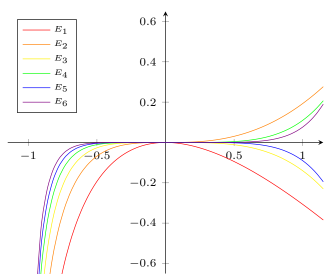
Lagrange's Remainder Theorem expresses that successive derivatives of a function at a point provide information about how the function looks like at nearby points. If $f : \mathbb{R} \to \mathbb{R}$ is n-times differentiable, then its n-th derivative is denoted by $f^{(n)}$.

Theorem 7. *Let $R > 0$ and let $f : (-R, R) \to \mathbb{R}$ be a function that is differentiable $N+1$ times with $N \in \mathbb{N}$. Set $a_n := \frac{f^{(n)}(0)}{n!}$ for each $n \in \{0, 1, \dots, N+1\}$. Also define $(-R, R) \ni x \mapsto S_N(x) := \sum_{n=0}^{N} a_n x^n$, and $E_N : (-R, R) \to \mathbb{R}$ by $E_N := f - S_N$. Then, for any $x \in (-R, R) \setminus \{0\}$, there exists a $c \in (-|x|, |x|)$ such that $E_N(x) = \frac{f^{(N+1)}(c)}{(N+1)!} x^{N+1}$ for that particular value of x.*

Given any differentiable function f, the theorem says that the value of f at some point $x \in (-R, R) \setminus \{0\}$ can be expressed in terms of two quantities. The first is a polynomial evaluated at x all of whose data comes from the (successive) derivatives of f at a single point, namely 0. This is "infinitesimal data." The second quantity is essentially what is left over. However, it, too, can be expressed as a polynomial of the next degree. This is no longer infinitesimal since the point c could in general be larger than 0, but it is at least between $-|x|$ and $|x|$. The importance of E_N is the actual expression for it in terms of $c \in (-|x|, |x|)$, which illustrates that if the higher order derivatives of f are not unreasonably large, then the error term is very small.

Consider the function $f : (-1, \infty) \to \mathbb{R}$ given by $(-1, \infty) \ni x \mapsto f(x) := \ln(x+1)$. The first few approximations using only the polynomial terms associated to the derivatives of f at 0 is given on the graph on the left while the difference between these Taylor approximations from the partial sums S_n and the actual function f is drawn on the graph on the right.





As the graph on the right indicates, the error term approaches zero as n increases. We can try to provide a bound for this error term using Lagrange's Remainder Theorem. Note that $f^{(N+1)}(x) = \frac{(-1)^N N!}{(x+1)^{N+1}}$ for $N \in \mathbb{N} \cup \{0\}$ and for all $x \in (-1, 1)$. Hence, fix $x \in (-1, 1)$. Then $\left|f^{(N+1)}(c)\right| = \frac{N!}{(c+1)^{N+1}}$ for all $c \in (-|x|, |x|)$ so that $\left|\frac{f^{(N+1)}(c)}{(N+1)!} x^{N+1}\right| \leq \frac{N!}{(N+1)!}\left(\frac{|x|}{c+1}\right)^{N+1} = \frac{1}{N+1}\left(\frac{|x|}{c+1}\right)^{N+1}$. If c were arbitrary, this could tend to infinity, so a blind estimate will not help (take for example $x = -0.9$ and $c = -0.8$). Therefore, let us instead restrict attention to $x \in \left(-\frac{1}{2}, \frac{1}{2}\right)$. In this case, $\frac{|x|}{c+1} < 1$ so that $\lim_{N\to\infty} \left|E_N(x)\right| = 0$. We can also compute c explicitly. Setting $\frac{f^{(N+1)}(c)}{(N+1)!} x^{N+1} = \ln(x+1) - \sum_{n=1}^{N} \frac{(-1)^n + 1}{n} x^n$ gives $(c+1)^{N+1} = \dfrac{(-1)^N x^{N+1}}{(N+1)\left(\ln(x+1) - \sum_{n=1}^{N} \frac{(-1)^n+1}{n} x^n\right)}$ so that

$$c = \frac{x}{\left[(N+1)\left(\ln(x+1) - \sum_{n=1}^{N} \frac{(-1)^n+1}{n} x^n\right)\right]^{1/(N+1)}} - 1.$$

The successor of this theorem is the Taylor series expansion for an infinitely differentiable function. It states that under suitable conditions, a large class of infinitely differentiable functions can be approximated very well by a power series whose coefficients are the higher derivatives of a function *at a single point*!

Exercises

Exercise 1. Let $f : [a, b] \to \mathbb{R}$ be a differentiable function.

i. Show that if $f'(x) \neq 0$ for all $x \in [a, b]$ then f is one-to-one. Hint: prove the contrapositive (this means to prove "not B implies not A" instead of directly proving "A implies B"—the two are equivalent), i.e. prove that "if f is not one-to-one, then there exists an $x \in [a, b]$ such that $f'(x) = 0$." Second hint: use Rolle's theorem.

ii. Provide an example to show that the converse statement need not be true. Namely, give an example of a one-to-one function $f : [a, b] \to \mathbb{R}$ for which there exists a point $x \in [a, b]$ with $f'(x) = 0$.

Exercise 2. Let $f : [a, c] \to \mathbb{R}$ be a differentiable function such that $f'(x) = m$ for all $x \in [a, c]$ for some constant $m \in \mathbb{R}$. Prove that there exists a $b \in \mathbb{R}$ such that $f(x) = mx + b$ for all $x \in [a, c]$. Hint: Define the function $g : [a, b] \to \mathbb{R}$ by $g(x) := mx$. What is $(f - g)'$ equal to? What does this say about $f - g$?

Exercise 3. Let $L, R \in \mathbb{R}$ with $L < R$ and let $f : [L, R] \to \mathbb{R}$ be a differentiable function such that $f'(x) = 2a_2 x + a_1$ for all $x \in [L, R]$ for some constants $a_2, a_1 \in \mathbb{R}$. Prove that there exists an $a_0 \in \mathbb{R}$ such that $f(x) = a_2 x^2 + a_1 x + a_0$ for all $x \in [L, R]$. Hint: look at the hint from the previous exercise and define a suitable function $g : [L, R] \to \mathbb{R}$.

Exercise 4. Let $L, R \in \mathbb{R}$ with $L < R$ and let $f : [L, R] \to \mathbb{R}$ be a differentiable function such that $f'(x) = \sum_{k=1}^{n} k a_k x^k$ for all $x \in [L, R]$ for some constants $a_1, a_2, \dots, a_n \in \mathbb{R}$. Prove that there exists an $a_0 \in \mathbb{R}$ such that $f(x) = \sum_{k=0}^{n} k! a_k x^k$ for all $x \in [L, R]$. Hint: look at the hint from the previous exercise and define a suitable function $g : [L, R] \to \mathbb{R}$.

Exercise 5. Is the converse of L'Hospital's rule: 0/0 case true? Namely, if $f, g : [a, b] \to \mathbb{R}$ are differentiable with $f(c) = g(c) = 0$ for some $c \in (a, b)$, does $\lim_{x\to c} \frac{f(x)}{g(x)} = L$ imply $\lim_{x\to c} \frac{f'(x)}{g'(x)} = L$?

Exercise 6. Compute the following limits.

i. $\lim_{x\to 0} \frac{\sin(x) - x}{x^3}$.

ii. $\lim_{x\to 0} \frac{\sin(x) - \left(x - \frac{1}{6}x^3\right)}{x^5}$.

iii. $\lim_{x\to 0} \frac{\sin(x) - \left(x - \frac{1}{6}x^3 + \frac{1}{120}x^5\right)}{x^7}$.

iv. $\lim_{x\to 0} \frac{\sin(x) - \sum_{n=1}^{N} \frac{(-1)^{n+1} x^{2n-1}}{(2n-1)!}}{x^{2N+1}}$ for any $N \in \mathbb{N}$. Here $k! := 1 \cdot 2 \cdot 3 \cdots (k-2) \cdot (k-1) \cdot k$ denotes k factorial.

v. $\lim_{x\to 0} \frac{\ln(1+x) - x}{x^2}$.

vi. $\lim_{x\to 0} \frac{\ln(1+x) - \left(x - \frac{x^2}{2}\right)}{x^3}$.

vii. $\lim_{x\to 0} \frac{\ln(1+x) - \left(x - \frac{x^2}{2} + \frac{x^3}{3}\right)}{x^4}$.

viii. $\lim_{x\to 0} \frac{\ln(1+x) - \sum_{n=1}^{N} \frac{(-1)^{n+1} x^n}{n}}{x^{N+1}}$.

ix. $\lim_{x\to 1} \Big((1-x)\ln\big(|\ln(x)|\big)\Big)$.

x. $\lim_{x\to 0} \frac{\ln\big(|\ln(x)|\big)}{\ln(|x|)}$.

Exercise 7. Let $(-1, \infty) \ni x \mapsto f(x) := \ln(x+1)$. Find c explicitly and compute the function E_1 for $x = \frac{1}{3}$. Do the same for E_2.

Exercise 8. Find c and E_N for any $x \in \mathbb{R}$ for the function $\cos$.

Curve plotting

Derivatives provide us with a lot of information about the shape of the graph of a function.

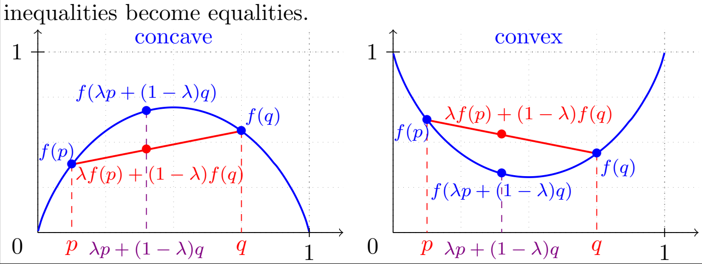
Theorem 1. *Let $f : (a, b) \to \mathbb{R}$ be a differentiable function.*

i. If $f'(x) > 0$ for all $x \in (a, b)$, then f is increasing, i.e. $f(x) < f(y)$ for $x, y \in (a, b)$ with $x < y$.

ii. If $f'(x) < 0$ for all $x \in (a, b)$, then f is decreasing, i.e. $f(x) > f(y)$ for $x, y \in (a, b)$ with $x < y$.

The proof of this theorem is another application of the Mean Value Theorem (MVT).

The second derivative of a function provides information about the curvature of the graph of that function. To state the theorem, it is useful to introduce the concept of convex and concave functions. A function is $f : (a, b) \to \mathbb{R}$ is <u>*strictly convex*</u> iff $f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y)$ for all $x, y \in (a, b)$ and for all $\lambda \in (0, 1)$. A function is $f : (a, b) \to \mathbb{R}$ is <u>*strictly concave*</u> iff $f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda)f(y)$ for all $x, y \in (a, b)$ and for all $\lambda \in (0, 1)$. If the word "strictly" is dropped, then the strict inequalities become equalities.



