Mathematical Analysis 1
(Engineering Sciences)[1]

Jonathan Ben-Artzi[2]

Fall 2025

[1]Version of January 12, 2026

[2]benartzi@mat.uniroma2.it

These are the lecture notes of course "Mathematical Analysis 1" for Engineering Sciences, taught at Università degli Studi di Roma "Tor Vergata" during the Fall semester of the 2025-2026 academic year. These notes are *notes*, and, as such may contains errors and typos; please let me know of any mistakes you find. The notes are meant to accompany the textbook, *not to replace it*. The textbook is

Claudio Canuto, Anita Tabacco. *Mathematical Analysis 1.* Pearson, 2022.

It is an English version of a standard Analysis 1 course given in Italian universities.

Jonathan Ben-Artzi

Contents

Chapter 1

Basic Notions

1.1 Sets

A **set** is a collection of elements. We denote sets with curly brackets $\{\cdots\}$, with the elements listed within the brackets. As an example, consider the set of students in the class *Mathematical Analysis 1*. This set can either be written explicitly,

$$X = \{\text{Bob}, \text{Lucy}, \text{Andrew}, \text{Giulia}\}$$

or it can be defined using a rule:

$$X = \{\text{all people who are students of } \textit{Mathematical Analysis 1}\}.$$

Some important sets of numbers that we will often encounter are

$$\begin{aligned}
\mathbb{N} &= \text{set of natural numbers} = \{0, 1, 2, \dots\} \\
\mathbb{Z} &= \text{set of integer numbers} = \{0, \pm 1, \pm 2, \dots\} \\
\mathbb{Q} &= \text{set of rational numbers} = \{\text{can you think how to define?}\} \\
\mathbb{R} &= \text{set of real numbers} = \{\text{can you think how to define?}\} \\
\mathbb{C} &= \text{set of complex numbers} = \{\text{can you think how to define?}\}
\end{aligned}$$

Another important set is the **empty set** which contains no elements. It is denoted $\emptyset$.

Basic notation

- **Element of**: if x is an element of X we write $x \in X$
- **Not element of**: if x is *not* an element of X we write $x \notin X$
- **Subset**: if A is a subset of X (i.e. any element of A is also an element of X) we write $A \subseteq X$ or $X \supseteq A$

  In this case it is possible that $A = X$.
- **Proper subset**: if A is a *proper* subset of X we write $A \subset X$ or $X \supset A$

  In this case $A \neq X$ (i.e. there exists $x \in X$ and $x \notin A$).

Lemma 1.1: For some set X and subsets $A, B \subseteq X$, if $A \subseteq B$ and $B \subseteq A$ then $A = B$.

Proof. By contradiction, assume that $A \neq B$. Then, without loss of generality, there exists $x \in X$ such that $x \in A$ and $x \notin B$. But then it is not true that $A \subseteq B$. The contradiction assumption must therefore be false, i.e. $A = B$. $\square$

Characteristic property

The elements of a subset $A \subseteq X$ can often be characterized by a mathematical property that they satisfy. This property is denoted $p(x)$, and we write

$$A = \{x \in X \mid p(x)\}.$$

For example, if $p(x) =$ 'x is even', then

$$A = \{x \in \mathbb{N} \mid x \text{ is even}\} = \{0, 2, 4, \dots\} \subseteq \mathbb{N}.$$

Operations on sets

- **Complement**: if $A \subseteq X$ then we define its *complement* to be

$$A^C = \mathcal{C}A = \{x \in X \mid x \notin A\}.$$

- **Union**: for two sets $A \subseteq X$ and $B \subseteq X$ we define their *union* to be

$$A \cup B = \{x \in X \mid x \in A \text{ or } x \in B\}$$

- **Intersection**: for two sets $A \subseteq X$ and $B \subseteq X$ we define their *intersection* to be

$$A \cap B = \{x \in X \mid x \in A \text{ and } x \in B\}$$

- **Difference**: for two sets $A \subseteq X$ and $B \subseteq X$ we define their *difference* to be

$$A \setminus B = \{x \in X \mid x \in A \text{ and } x \notin B\}$$

- **Symmetric Difference**: for two sets $A \subseteq X$ and $B \subseteq X$ we define their *symmetric difference* to be

$$A \Delta B = (A \cup B) \setminus (A \cap B)$$

- **Disjoint Union**: for two sets $A \subseteq X$ and $B \subseteq X$ whose intersection is empty, we often replace the symbol $\cup$ by

$$A \sqcup B \quad \text{or} \quad A \dot{\cup} B$$

Lemma 1.2 (Properties of $\cap$ and $\cup$)**:** For some set X and subsets $A, B, C \subseteq X$ the operations $\cap$ and $\cup$ satisfy:

1. Boolean properties: $A \cap A^C = \emptyset$ and $A \cup A^C = X$.

2. Commutativity: $A \cap B = B \cap A$ and $A \cup B = B \cup A$.

3. Associativity: $(A \cap B) \cap C = A \cap (B \cap C)$ and $(A \cup B) \cup C = A \cup (B \cup C)$.

4. Distributivity: $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ and $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.

5. De Morgan laws: $(A \cap B)^C = A^C \cup B^C$ and $(A \cup B)^C = A^C \cap B^C$.

Proof. We prove the first of the De Morgan laws. The rest is an exercise.

We want to show that two sets are the same: $(A \cap B)^C = A^C \cup B^C$. To do this, we will show that the set on the left is contained in (or equal to) the set on the right, and vice versa. I.e., we shall show that $(A \cap B)^C \subseteq A^C \cup B^C$ and $(A \cap B)^C \supseteq A^C \cup B^C$.

(i) To show that $(A \cap B)^C \subseteq A^C \cup B^C$, we note the following implications:

$$\begin{aligned} & x \in (A \cap B)^C \\ & \quad \Downarrow \\ & x \notin A \cap B = \{y \in X \mid y \in A \text{ and } y \in B\} \\ & \quad \Downarrow \\ & x \notin A \text{ or } x \notin B \\ & \quad \Downarrow \\ & x \in A^C \text{ or } x \in B^C. \end{aligned}$$

Since $A^C \subseteq A^C \cup B^C$ and $B^C \subseteq A^C \cup B^C$, we conclude that necessarily $x \in A^C \cup B^C$. Hence $(A \cap B)^C \subseteq A^C \cup B^C$.

(ii) Conversely, we can show $(A \cap B)^C \supseteq A^C \cup B^C$. Assume that $x \in A^C \cup B^C$ and *by contradiction*, assume that $x \notin (A \cap B)^C$. Then we have the implications:

$$\begin{aligned} & x \notin (A \cap B)^C \\ & \quad \Downarrow \\ & x \in A \cap B = \{y \in X \mid y \in A \text{ and } y \in B\} \\ & \quad \Downarrow \\ & x \in A \text{ and } x \in B \\ & \quad \Downarrow \\ & x \notin A^C \text{ and } x \notin B^C. \end{aligned}$$

But this is in contradiction to the assumption that $x \in A^C \cup B^C$. Therefore the contradiction assumption $x \notin (A \cap B)^C$ is not true, hence $x \in (A \cap B)^C$.

We have shown that $(A \cap B)^C \subseteq A^C \cup B^C$ and that $(A \cap B)^C \supseteq A^C \cup B^C$, so by Lemma 1.1 the two sets must be equal, completing the proof. $\square$

Power set

For a given set X, we define its power set $\mathcal{P}(X)$ to be the set of all subsets of X:

$$\mathcal{P}(X) = \{A \mid A \subseteq X\}.$$

In particular, $\emptyset \in \mathcal{P}(X)$ and $X \in \mathcal{P}(X)$.

1.2 Elements of mathematical logic

The building blocks of mathematical logic are **formulas**, which can be either *true* or *false*. Here are some examples:

$$
\begin{aligned}
p &= \text{`Blue is a color'} \\
q &= \text{`15 is the square of a natural number'} \\
r &= \text{`the number 3 belongs to the set } X\text{'}
\end{aligned}
$$

Then p is true, q is false, and we have no way of knowing whether r is true or false without knowing something about the set X.

1.2.1 Connectives

Connectives are the tools to build new formulas from existing ones. We briefly mention them:

Logical negation $\neg p$ ('not p') is the negation of the formula p

Logical conjunction $