An example of a concave function is the *Shannon entropy* of a probability distribution. For two events, this takes the form $(0, 1) \ni p \mapsto H(p) := -p\ln(p) - (1-p)\ln(1-p)$. H can be extended continuously to $[0, 1]$ by defining $H(0) := 0$ and $H(1) := 0$.

The second derivative of a function, if non-zero, describes the concavity or convexity of a function.

Theorem 2. *Let $f : (a, b) \to \mathbb{R}$ be a twice differentiable function.*

i. If $f''(x) > 0$ for all $x \in (a, b)$, then f is strictly convex.

ii. If $f''(x) < 0$ for all $x \in (a, b)$, then f is strictly concave.

Be aware, however, that a function can be concave or convex even if its second derivative is zero at a particular point. For example, consider the function $\mathbb{R} \ni x \mapsto p_4(x) := x^4$. The second derivative of this function at $x = 0$ is 0 and yet it is convex.

Combining the previous results, we have the following incredibly useful result about first and second derivatives and optimization. It is called the <u>*second derivative test*</u>.

Theorem 3. *Let $f : (a, b) \to \mathbb{R}$ be a twice differentiable function such that f'' is continuous near $c \in (a, b)$.*

i. If $f'(c) = 0$ and $f''(c) > 0$, then f has a local minimum at c.

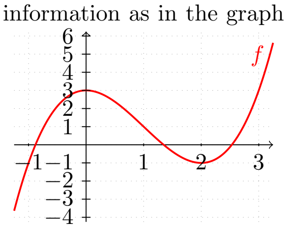
ii. If $f'(c) = 0$ and $f''(c) < 0$, then f has a local maximum at c.

Let $f : \mathbb{R} \to \mathbb{R}$ be a twice differentiable function whose second derivative is continuous and suppose that one is given information about f and its derivatives as in the following tables. A rough sketch of the graph of f can be obtained from this information as in the graph on the right.

$f(0) = 3$
$f(2) = -1$

$(-\infty, 0)$	$(0, 2)$	$(2, \infty)$
$f'(x) > 0$	$f'(x) < 0$	$f'(x) > 0$

$(-\infty, 1)$	$(1, \infty)$
$f''(x) < 0$	$f''(x) > 0$



[Reference: Herb Gross: II.7 Curve Plotting]

Even, odd, and periodic functions

Define $\mathbb{R} \ni x \mapsto r(x) := -x \in \mathbb{R}$ be the function that reflects all numbers across 0. A function $f : \mathbb{R} \to \mathbb{R}$ is <u>*even*</u> iff $f \circ r = f$ and <u>*odd*</u> iff $f \circ r = r \circ f$. Given a real number $L \in \mathbb{R}$, defined $\mathbb{R} \ni x \mapsto s_L(x) := x + L \in \mathbb{R}$ be the function that shifts all numbers by L. A function $f : \mathbb{R} \to \mathbb{R}$ is <u>*periodic*</u> iff there exists a number $L \in \mathbb{R}$ such that $f \circ s_L = f$. The smallest positive number L such that $f \circ s_L = f$ is called the <u>*period*</u> of f. Knowing when a function is even, odd, or periodic will be incredibly helpful when we discuss definite integrals.

Exercises

Exercise 1. Prove the Shannon entropy for probability distributions on two events is convex by using the definition. Then prove this using calculus by taking the second derivative. Which proof is easier?

Exercise 2. If $f : (a, b) \to \mathbb{R}$ is a twice differentiable function that is strictly convex on all of (a, b). Does this imply $f'(x) > 0$ for all $x \in (a, b)$? Explain.

Exercise 3. Show that if a function $f : \mathbb{R} \to \mathbb{R}$ is odd, then $f(0) = 0$. Check that the function defined by the formula $h(x) = \frac{1}{x}$ is odd. Do these two facts contradict each other? Explain.

Exercise 4. Show that if a function is both even and odd, then it is the 0 function.

Exercise 5. Show that if a function is periodic and convex, then it is constant.

Exercise 6. Show that if a function $f : \mathbb{R} \to \mathbb{R}$ is periodic and $\lim_{x \to \infty} f(x)$ exists, then f is constant.

Exercise 7. Show that if a function $f : \mathbb{R} \to \mathbb{R}$ is even and differentiable at 0, then $f'(0) = 0$.

Exercise 8. Show that if a function is periodic and differentiable, then the derivative is also periodic. If the derivative of a function is periodic, is the function periodic? Explain.

Exercise 9. Find a degree 3 polynomial $\mathbb{R} \ni x \mapsto p(x) := a_0 + a_1 x + a_2 x^2 + a_3 x^3$ whose values and derivatives match those of f in the table after Theorem 3 in this handout. Are there multiple solutions to this problem? If so, find another polynomial $\mathbb{R} \ni x \mapsto q(x) := b_0 + b_1 x + b_2 x^2 + b_3 x^3$ satisfying the conditions in that table. Otherwise, prove that your polynomial is the unique one satisfying these conditions.

Exercise 10. The *ideal gas law* is a relationship between the pressure P, temperature T (in Kelvin), the volume V, and the number n of moles of an ideal gas in some container. It is given by $PV = nRT$, where R is a constant, called the universal gas constant. At constant temperature T, the pressure is inversely proportional to the volume $P = \frac{nRT}{V}$. This function has no critical points on its domain $(0, \infty)$. Furthermore, it is positive, convex, and has limits given by $\lim_{V \to 0} P(V) = \infty$ and $\lim_{V \to \infty} P(V) = 0$ (verify this!). Not all gases can be adequately modeled as an ideal gas. For this reason, other models are needed. One such model is the *Van der Waals gas law*. It is given by $\left(P + \frac{an^2}{V^2}\right)(V - bn) = nRT$, where a and b are positive constants (which are computed empirically by data fitting and therefore depend on the gas). Solve this equation for P and compare these two functions. For example, discuss the critical points, concavity, etc. of the Van der Walls pressure in terms of the volume (assume that T is constant). Let V_c be the volume, P_c the pressure, and T_c the temperature at which $P(V_c) = P_c$, $P'(V_c) = 0$, and $P''(V_c) = 0$. Solve for a and b in terms of V_c, T_c, and P_c. Rewrite P in terms of these other constants. Compare and contrast the graphs of P depending on the three possibilities: i. $T > T_c$ ii. $T = T_c$ and iii. $T < T_c$.

Optimization

An *optimization problem* consists of a function $f : \mathbb{R}^n \to \mathbb{R}$ subject to *constraints* given by level sets of functions $g_1, \dots, g_k : \mathbb{R}^n \to \mathbb{R}$ for values $c_1, \dots, c_k \in \mathbb{R}$, respectively. A *solution of an optimization problem* is a vector $\vec{x} := (x_1, \dots, x_n) \in \mathbb{R}^n$ for which $g_1(\vec{x}) = c_1, \dots, g_k(\vec{x}) = c_k$ satisfying the following maximization (or minimization) condition: for any other vector $\vec{y} \in \mathbb{R}^n$ satisfying $g_1(\vec{y}) = c_1, \dots, g_k(\vec{y}) = c_k$, then $f(\vec{x}) \geq f(\vec{y})$ (or $f(\vec{x}) \leq f(\vec{y})$). The variables $x_1, \dots, x_n$ represent some physical parameters while the level sets $g_1^{-1}(\{c_1\}), \dots, g_k^{-1}(\{c_k\})$ describe constraints that these parameters must satisfy. An optimization solution is a set of parameters that maximizes (or minimizes) f but are still within the level sets. The same definition makes sense if the domains of $f, g_1, \dots, g_k$ are replaced by subsets of $\mathbb{R}^n$.

If $k = n - 1$ and we are fortunate enough, we can use the constraint equations to solve for all of the variables in terms of a single one, say x_1. Then $f(x_1, \dots, x_n) = f(x_1, x_2(x_1), \dots, x_n(x_1))$. If f and all of the functions $x_2, \dots, x_n$ of x_1 are differentiable, maximizing or minimizing this function is obtained by finding the critical points by taking the derivative and setting it equal to zero, i.e. $\frac{d}{dx_1} f(x_1, x_2(x_1), \dots, x_n(x_1)) = 0$. One then analyzes the critical points to determine if they are maxima or minima.

If $k \neq n - 1$ or if it is not easy (or possible) to solve for all of the parameters in terms of a single parameter, then other techniques are needed. This is covered in multivariable calculus under the topic "Lagrange multipliers."

A cylindrical can is made to hold L liters of a liquid. What radius and height would minimize the cost needed to manufacture the can? Let r and h denote the radius and height, respectively. The cost is proportional to the surface area of the can, which is given by $A(r, h) := 2\pi r^2 + 2\pi r h$, where the first term represents the top and bottom lid while the second term represents the cylindrical shell. Since the cylindrical can holds L liters, its volume is constant. Hence, the level set $V^{-1}(\{L\})$ of the volume function $(0, \infty) \times (0, \infty) \ni (r, h) \mapsto V(r, h) := \pi r^2 h$ defines the constraint. In terms of the notation from the definition, $n = 2$, $x_1 = r$, $x_2 = h$, $f = A$, $k = 1$, $g_1 = V$, and $c_1 = L$. The constraint equation takes the form $\pi r^2 h = L$ which allows us to solve for h since it is given by $h = \frac{L}{\pi r^2}$. Note that L liters $= L$ cubic decimeters so that we do not have to include any factors of 10. We just have to express our answers in terms of decimeters. Therefore, $A(r, h(r)) = 2\pi r^2 + 2\pi r \left(\frac{L}{\pi r^2}\right) = 2\pi r^2 + \frac{2L}{r}$ and its derivative is $\frac{d}{dr} A(r, h(r)) = 4\pi r - \frac{2L}{r^2}$. r_o is a critical point of A if and only if $4\pi r_o - \frac{2L}{r_o^2} = 0$. Solving for r_o gives $r_o = \left(\frac{L}{2\pi}\right)^{1/3}$ decimeters. r_o is a minimum for many reasons:

i. $A''(r_0, h(r_0)) > 0$,

ii. $\lim\limits_{r \to 0} A(r, h(r)) = \infty$ and $\lim\limits_{r \to \infty} A(r, h(r)) = \infty$ and continuity imply A attains a minimum,

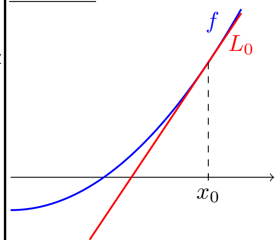
iii. $A'(r_0 + \epsilon, h(r_0 + \epsilon)) > 0$ and $A'(r_0 - \epsilon, h(r_0 - \epsilon)) < 0$ for sufficiently small $\epsilon > 0$.

Plugging r_o into h gives $h(r_o) = \frac{L}{\pi}\left(\frac{L}{2\pi}\right)^{-2/3} = \left(\frac{L}{4\pi}\right)^{1/3}$ decimeters.

Fix $a, b > 0$. What is the area of the largest rectangle that can be inscribed in the ellipse given by the level set equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$? Let (x, y) be one of the corners of the rectangle, which we may assume is on the ellipse and is in the first quadrant. Solving for y as a function of x gives $y = b\sqrt{1 - \frac{x^2}{a^2}}$. The area of the rectangle is therefore $A = 4xy = 4bx\sqrt{1 - \frac{x^2}{a^2}}$. Maximizing this gives $0 = \frac{dA}{dx} = 4b\sqrt{1 - \frac{x^2}{a^2}} - \frac{4bx^2}{a^2\sqrt{1 - \frac{x^2}{a^2}}}$. Multiplying by $\sqrt{1 - \frac{x^2}{a^2}}$ and rearranging gives $1 - \frac{x^2}{a^2} = \frac{x^2}{a^2}$ so $x = \frac{a}{\sqrt{2}}$. Hence $y = b\sqrt{1 - \frac{1}{2}} = \frac{b}{\sqrt{2}}$ so that the largest area is $A = xy = \frac{ab}{2}$.

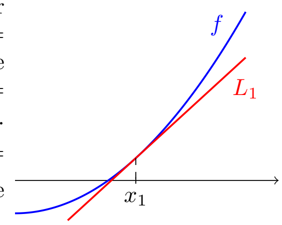
Newton's method

Let $f : \mathbb{R} \to \mathbb{R}$ be a function and let $x_0 \in \mathbb{R}$. *Newton's approximation sequence* is the sequence of real numbers $x_1, x_2, \dots$ defined as follows.



Let $L_0 : \mathbb{R} \to \mathbb{R}$ be the linear approximation to f at x_0, i.e. $L_0(x) = f'(x_0)x + \big(f(x_0) - f'(x_0)x_0\big)$. Let x_1 be the root of this line, i.e. x_1 satisfies $0 = L_0(x_1) = f'(x_0)x_1 + \big(f(x_0) - f'(x_0)x_0\big)$. Solving for x_1 gives $x_1 = \frac{f'(x_0)x_0 - f(x_0)}{f'(x_0)} = x_0 - \frac{f(x_0)}{f'(x_0)}$. Note that $f'(x_0)$ must not be zero in order for x_1 to be defined.

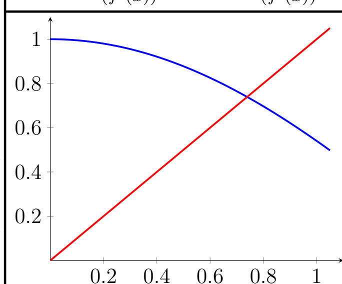
Similarly, let $L_1 : \mathbb{R} \to \mathbb{R}$ be the linear approximation to f at x_1, i.e. $L_1(x) = f'(x_1)x + \big(f(x_1) - f'(x_1)x_1\big)$. Let x_2 be the root of this line, i.e. x_2 satisfies $0 = L_1(x_2) = f'(x_1)x_2 + \big(f(x_1) - f'(x_1)x_1\big)$. Solving for x_2 gives $x_2 = \frac{f'(x_1)x_1 - f(x_1)}{f'(x_1)} = x_1 - \frac{f(x_1)}{f'(x_1)}$. Note that $f'(x_1)$ must not be zero in order for x_2 to be defined.



We continue in this fashion, which requires the derivatives of f evaluated at x_n to always be nonzero since $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$. Under suitable conditions, the sequence converges $\lim\limits_{n \to \infty} x_n$ to a root of the function f. I am not aware of necessary and sufficient conditions for this to happen, but here are some thoughts. Let $[a, b] \subseteq \mathbb{R}$ be a closed interval on which f has a single root and suppose $f'(x) \neq 0$ for all $x \in [a, b]$. Define $[a, b] \ni x \mapsto g(x) := x - \frac{f(x)}{f'(x)}$. In order to have some control over the successive approximations to the root, it is sufficient to demand the image of g to be in $[a, b]$ as well. In addition, notice that if x_r is the root of f, then $g(x_r) = x_r$.

Theorem 1. *Let $f : \mathbb{R} \to \mathbb{R}$ and let $[a, b] \subseteq \mathbb{R}$ be a closed interval containing a single root of f. Suppose that f is twice differentiable with $f'(x) \neq 0$ for all $x \in [a, b]$ and suppose there exists a number $k \in [0, 1)$ such that $\left|\frac{f(x)f''(x)}{(f'(x))^2}\right| \leq k$ for all $x \in [a, b]$. Then for any $x_0 \in [a, b]$, the sequence $x_0, x_1 := g(x_0), x_2 := g(g(x_0)), \dots$ converges to the unique root of f.*

Proof. This is a consequence of Exercise 18.23 in my notes on Analysis II applied to the function $[a, b] \ni x \mapsto g(x) := x - \frac{f(x)}{f'(x)}$ since $g'(x) = 1 - \frac{(f'(x))^2 - f''(x)f(x)}{(f'(x))^2} = \frac{f''(x)f(x)}{(f'(x))^2}$. ■



The function $\mathbb{R} \ni x \mapsto \cos(x) - x$ has a single root. To find the root of f up to a certain number of decimal places, make a guess such as $x_0 = 0.8$ and plug into the function $g(x) := x - \frac{f(x)}{f'(x)} = x + \frac{\cos(x) - x}{\sin(x) + 1}$. By iterating this, we get the sequence $x_0, x_1 = g(x_0), x_2 = g(g(x_0)), \dots$, the first values of which are given below.

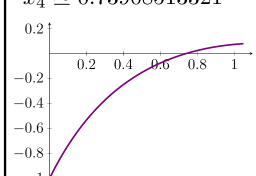
$x_0 = 0.8$

$x_1 \approxeq 0.73985330637$

$x_2 \approxeq 0.73908526340$

$x_3 \approxeq 0.73908513321$

$x_4 \approxeq 0.73908513321$



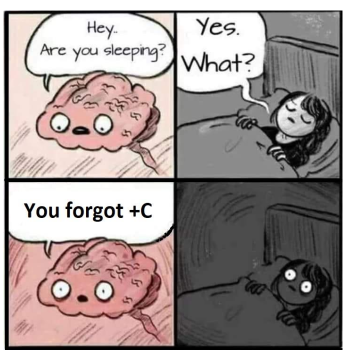
This shows that even after four iterations, the root of f is 0.73908513321 up to 11 decimal places. Note that if you had chosen a different initial starting position for x_0, your answer is still *guaranteed* to converge to the root because $g'(x) = \frac{\cos(x)(x - \cos(x))}{(1 + \sin(x))^2}$ provided that the initial starting point is chosen in $(0, 1]$. This is because $|g'(x)| \leq k$ for some $k \in (0, 1)$ for all $x \in (0, 1]$. Notice $g'(0) = -1$ so this condition is not satisfied. Theorem 1 is not valid for x satisfying $|g'(x)| \geq 1$.

Antiderivatives

Let $f : \mathbb{R} \to \mathbb{R}$ be a function. An *antiderivative* of f is a differentiable function $F : \mathbb{R} \to \mathbb{R}$ such that $F' = f$. Note that if F and G are antiderivatives of f, then by Theorem 4 from Week #08, $F - G$ is a constant. In most applications, this constant is important.

If differentiation is viewed as a function whose domain is the set of differentiable functions, we can ask if there is an inverse. Let $\mathcal{S}$ be the set of differentiable functions on $\mathbb{R}$, let $\mathcal{T}$ be the set of all functions on $\mathbb{R}$, and let $\mathcal{D} : \mathcal{S} \to \mathcal{T}$ be the assignment sending any differentiable function f to its derivative $\mathcal{D}(F) := F'$. In this notation, the linearity of differentiation means $\mathcal{D}(F + G) = \mathcal{D}(F) + \mathcal{D}(G)$ and $\mathcal{D}(cF) = c\mathcal{D}(F)$ for all $F, G \in \mathcal{S}$ and $c \in \mathbb{R}$. Let $\mathcal{R}$ be the range of $\mathcal{D}$. Then $\mathcal{D} : \mathcal{S} \to \mathcal{R}$ is onto. Therefore, there exists a function $\mathcal{I} : \mathcal{R} \to \mathcal{S}$ such that $\mathcal{D} \circ \mathcal{I} = \mathrm{id}_{\mathcal{R}}$ (every surjective function has a right inverse—take this for granted). Is $\mathcal{D}$ one-to-one? No, because if F and G satisfy $\mathcal{D}(F) = \mathcal{D}(G)$, i.e. $F' = G'$, then $F - G$ is a constant. Therefore, for any function $f \in \mathcal{R}$ let F be *any* antiderivative of f. Then the level set $\mathcal{D}^{-1}(\{f\}) = \{F \in \mathcal{S} \ : \ F' = f\}$ is given by $\mathcal{D}^{-1}(\{f\}) = \{F + C \ : \ C \in \mathbb{R},\ F' = f\}$. Putting all of this together says that the set of *all* solutions F to $D(F) = f$ is equal to a particular solution, some differentiable function F such that $F' = f$, plus any constant. [Reference: Herb Gross: II.10 Inverse Differentiation and Herb Gross: II.11 The "Definite" Indefinite Integral].

One of the theorems on derivatives we discovered is the power rule which says $p_n' = np_{n-1}$, where $p_k(x) := x^k$ for all $k \in \{0, 1, 2, \dots\}$. Therefore, we can ask: "what is the set of all differentiable functions F such that $F' = p_n$?" By thinking backwards, we know $\mathcal{D}(p_{n+1}) = (n+1)p_n$ so solving for p_n gives $p_n = \frac{1}{n+1}\mathcal{D}(p_{n+1}) = \mathcal{D}\left(\frac{p_{n+1}}{n+1}\right)$. Therefore, one particular solution is $\frac{p_{n+1}}{n+1}$. Therefore, the general solution to $F' = p_n$ is the set $\left\{\frac{p_{n+1}}{n+1} + C \ : \ C \in \mathbb{R}\right\}$. For example, if $n = 4$, then we want a function F such that $F'(x) = x^4$. Since $\frac{d}{dx}x^5 = 5x^4$, we divide by 5 to obtain $\frac{d}{dx}\left(\frac{x^5}{5}\right) = x^4$ but also $\frac{d}{dx}\left(\frac{x^5}{5} + C\right) = x^4$ for any $C \in \mathbb{R}$.



Let us examine why $+C$ is so important. Imagine you drop a rock off the top of a cliff at the beginning of your timer. You know the velocity of the rock as it falls is given by $v(t) = -gt$, where $g = 9.81$ meters per second. What is the height of the rock as measured by someone at the bottom of the cliff at time t? To find the height, we solve $\frac{dx}{dt} = -gt$. One solution is $x(t) = -\frac{gt^2}{2}$ but if we plug in $t = 0$, we get $x(0) = 0$, which is not possible. Where is the rock at time $t = 0$? Obviously, it's at the top of the cliff. If we include $+C$ in our antiderivative, we get $x(t) = -\frac{gt^2}{2} + C$. If h is the height of the cliff above the ground, then $x(0) = h$. This forces $C = h$ in this case. Therefore, $x(t) = -\frac{gt^2}{2} + h$.

Imagine a force acting on a particle of mass m subject to motion along a straight line. By Newton's force equation $F = ma$, one can solve for $a = \frac{F}{m}$. Since $a = \frac{dv}{dt} = \frac{d^2x}{dt^2}$ is the second derivative of the displacement, one can solve for the displacement as a function of time. For example, suppose $F(t) = k\cos(t)$ for all $t \in \mathbb{R}$ with k some constant and the initial position of the mass is x_0 and the initial velocity is v_0. Antidifferentiating $\frac{k\cos(t)}{m}$ once gives $v(t) = \frac{k\sin(t)}{m} + c_1$. At $t = 0$, we have $v_0 = v(0) = c_1$ so $v(t) = \frac{k\sin(t)}{m} + v_0$. Antidifferentiating again gives $x(t) = \frac{-k\cos(t)}{m} + v_0 t + c_2$. At $t = 0$, we have $x_0 = -\frac{k}{m} + c_2$ so $c_2 = x_0 + \frac{k}{m}$. Therefore, $x(t) = x_0 + v_0 t + \frac{k}{m}\big(1 - \cos(t)\big)$ for all $t \in \mathbb{R}$.

Exercises

Exercise 1. Find a variety of five or six cans and measure them (in units of decimeters). Compute their area and volume. Then compute the optimal radius, height, and area for that given volume. Make a table of your results and discuss the significance of your findings.

Exercise 2. Fix $r > 0$. What is the area of the largest isosceles triangle that can be inscribed in a circle of radius r? Recall, an isosceles triangle is a triangle for which two sides have the same length.

Exercise 3. Fix $r > 0$. Consider the following problem: "What is the area of the largest triangle that can be inscribed in a circle of radius r?" Set up the problem to identify the function you are optimizing. Identify all of the variables that appear in the equation. Identify all of the constraints. Are there enough constraints for you to solve this problem using the techniques introduced here? Explain.

Exercise 4. Fix $a, b > 0$. What is the area of the largest isosceles triangle that can be inscribed in the ellipse given by the level set equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$?

Exercise 5. Using Newton's method of approximation, find $x \in \mathbb{R}$ such that $e^{-x} = x$ to six decimal points.

Exercise 6. Does Newton's method work for finding the root of $\mathbb{R} \ni x \mapsto \arctan(1 + x) - x$? If yes, find it. Else, explain.

Exercise 7. Apply Newton's method to find $x \in (0, 1)$ such that $x^2\tan(x) = x$.

Exercise 8. Let $\mathcal{S}$ be the set of differentiable functions, let $\mathcal{D}$ be the assignment sending a differentiable function to its derivative, let $\mathcal{R}$ be the range of $\mathcal{D}$ and let $\mathcal{I}$ be a choice of right inverse for $\mathcal{D}$, i.e. $\mathcal{I} : \mathcal{R} \to \mathcal{S}$ satisfies $\mathcal{D} \circ \mathcal{I} = \mathrm{id}_{\mathcal{R}}$ (one does this by making a choice for the antiderivative for any function in $\mathcal{R}$).

i. Let $f, g \in \mathcal{R}$. Prove $\mathcal{I}(f + g) - \big(\mathcal{I}(f) + \mathcal{I}(g)\big)$ is some constant.

ii. Let $f \in \mathcal{R}$ and $c \in \mathbb{R}$. Prove $\mathcal{I}(cf) - c\mathcal{I}(f)$ is a constant.

Hint: Apply $\mathcal{D}$ to both of these equations and use linearity.

Exercise 9. Fix $a, m > 0$. Find an equation relating a and m such that the graphs of $f(x) := e^{ax}$ and $g(x) := mx$ intersect in exactly one point. Can you use Newton's method to find the point x at which these two functions intersect? If so, express x as a function of a. If not, do so when $a = 1$.

Exercise 10. If the force on a particle of mass m is $F = k$ for some constant k, and if the initial position is x_0 and the initial velocity is v_0, find the displacement $x(t)$ as a function of time t. Describe one situation where the force is constant in this way.

Exercise 11. Find the general antiderivatives of the following functions. If f_k denotes the function given, F_k denotes its antiderivative. Then find the exact antiderivative if the value of the antiderivative is given at the point specified.

i. $\mathbb{R} \ni x \mapsto f_1(x) := |x|$. $F_1(2) = 5$.

ii. $\mathbb{R} \ni x \mapsto f_2(x) := \sqrt{|x|}$. $F_2(1) = 1$.

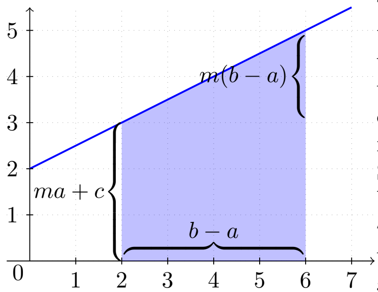
iii. $\mathbb{R} \ni x \mapsto f_3(x) := \begin{cases} 0 & \text{if } x = 0 \\ x/|x| & \text{otherwise} \end{cases}$. $F_3(0) = -1$.

iv. $\mathbb{R} \ni x \mapsto f_4(x) := x\cos(x^2)$. $F_4(0) = 1$.

v. $\mathbb{R} \ni x \mapsto f_5(x) := 0$. $F_5(0) = 17$.

For example, if $\mathbb{R} \ni x \mapsto f(x) := x$ and $F(1) = 1$ then the general antiderivative is $F(x) = \frac{x^2}{2} + C$ with C an arbitrary constant. The specific solution is obtained from $1 = F(1) = \frac{1}{2} + C$ so $C = \frac{1}{2}$ so the specific antiderivative is $F(x) = \frac{x^2}{2} + \frac{1}{2}$.

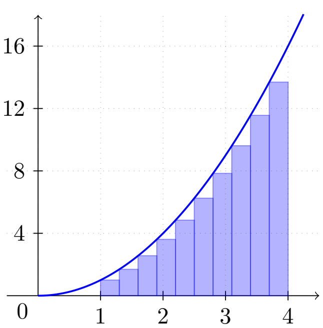
Area and distance

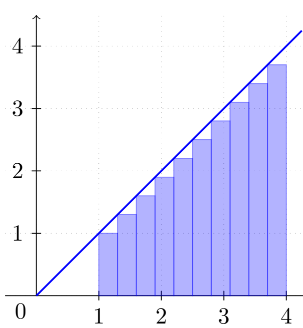


To compute the shaded area under the curve on the left, we could break up the shaded trapezoid into two parts and compute the area of the resulting triangle and the rectangle. Suppose the curve is described by the line $y(x) = mx + c$ with $m, b \geq 0$. Suppose we wanted to know the area under it from $x = a$ to $x = b$.

The dimensions of the resulting triangle and rectangle are shown above. The area is therefore $(b-a)(ma+c) + \frac{1}{2}(b-a)(m(b-a)) = \frac{1}{2}m(b^2-a^2) + c(b-a)$. Interestingly, notice that this expression is the difference of the antiderivative of $f(x) = mx + c$ evaluated at $x = b$ and $x = a$. Indeed, the antiderivative is $F(x) = \frac{1}{2}mx^2 + cx + C$ (where C is an unknown constant) and the area equals $F(b) - F(a)$ (notice that the constant C gets cancelled). Also notice that the area under the curve is the area under the curve $\mathbb{R} \ni x \mapsto mx$ plus the area under the curve $\mathbb{R} \ni x \mapsto c$ between $x = a$ and $x = b$.

How do we compute areas under curves that are not of the form $y(x) = mx + c$? For example, how can we compute the area under the curve $f(x) = x^2$ from $x = a$ to $x = b$? One method is to approximate the area by shapes for which the area formula takes a particularly simple form. This occurs for rectangles. Using more rectangles will provide a more accurate formula. To do this calculation, we will revisit computing the area under a straight line curve as well.



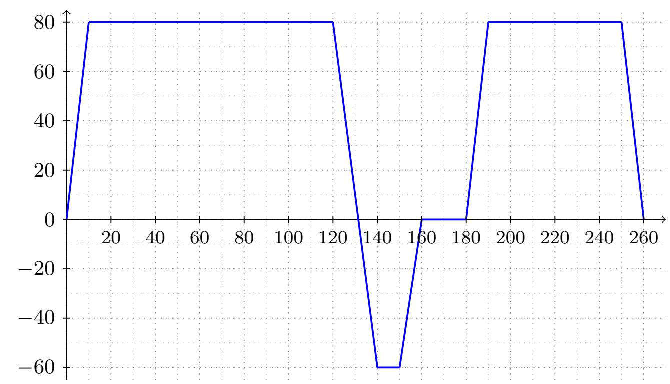


To compute the area under the straight line curve $g(x) := x$ from $x = a$ to $x = b$, set $L := b - a$ and fix $n \in \mathbb{N}$ to be the number of rectangles we will use. If we demand the width to be the same for all of them, then they are L/n in width. The bottom left corners of the rectangles are given by $a + \frac{kL}{n}$ as k varies from 0 to $n-1$ while the top left corners are given by the same formula. If we use $f(x) := x^2$, then the top left corners are given by $\left(a + \frac{kL}{n}\right)^2$. The area of a given such rectangle is $\left(\frac{L}{n}\right)\left(a + \frac{kL}{n}\right)$ for g and $\left(\frac{L}{n}\right)\left(a + \frac{kL}{n}\right)^2$ for f. The sum of the areas of the rectangles is therefore $A_n(g;a,b) := \sum_{k=0}^{n-1}\left(\frac{L}{n}\right)\left(a + \frac{kL}{n}\right) = \frac{La}{n}\sum_{k=0}^{n-1} 1 + \frac{L^2}{n^2}\sum_{k=0}^{n-1} k = \frac{Lan}{n} + \frac{L^2 n(n-1)}{2n^2} = L\left(a + \frac{L}{2}\left(\frac{n-1}{n}\right)\right)$. Therefore, $\lim_{n\to\infty} A_n = L\left(a + \frac{L}{2}\right) = \frac{b^2}{2} - \frac{a^2}{2}$, which is what we expected. Now, for the $f(x) = x^2$ curve, we have $A_n(f;a,b) := \sum_{k=0}^{n-1}\left(\frac{L}{n}\right)\left(a + \frac{kL}{n}\right)^2 = \frac{La^2}{n}\sum_{k=0}^{n-1} 1 + \frac{2L^2 a}{n^2}\sum_{k=0}^{n-1} k + \frac{L^3}{n^3}\sum_{k=0}^{n-1} k^2 = \frac{La^2 n}{n} + \frac{2L^2 an(n-1)}{2n^2} + \frac{L^3(n-1)n(2n-1)}{6n^3}$ so $\lim_{n\to\infty} A_n(f;a,b) = La^2 + L^2 a + \frac{L^3}{3} = \frac{b^3}{3} - \frac{a^3}{3}$.

To compute the area under the graph of the exponential function from a to b, we apply a similar procedure: $A_n(\exp;a,b) := \sum_{k=0}^{n-1}\left(\frac{L}{n}\right)\exp\left(a + \frac{kL}{n}\right) = \frac{Le^a}{n}\sum_{k=0}^{n-1}\left(e^{L/n}\right)^k = \frac{Le^a}{n}\left(\frac{1 - e^{L(n-1)/n}}{1 - e^{L/n}}\right)$. To compute the limit $\lim_{n\to\infty} A_n(\exp;a,b)$, we will first compute $\lim_{n\to\infty} n(1 - e^{L/n})$. For this, it is convenient to turn this limit into one involving a continuous variable so L'Hospital's rule can be applied. This gives $\lim_{n\to\infty} n(1 - e^{L/n}) = \lim_{x\to 0}\frac{1 - e^{Lx}}{x} = \lim_{x\to 0}\frac{-Le^{Lx}}{1} = -L$. Therefore, $\lim_{n\to\infty} A_n(\exp;a,b) = -e^a(1 - e^L) = -e^a(1 - e^{b-a}) = e^b - e^a$.

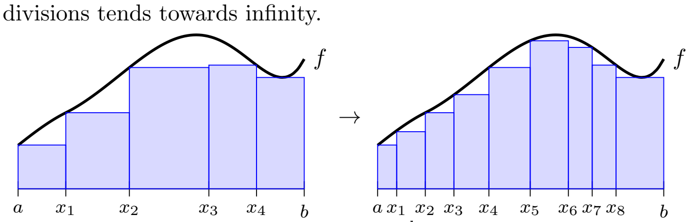
Vicky drove a hovering vehicle from NYC to Boston along a straight line and her velocity was tracked at all times and is shown on the graph. At one point during the trip, she passed by a Wendy's. A few minutes after passing, she regretted not stopping. She slows down and reverses direction to pick up some nuggets and two junior bacon cheeseburgers. Then she proceeds to go towards Boston. What is her displacement from NYC during the following times?

i. 2 hours
ii. 2 hours and 40 minutes
iii. 3 hours and 20 minutes



The definite integral

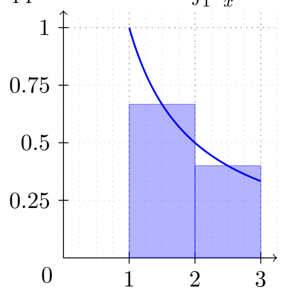
The previous example shows the importance of signs when computing areas under curves. If a function is negative over a given interval, then its area will be assigned a negative value. Let $f : [a,b] \to \mathbb{R}$ be a piecewise continuous function over an interval with at most finitely many points of discontinuity. Then $\lim_{n\to\infty}\sum_{k=0}^{n-1}\left(\frac{b-a}{n} f\left(a + \frac{k(b-a)}{n}\right)\right)$ converges. Its value is called the *definite integral* of f over $[a,b]$ and is denoted by $\int_a^b f(x)\,dx$. a and b are called the *limits of integration*. Such a limit might or might not exist for functions that are not piecewise continuous. In such cases, one can generalize the summation to allow for rectangles of arbitrary widths and evaluated at arbitrary points. One then defines the integral by summing the areas for successively finer *partitions* and taking the limit as the number of divisions tends towards infinity.

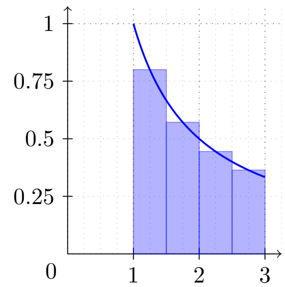


A function $f : [a,b] \to \mathbb{R}$ for which $\int_a^b f(x)\,dx$ converges is called *(Riemann) integrable* over $[a,b]$. It is not easy to identify the class of functions that are integrable. [Reference: Herb Gross: II.11 The "Definite" Indefinite Integral and Herb Gross: IV.1 The Definite Integral].

One of the immediate properties of the integral of an integrable function f over an interval $[a,b]$ is for any $c \in (a,b)$, f is integrable over $[a,c]$ and $[c,b]$ and the integrals are related by $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$. It is useful to define $\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$ so that interchanging limits reverses the sign of the area. Furthermore, if g is another integrable function over $[a,b]$ and $\lambda \in \mathbb{R}$ then $\int_a^b \big(f(x)+g(x)\big)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$ and $\int_a^b \lambda f(x)\,dx = \lambda \int_a^b f(x)\,dx$. If $f(x) \le g(x)$ for all $x \in [a,b]$, then $\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$. If $L \le f(x) \le M$ for all $x \in [a,b]$ for some numbers $L, M \in \mathbb{R}$, then $L(b-a) \le \int_a^b f(x)\,dx \le M(b-a)$.

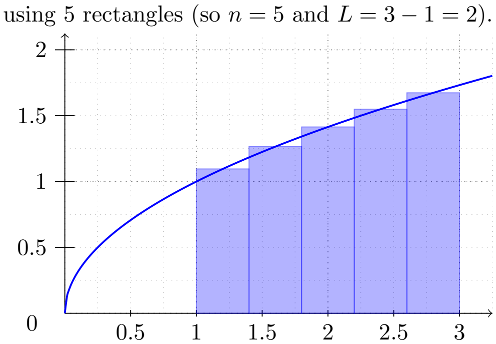
Sometimes computing the integral from the definition is challenging. For instance, what is the integral of the function $(0,\infty) \ni x \mapsto \frac{1}{x}$ on an interval $[a,b]$ with $0<a<b$? Using the definition, it is given by $\displaystyle\int_a^b \frac{1}{x}\,dx = \lim_{n\to\infty}\sum_{k=0}^{n-1}\left(\frac{L}{n}\right)\left(\frac{1}{a+\frac{kL}{n}}\right)$. One option is to approximate and use midpoints instead of the left endpoint as we have been doing. The resulting sum for an arbitrary function f would look like $\displaystyle A_n(f;a,b) := \sum_{k=0}^{n-1}\left(\frac{L}{n}\right) f\left(a+\frac{(2k+1)L}{2n}\right)$ For example, below are two approximations of $\int_1^3 \frac{1}{x}\,dx$.





The first one gives $\frac{2}{3}+\frac{2}{5}=\frac{16}{15}\approx 1.06667$ while the second one gives $\frac{1}{2}\left(\frac{4}{5}+\frac{4}{7}+\frac{4}{9}+\frac{4}{11}\right)=\frac{3782}{3465}\approx 1.09149$. The exact value, as we will derive next week, is $\int_1^3 \frac{1}{x}\,dx = \ln(3)\approx 1.09861$.

Let's use the midpoint approximation method to approximate $\int_1^3 \sqrt{x}\,dx$ using 5 rectangles (so $n=5$ and $L=3-1=2$).



The sum of the areas of these rectangles is $\displaystyle\sum_{k=0}^{4}\left(\frac{3-1}{5}\right)\sqrt{1+\frac{(2k+1)(3-1)}{2(5)}} = \left(\frac{2}{5}\right)^{3/2}\sum_{k=0}^{4}\sqrt{3+k} = \left(\frac{2}{5}\right)^{3/2}\left(\sqrt{3}+\sqrt{4}+\sqrt{5}+\sqrt{6}+\sqrt{7}\right) \approx 2.79883$. Compared to the actual definite integral, which is (as we will see next week) $\frac{2}{3}\left(3^{3/2}-1^{3/2}\right)\approx 2.79743$, it is a decent approximation.

There are also many other approximation techniques for computing Riemann sums. Another common one is to use trapezoids instead of rectangles. All such methods yield the same direct integrals for Riemann integrable functions. Yet another method mentioned earlier involves arbitrary partitions. A vastly more general method re-examines the meaning of "volume" or "area" through the use of measures. This is covered in a course on measure theory.

Exercises

Exercise 1. Give an example of a piecewise continuous function $f : [0,1] \to \mathbb{R}$ that is not integrable. Hint: it must have an infinite number of points at which it is discontinuous.

Exercise 2. Assume $0<a<b$. Using n lower rectangles, compute the area under the graph of $\mathbb{R} \ni x \mapsto x^3$ between $x=a$ and $x=b$. Then, take the $n\to\infty$ limit of your result to compute $\int_a^b x^3\,dx$. Hint: $\displaystyle\sum_{k=0}^{n-1} k^3 = \frac{n^2(n-1)^2}{4}$.

Exercise 3. Assume $0<a<b$. Using n lower rectangles, compute the signed area under the graph of $(0,\infty)\ni x \mapsto \ln(x)$ between $x=a$ and $x=b$. How does this answer change if you use n midpoint rectangles? Approximate $\int_2^4 \ln(x)\,dx$ by using 4-midpoint rectangles. Compare your answer to the actual value $\int_2^4 \ln(x)\,dx = -2+\ln(64)\approx 2.15889$.

Exercise 4. Let $a,b\in\mathbb{R}$ with $a<b$. If $\int_a^b \ln(x)\,dx =: A$, what is $\int_a^b \ln\left(\frac{1}{x^{2018}}\right)dx$ in terms of A? What about $\int_a^b \ln(2018x)\,dx$?

Exercise 5. Let $a,b\in\mathbb{R}$ with $a<b$. If $f:\mathbb{R}\to\mathbb{R}$ is an integrable function and $c\in\mathbb{R}$ what should the limits of integration be in the integral $\int_?^{??} f(x+c)\,dx$ so that $\int_?^{??} f(x+c)\,dx = \int_a^b f(x)\,dx$?

Exercise 6. Let $a>0$. If f is an even integrable function, can the expression $\int_{-a}^a f(x)\,dx$ be simplified? If g is an odd integrable function, what does $\int_{-a}^a g(x)\,dx$ equal?

Exercise 7. A probability density function (pdf) is a function $f:\mathbb{R}\to\mathbb{R}$ such that $f(x)\ge 0$ for all $x\in\mathbb{R}$ and $\displaystyle\lim_{\substack{a\to-\infty\\ b\to\infty}}\int_a^b f(x)\,dx = 1$. Show that if g is another pdf, then $\lambda f+(1-\lambda)g$ is a pdf for all $\lambda\in[0,1]$. Such a pdf is called the convex combination of the pdf's f and g.

Exercise 8. A cannon ball is shot out of a cannon on top of a cliff of height h above sea level with initial velocity v at an angle of θ with respect to the horizontal direction. If g represents the gravitational acceleration (given by 9.81 meters per second), than the acceleration as a function of time for the cannon ball is $a(t)=-g$. Using the initial conditions given, derive a formula for the vertical displacement of the cannon ball as a function of time.

Exercise 9. Using an approximation from the first week of class, compute $\int_0^1 \sqrt{1+x^2}\,dx$. Hint: the graph of $[0,1]\ni x\mapsto\sqrt{1+x^2}$ describes one quarter of a circle. In the first week, we computed the area of a polygon inscribed inside a circle. As the number of sides of this polygon increases, its area tends towards a number. Compare that calculation to this problem.

Exercise 10. Using the midpoint approximation method, approximate $\int_{-1}^1 \exp(-x^2)\,dx$ using $n=5$ rectangles. Do the same for $\int_{-2}^2 \exp(-x^2)\,dx$ using $n=9$ rectangles. Write a program to compute the area for an arbitrary number of rectangles. Then, do the same for $\int_{-2}^2 \exp(-x^2)\,dx$ using $n=27$ rectangles. How close is your answer to $\frac{\sqrt{\pi}}{2}$? Write a general mathematical formula for the midpoint formula using n rectangles for $\int_{-a}^a \exp(-x^2)\,dx$ given an arbitrary $a>0$.

Exercise 11. Let $f:[a,b]\to\mathbb{R}$ be a continuous function. Define $g:[a,b]\to\mathbb{R}$ by $[a,b]\ni x\mapsto g(x) := \int_a^x f(x')\,dx'$ (the variable x' is used to distinguish it from the limits of integration). Is g differentiable on (a,b)? To check this, write the difference quotient from the definition of the derivative, i.e write the formula for $g'(x)$. Simplify this difference quotient by using the rule $\int_a^b = \int_a^c + \int_c^b$ for $a<c<b$. Argue why the numerator of this quotient is approximately $\epsilon f(x)$. Prove $g'(x)=f(x)$. This is an important exercise and illustrates a surprising relationship between integrals and derivatives. It is one of the fundamental theorems of calculus.

The Fundamental Theorem of Calculus

Based on the computations from last week, we noticed a surprising "coincidence." Namely, if $f : [a,b] \to \mathbb{R}$ is integrable and $F : [a,b] \to \mathbb{R}$ is any antiderivative of f, then $\int_a^b f(x)\ dx = F(b) - F(a)$. This is more than a coincidence. It is a fundamental fact of calculus that relates rates of change to areas/distances. It is so important it is called the <u>*Fundamental Theorem of Calculus*</u>. It actually has two forms which together show that differentiation and integration are in some sense inverses of each other (though technically, they aren't—we will explain exactly what we mean soon).

Theorem 1. *Let $f : [a,b] \to \mathbb{R}$ be continuous. Then the function $[a,b] \ni x \mapsto g(x) := \int_a^x f(x)\ dx$ is continuous and is differentiable when restricted to (a,b). In fact, its derivative is given by $g'(x) = f(x)$ for all $x \in (a,b)$.*

Proof. The proof of this is a direct consequence of the definition of the derivative and the properties of the integral. We will not prove this here rigorously but instead will give the idea. To prove continuity of g at c, one must show $\displaystyle\lim_{x\to c}\int_a^x f(x')\ dx' = \int_a^c f(x')\ dx'$. First, consider the limit from the right. Proving continuity from the right is equivalent to proving $\displaystyle\lim_{\delta\to 0}\int_a^{c+\delta} f(x')\ dx' - \int_a^c f(x')\ dx' = 0$, but the left equals $\displaystyle\lim_{\delta\to 0}\int_c^{c+\delta} f(x')\ dx'$ by properties of the integral. Since f is continuous, its value in the small range $[c, c+\delta]$ is approximately $f\left(c+\frac{\delta}{2}\right)$. Hence, $\displaystyle\lim_{\delta\to 0}\int_c^{c+\delta} f(x')\ dx' \approx \lim_{\delta\to 0}\int_c^{c+\delta} f\left(c+\frac{\delta}{2}\right)\ dx' = \lim_{\delta\to 0} f\left(c+\frac{\delta}{2}\right)\int_c^{c+\delta} dx' = \lim_{\delta\to 0} f\left(c+\frac{\delta}{2}\right)\delta = 0$. A similar argument applies to the limit on the left. To prove differentiability of g at c, note that the difference quotient numerator is given by $g(c+\epsilon) - g(c) = \int_a^{c+\epsilon} f(x')\ dx' - \int_a^c f(x')\ dx' = \int_c^{c+\epsilon} f(x')\ dx'$. For a similar reason to the above, this integral is approximately given by $f\left(c+\frac{\epsilon}{2}\right)\epsilon$. Hence, the difference quotient looks like $\displaystyle g'(c) = \lim_{\epsilon\to 0}\frac{g(c+\epsilon)-g(c)}{\epsilon} \approx \lim_{\epsilon\to 0}\frac{f\left(c+\frac{\epsilon}{2}\right)\epsilon}{\epsilon} = f(c)$. ■

Theorem 2. *Let $f : [a,b] \to \mathbb{R}$ be and integrable function and let $F : [a,b] \to \mathbb{R}$ be any antiderivative of f, then $\int_a^b f(x)\ dx = F(b) - F(a)$.*

Proof. Set $[a,b] \ni x \mapsto g(x) := \int_a^x f(x)\ dx$ and apply the previous theorem. ■

There are so many consequences of this theorem. The most immediate one is the ability to compute Riemann sums that were too difficult to compute explicitly. Here are several examples.

i. $\int_a^b \frac{1}{x}\ dx = \ln(b) - \ln(a) = \ln\left(\frac{b}{a}\right)$

ii. $\int_a^b \sqrt{x}\ dx = \frac{2}{3}x^{3/2}\Big|_b^a = \frac{2}{3}\left(b^{3/2} - a^{3/2}\right)$

iii. $\int_a^b \sin(x)\ dx = -\cos(x)\Big|_b^a = \cos(a) - \cos(b)$

iv. $\int_a^b x^n\ dx = \frac{x^{n+1}}{n+1}\Big|_b^a = \frac{1}{n+1}\left(b^{n+1} - a^{n+1}\right)$

In the above, we have used the notation $F(x)\Big|_b^a := F(b) - F(a)$.

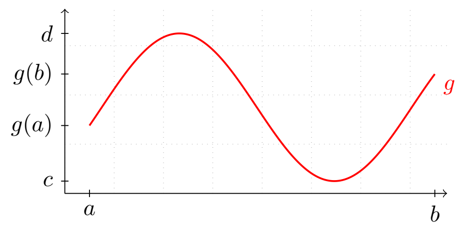
Another consequence is the following. Suppose $a, b : \mathbb{R} \to \mathbb{R}$ are differentiable and $f : \mathbb{R} \to \mathbb{R}$ is integrable and continuous (we are used to writing a and b as constants, but they are now functions). Let F denote an antiderivative of f. Then $\frac{d}{dx}\int_{a(x)}^{b(x)} f(y)\ dy = \frac{d}{dx}\big(F(b(x)) - F(a(x))\big) = F'(b(x))b'(x) - F'(a(x))a'(x) = f(b(x)) - f(a(x))$. This is a sort of generalized fundamental theorem of calculus, but it is really a consequence of it and the chain rule.

The substitution rule is the chain rule

Theorem 3. *Let $g : [a,b] \to [c,d]$ be a differentiable function whose range is $[c,d]$ and let $f : [c,d] \to \mathbb{R}$ be an integrable function. Then* $\displaystyle\int_a^b f\big(g(x)\big)g'(x)\ dx = \int_{g(a)}^{g(b)} f(y)\ dy.$

Proof. If f is continuous and F denotes an antiderivative of f, then $\int_a^b f\big(g(x)\big)g'(x)\ dx = \int_a^b (F\circ g)'(x)\ dx = (F\circ g)(b) - (F\circ g)(a) = \int_{g(a)}^{g(b)} f(y)\ dy$ showing this theorem is a consequence of the Chain Rule and the Fundamental Theorem of Calculus. ■

A natural question might be why aren't the limits of integration on the right c and d, namely why doesn't it say $\int_c^d f(y)\ dy$? Consider the following example illustrating the generic situation.



Notice that the integration actually proceeds in several steps depending on how g increases and decreases over its domain $[a,b]$: $\int_{g(a)}^d + \int_d^c + \int_c^{g(b)}$. The middle integral backtracks. To see how all these integrals simplify, break each of them into parts: $\int_{g(a)}^d + \int_d^c + \int_c^{g(b)} = \int_{g(a)}^{g(b)} + \int_{g(b)}^d + \int_d^{g(b)} + \int_{g(b)}^{g(a)} + \int_{g(a)}^c + \int_c^{g(a)} + \int_{g(a)}^{g(b)} = \int_{g(a)}^{g(b)}$. In other words, all of the other intervals cancel and we are left with $[g(a), g(b)]$.

Let us work through several examples.

i. For $\int_a^b x\sqrt{1+x^2}\ dx$, set $g(x) = 1+x^2$ and $f(y) = \sqrt{y}$ so $g'(x) = 2x$ and $\int_a^b x\sqrt{1+x^2}\ dx = \frac{1}{2}\int_a^b f\big(g(x)\big)g'(x)\ dx = \frac{1}{2}\int_{1+a^2}^{1+b^2}\sqrt{y}\ dy = \frac{1}{3}y^{2/3}\Big|_{1+a^2}^{1+b^2} = \frac{1}{3}\left((1+a^2)^{3/2} - (1+b^2)^{3/2}\right)$.

ii. For $\int_a^b \big(\sin(x)\big)^5\cos(x)\ dx$, set $g(x) = \sin(x)$ and $f(y) = y^5$ so $g'(x) = \cos(x)$ and $\int_a^b \cos(x)\big(\sin(x)\big)^5 dx = \int_a^b f\big(g(x)\big)g'(x)\ dx = \int_{\sin(a)}^{\sin(b)} y^5\ dy = \frac{1}{6}y^6\Big|_{\sin(a)}^{\sin(b)} = \frac{1}{6}\Big(\big(\sin(b)\big)^6 - \big(\sin(a)\big)^6\Big)$.

iii. To compute $\int_a^b \cos^2(x)\ dx$, first note that $\cos(2x) = \cos^2(x) - \sin^2(x) = \cos^2(x) - \big(1-\cos^2(x)\big) = 2\cos^2(x) - 1$ so $\cos^2(x) = \frac{1}{2}\left(\cos(2x)+1\right)$. Hence, $\int_a^b \cos^2(x)\ dx = \int_a^b \frac{1}{2}\left(\cos(2x)+1\right)dx = \frac{1}{2}\int_a^b \cos(2x)\ dx + \frac{1}{2}\int_a^b dx = \frac{1}{2}\int_a^b \cos(2x)\ dx + \frac{1}{2}(b-a)$. To compute the first integral, set $g(x) = 2x$ and $f(y) = \cos(y)$. Then $g'(x) = 2$ and $\frac{1}{2}\int_a^b \cos(2x)\ dx = \frac{1}{4}\int_a^b f\big(g(x)\big)g'(x)\ dx = \frac{1}{4}\int_{2a}^{2b}\cos(y)\ dy = \frac{1}{4}\big(\sin(2b) - \sin(2a)\big)$. Putting these together gives $\int_a^b \cos^2(x)\ dx = \frac{1}{4}\big(\sin(2b) - \sin(2a)\big) + \frac{1}{2}(b-a)$.

iv. For $a, b \in (0,\infty)$ with $a < b$, to evaluate $\int_a^b \frac{(\ln(x))^4}{x}\ dx =$, set $g(x) = \ln(x)$ and $f(y) = y^4$ so $g'(x) = \frac{1}{x}$ and $\int_a^b \frac{(\ln(x))^4}{x}\ dx = \int_a^b f\big(g(x)\big)g'(x)\ dx = \int_{\ln(a)}^{\ln(b)} y^4\ dy = \frac{1}{5}\big((\ln(b))^5 - (\ln(a))^5\big)$.

v. For $\int_a^b \sin(x)e^{\cos(x)}\ dx$, set $g(x) = \cos(x)$ and $f(y) = e^y$ so $g'(x) = -\sin(x)$. Then $\int_a^b \sin(x)e^{\cos(x)}\ dx = -\int_a^b f\big(g(x)\big)g'(x)\ dx = -\int_{\cos(a)}^{\cos(b)} e^y\ dy = e^{\cos(a)} - e^{\cos(b)}$.

The notation $\int f(x)\ dx$ will be used (without limits of integration) to denote an antiderivative of f. In particular, it must include an integration constant. For example, $\int x\ dx = \frac{1}{2}x^2 + C$, where C is a constant. Furthermore, we often us the variable u to denote g. For example in $\int x\sqrt{1+x^2}\ dx$, set $u = 1+x^2$ so that $du = 2x$. Then $\int x\sqrt{1+x^2}\ dx = \frac{1}{2}\int\sqrt{u}\ du = \frac{1}{3}u^{3/2} + C = \frac{1}{3}(1+x^2)^{3/2} + C$. This facilitates computations. This is sometime called <u>*u-substitution*</u>.

The u-substitution rule can be used to simplify other integrals that cannot be immediately computed using the rule.

i. For example, to compute $\int (1+x^2)^{1/2} x^5 \; dx$, set $u = 1 + x^2$ so that $du = 2x dx$ and $x^2 = u-1$. Then $\int (1+x^2)^{1/2} x^5 \; dx = \int (1+x^2)^{1/2} x^4 x \; dx = \frac{1}{2}\int u^{1/2}(u-1)^2 \; du = \frac{1}{2}\int \left(u^{5/2} - 2u^{3/2} + u^{1/2}\right) du = \frac{1}{2}\left(\frac{2}{7}u^{7/2} - \frac{4}{5}u^{5/2} + \frac{2}{3}u^{3/2}\right) + C = \frac{1}{7}(1+x^2)^{7/2} - \frac{2}{5}(1+x^2)^{5/2} + \frac{1}{3}(1+x^2)^{3/2} + C$.

ii. As another example, consider $\int \frac{x^2\,dx}{(3-5x)^2}$. Setting $u = 3-5x$ we have $x = \frac{3-u}{5}$ together with $du = -5dx$ and $dx = -\frac{1}{5}du$. The integral then becomes $\int \frac{x^2\,dx}{(3-5x)^2} = -\frac{1}{5}\int \left(\frac{3-u}{5}\right)^2 \frac{du}{u^2} = -\frac{1}{125}\int \left(\frac{9}{u^2} - \frac{6}{u} + 1\right) \quad du = -\frac{1}{125}\left(-\frac{9}{u} - 6\ln(u) + u\right) + C = -\frac{1}{125}\left(-\frac{9}{3-5x} - 6\ln(3-5x) + 3 - 5x\right) + C$.

Warning: When applying u-substitution, you might apply it differently than your neighbor and your answers may differ. However, your two answers should only differ by an overall integration constant.

Sometimes, clever tricks make computing integrals easier. Also, being able to identify which substitutions will work takes a lot of practice.

i Let $f : [0,\pi] \to \mathbb{R}$ be continuous. First note that $\int_0^\pi x f(\sin(x))\; dx = \frac{\pi}{2}\int_0^\pi f(\sin(x))\; dx$. To see this, set $u = \pi - x$. Then $du = -dx$ and $x = \pi - u$ so that $\int_0^\pi x f(\sin(x))\; dx = \int_\pi^0 (\pi-u) f(\sin(\pi-u))(-du) = \int_\pi^0 (\pi-u) f(\sin(u))(-du) = \pi\int_0^\pi f(\sin(u))\; du - \int_0^\pi u f(\sin(u))\; du$. Notice that the right integral is the same as what we initially started with since u is just a dummy variable. Solving for this gives $\int_0^\pi x f(\sin(x))\; dx = \frac{\pi}{2}\int_0^\pi f(\sin(x))\; dx$.

ii This identity can be used to solve some complicated integrals such as $\int_0^\pi \frac{x\sin(x)}{1+\cos^2(x)}\; dx$. Since $1+\cos^2 = 2+\sin^2$, this integral becomes $\int_0^\pi \frac{x\sin(x)}{2-\sin^2(x)}\; dx$. If we define $f(y) := \frac{y}{2-y^2}$, this integral is of the form $\int_0^\pi x f(\sin(x))\; dx$. By the preceding identity, our integral becomes $\int_0^\pi \frac{x\sin(x)}{1+\cos^2(x)}\; dx = \frac{\pi}{2}\int_0^\pi f(\sin(x))\; dx = \frac{\pi}{2}\int_0^\pi \frac{\sin(x)}{1+\cos^2(x)}\; dx$. To proceed, set $u = \cos(x)$. Then $du = -\sin(x)dx$ and the integral becomes $\int_0^\pi \frac{x\sin(x)}{1+\cos^2(x)}\; dx = -\frac{\pi}{2}\int_1^{-1} \frac{du}{1+u^2} = \frac{\pi}{2}\big(\arctan(1) - \arctan(-1)\big) = \frac{\pi}{2}\left(\frac{\pi}{4} + \frac{\pi}{4}\right) = \frac{\pi^2}{4}$.

Separation of variables

A substance decays at a rate proportional to itself. If the initial amount of the substance is A, find an equation describing the amount of the substance as a function of time. To solve this, let x denote the substance and let t denote time. The initial sentence says $\frac{dx}{dt} = -kx(t)$ with k some positive constant (the reaction constant). To find $x(t)$, a useful trick is to treat dx and dt as separate measures and move all x's to one side and t's to the other. This gives $\frac{dx}{x} = -kdt$. Integrating both sides gives $\int \frac{dx}{x} = -k\int dt$. Performing these antiderivatives gives $\ln(x) = -kt + c$, where c is some integration constant. Solving for x gives $x(t) = e^{-kt+c} = Ce^{-kt}$, where $C = e^c$ is just another way of rewriting the constant of integration. Since $x(t=0) = A$, this lets us solve for C. In fact, $A = C$ so $x(t) = Ae^{-kt}$. Notice that this solution is reasonable since the amount of the substance is decreasing as a function of time. Furthermore, initially, the decay is quick and then it slows down as the amount of the substance decreases.

Recall from Week #07 the differential equation describing the von Bertalanffy growth equation given by $\frac{dL}{dt} = \frac{a}{3} - \frac{b}{3}L = k(L_\infty - L)$, where $k := \frac{b}{3}$ and $L_\infty := \frac{a}{b}$. By separating variables, this becomes $\frac{dL}{L-L_\infty} = -kdt$. Integrating gives $\ln(L - L_\infty) = -kt + c$. Exponentiating gives $L(t) - L_\infty = Ce^{-kt}$. If the initial size of the organism is L_0, this gives the constraint $L_0 - L_\infty = C$. Hence $L(t) = (L_0 - L_\infty)e^{-kt} + L_\infty = L_0 e^{-kt} + L_\infty(1-e^{-kt})$. The solution is consistent with what we would expect. The organism grows rather quickly initially and then continues to grow until it reaches its largest size asymptotically.

Exercises

Exercise 1. Compute the following integrals.

i. $\int_a^b \frac{dx}{1+2x+x^2}$.

ii. For $a,b \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and $a<b$, evaluate $\int_a^b \frac{\exp(\tan(x))}{\cos^2(x)}\; dx$.

iii. For $a,b \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ and $a<b$, evaluate $\int_a^b \frac{2x\exp(\tan(x^2))}{\cos^2(x^2)}\; dx$.

iv. $\int_a^b \sin^2(x)\; dx$.

v. $\int_a^b xe^{3x^2}\; dx$.

vi. $\int_a^b \frac{dx}{5+4x+x^2}$.

Exercise 2. Let $f:\mathbb{R}\to\mathbb{R}$ be continuous and integrable. Define $g(x) := \int_0^x f(y)\; dy$. Describe the critical points of g. Identify when the critical points are local minima and maxima.

Exercise 3. Let $f:\mathbb{R}\to\mathbb{R}$ be a continuous and integrable function and let $c\in\mathbb{R}$. For each $n\in\mathbb{N}$, let $\delta_n:\mathbb{R}\to\mathbb{R}$ be the function defined by $\delta_n(x) := \begin{cases} 2n & \text{if } -\frac{1}{n} \le x \le \frac{1}{n} \\ 0 & \text{otherwise}\end{cases}$. Draw the graph of δ_n for $n = 1,2,3$ and 3. Compute $\int_{-\infty}^\infty \delta_n(x)\; dx$. Prove $\int_{-\infty}^\infty f(x)\delta_n(x)\; dx = 2n\int_{-1/n}^{1/n} f(x)\; dx$. What is $\lim_{n\to\infty}\int_{-\infty}^\infty f(x)\delta_n(x)\; dx$? The sequence of functions $\delta_1, \delta_2, \delta_3, \dots$ is a sequence of functions converging to the Dirac delta "function" (technically, the Dirac delta distribution).

Exercise 4. The work done on a gas during a process that changes the volume but keeps the gas at constant temperature is given by $W = -\int_{V_1}^{V_2} P(V)\; dV$, where $P(V)$ is the pressure of the gas assumed to be a function of the volume. Compute the work under such a process when the volume of the gas changes from V_1 to V_2 under the following assumptions on the gas.

i. For the ideal gas model, the pressure is given by $P(V) = \frac{nRT}{V}$, where n is the number of moles of the gas, R is the ideal gas constant, and T is the temperature (in Kelvin). For reference (not needed to solve this part of this problem), $R = 8.3144598$ Joules per mole per Kelvin ($\mathrm{Jmol^{-1}K^{-1}}$).

ii. For the Van der Walls gas model, the pressure is related to the volume V by $\left(P + \frac{an^2}{V^2}\right)(V - bn) = nRT$, where a and b are positive constants and depend on the gas one is working with (the constants are obtained empirically). In Week #09, you were asked to solve this equation for P. The result is $P(V) = \frac{nRT}{V-bn} - \frac{an^2}{V^2}$.

iii. For the Redlich-Kwong gas model, the pressure is related to the volume V by $P = \frac{nRT}{V-Bn} - \frac{An^2}{V(V+Bn)\sqrt{T}}$, where A and B are positive constants and depend on the gas one is working with (the constants are obtained empirically).

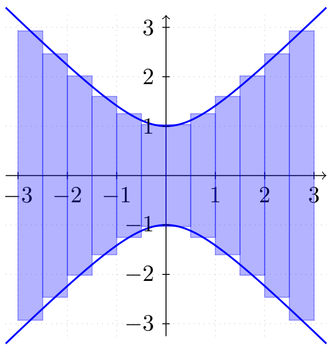
Exercise 5. Solve the differential equation $\frac{dx}{dt} = x^2$ by separation of variables with the initial condition $x(t=0) = -1$.

Exercise 6. Solve the differential equation $\frac{dx}{dt} = -tx$ by separation of variables with the initial condition $x(t=0) = \sqrt{\frac{2}{\pi}}$. This solution is a type of Gaussian distribution.

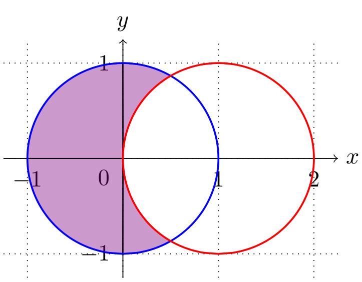
Exercise 7. A Schwarzschild black hole emits Hawking radiation at a rate proportional to the inverse square of its mass $\frac{dm}{dt} = -\frac{\alpha}{m^2}$, where $\alpha = \frac{\hbar c^4}{15360\pi G^2}$ is a constant ($\hbar$ is Planck's reduced constant, c is the speed of light, and G is Newton's gravitational constant) and m is the mass of the black hole. As a result, the mass of the black hole decreases as a function of time. If the initial mass of the black hole is m_0, find the equation describing its mass $m(t)$ as a function of time. When has all of the mass of the black hole evaporated? For a black hole whose initial mass is the solar mass, what is the evaporation time? Compare this to the current age of the universe. What is the significance of this result?

Area between curves

We can think of the definite integral as computing the area between the graph of a function and the function whose value is 0 for all inputs (the x-axis). More generally, we can ask to compute the area between the graphs of two functions. If $f,g:[a,b]\to\mathbb{R}$ are two integrable functions, the (signed) area between their graphs is just the integral of the difference between their graphs. In other words, The area between the graphs of f and g is $\int_a^b\big(f(x)-g(x)\big)\,dx$. Note that the order of which function has the positive sign matters. The other way $\int_a^b\big(g(x)-f(x)\big)\,dx$ would give the area between the graphs of g and f. If one is only interested in the magnitude of the area (ignoring the signs if one graph is above the other), then this area is given by $\int_a^b|f(x)-g(x)|\,dx$.



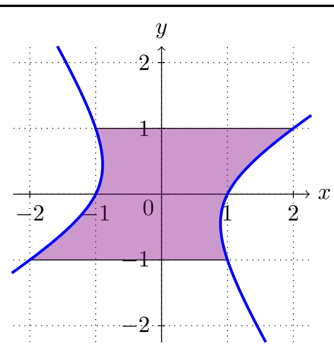
Let $n\in\mathbb{N}$. The area between $[0,1]\ni x\mapsto x^{1/n}$ and $[0,1]\ni x\mapsto x^n$ is given by $\int_0^1\left(x^{1/n}-x^n\right)\,dx=\frac{n}{n+1}x^{(n+1)/n}-\frac{1}{n+1}x^{n+1}\Big|_0^1=\frac{n-1}{n+1}$. Then $\lim\limits_{n\to\infty}\displaystyle\int_0^1\left(x^{1/n}-x^n\right)\,dx=\lim\limits_{n\to\infty}\frac{1-\frac{1}{n}}{1+\frac{1}{n}}=1.$



One can also use integration to find the area between level sets. The level sets shown on the left are those of $x^2+y^2=1$ and $(x-1)^2+y^2=1$. One can do this by finding the area of the top part first and then doubling. The top blue curve is the graph of the function $[0,\frac12]\ni x\mapsto\sqrt{1-x^2}$. The top red curve is given by $[0,\frac12]\ni x\mapsto\sqrt{1-(x-1)^2}$.

The area of the top half splits into two terms $\int_{-1}^0\sqrt{1-x^2}\,dx+\int_0^{1/2}\left(\sqrt{1-x^2}-\sqrt{1-(x-1)^2}\right)\,dx$. Set $x=\cos\theta$ or the first integral and set $x-1=\cos\theta$ for the second part of the second one. After combining like terms, our integrals become $-\int_\pi^{\pi/3}\sin^2(\theta)\,d\theta+\int_\pi^{2\pi/3}\sin^2(\theta)\,d\theta=\int_{\pi/3}^{2\pi/3}\sin^2(\theta)\,d\theta$. To compute this, recall $\cos(2\theta)=\cos^2(\theta)-\sin^2(\theta)=1-2\sin^2(\theta)$ so $\sin^2(\theta)=\frac{1-\cos(2\theta)}{2}$. Therefore, this integral becomes $\int_{\pi/3}^{2\pi/3}\sin^2(\theta)\,d\theta=\frac12\int_{\pi/3}^{2\pi/3}(1-\cos(2\theta))\,d\theta=\frac12\left(\frac{2\pi}{3}-\frac{\pi}{3}\right)-\frac14\left(\sin\left(\frac{4\pi}{3}\right)-\sin\left(\frac{2\pi}{3}\right)\right)=\frac{\pi}{6}-\frac14\left(-\frac{\sqrt3}{2}-\frac{\sqrt3}{2}\right)=\frac{\pi}{6}+\frac{\sqrt3}{4}$. Hence, the area is twice this, which is $\frac{\pi}{3}+\frac{\sqrt3}{2}$.

Let $f:\mathbb{R}^2\to\mathbb{R}$ be defined by the formula $f(x,y):=x^2-xy-y^2$. The level set $f^{-1}(\{1\})$ is drawn on the graph. To compute the shaded area, first solve for the curves as functions in the simplest way possible. This is achieved by solving for x as a function of y. The solutions are $\mathbb{R}\ni y\mapsto x_\pm(y)=\frac{y}{2}\pm\frac{\sqrt{4+5y^2}}{2}$. The difference of these two functions is $x_+(y)-x_-(y)=\sqrt{4+5y^2}$.



The area is therefore $\int_{-1}^1\sqrt{4+5y^2}\,dy$. It will be convenient to note that this integral equals $\int_{-1}^1\sqrt{4+5y^2}\,dy=2\int_0^1\sqrt{4+5y^2}\,dy$.

This integral can be computed by substituting $u=\frac{\sqrt5}{2}y$ giving $\frac{8}{\sqrt5}\int_0^{\sqrt5/2}\sqrt{1+u^2}\,du$. Then, substituting $u=\sinh(v)$ gives $=\frac{8}{\sqrt5}\int_0^{\sinh^{-1}(\sqrt5/2)}\cosh^2(v)\,dv=\frac{4}{\sqrt5}\int_0^{\sinh^{-1}(\sqrt5/2)}(1+\cosh(2v))\,dv=\frac{4}{\sqrt5}\left(\sinh^{-1}(\sqrt5/2)+\frac12\sinh(2\sinh^{-1}(\sqrt5/2))\right)=\frac{4}{\sqrt5}\left(\sinh^{-1}(\sqrt5/2)+\sinh(\sinh^{-1}(\sqrt5/2))\cosh(\sinh^{-1}(\sqrt5/2))\right)=\frac{4}{\sqrt5}\left(\sinh^{-1}(\sqrt5/2)+\frac{\sqrt5}{2}\sqrt{1+\frac54}\right)=\frac{4}{\sqrt5}\sinh^{-1}\left(\frac{\sqrt5}{2}\right)+3$. We used several identities above including $1+\sinh^2=\cosh^2$, $\cosh^2(v)=\frac12(\cosh(2v)+1)$, $\sinh(2w)=2\sinh(w)\cosh(w)$, and $\cosh(\sinh^{-1}(z))=\sqrt{1+z^2}$.

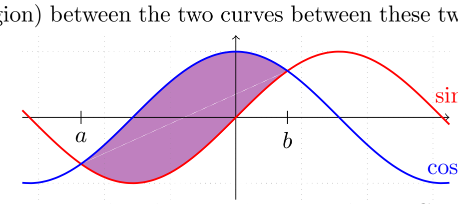
Volumes and Surface Area

Let $f:[a,b]\to\mathbb{R}$ be a continuous function that is strictly positive or strictly negative. The *surface of revolution* of f is the surface obtained from rotating the graph of f about the x-axis (the domain) and viewing $[a,b]\times\mathbb{R}$ as inside of $[a,b]\times\mathbb{R}\times\mathbb{R}$. The circumference of this surface at $x\in[a,b]$ is given by $2\pi f(x)$ while the area of the disc enclosed by this surface of revolution at x is $\pi f(x)^2$. The volume of the enclosed region is therefore $\pi\int_a^b f(x)^2\,dx$. To compute the surface area, we assume in addition that f is differentiable and let us write $f(x)=y(x)$. If the surface of revolution is broken up into infinitesimal cylinders, the radius of these cylinders is $f(x)$ while the length is (by Pythagorean's theorem) $\sqrt{dx^2+dy^2}=\sqrt{1+\left(\frac{dy}{dx}\right)^2}dx$. Therefore, the surface area is $2\pi\int_a^b f(x)\sqrt{1+f'(x)}\,dx$.

For example, fix $r>0$. Then $[-r,r]\ni x\mapsto\sqrt{r^2-x^2}$ has surface of revolution a sphere of radius r. The volume is $\pi\int_{-r}^r\left(r^2-x^2\right)\,dx=\pi\left(2r^3-\frac23r^3\right)=\frac{4\pi r^3}{3}$. The surface area is $2\pi\int_{-r}^r\sqrt{r^2-x^2}\sqrt{1+\frac{x^2}{r^2-x^2}}\,dx=4\pi r^2$.

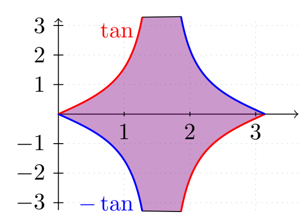
Exercises

Exercise 1. Let a and b be the points of intersection of the graphs of sin and cos depicted in the graph below. Compute the area (the shaded region) between the two curves between these two intersection points.



Exercise 2. Let $m,\lambda,a,b>0$ and suppose $b>a$. Compute the area between the graphs of $\mathbb{R}\ni x\mapsto g(x):=e^{\lambda x}$ and $\mathbb{R}\ni x\mapsto f(x):=mx$ on the interval $[a,b]$. What should the relationship between m and λ be so that a part of the graph of f is above the graph of g? If this relationship is satisfied, suppose that a is the first time f intersects g, i.e. $f(a)=g(a)$, and suppose that b is the second time f intersects g, i.e. $f(b)=g(b)$. Find an expression for the area involving only a,b,m. Can you find an expression for the area involving only m,λ?

Exercise 3. Consider the graphs of $\tan(x)$ and $-\tan(x)$ on their domains. Is the diamond-shaped area (shaded on the right) between these curves finite? If so, compute it. Otherwise, prove the area is infinite.



Exercise 4. Let $a>1$. Gabriel's horn is the surface of revolution obtained from the graph of $[1,a]\ni\mapsto\frac{1}{x}$ by rotating it about the x axis. Compute the surface area of the resulting surface of revolution (do not include the "caps" at the two ends). Compute the volume enclosed by this surface. Then, compute their limits as $a\to\infty$. Which, if any, of the resulting limits are finite? This result shows us that sometimes it is possible to bake a cake but it might not be possible to frost it.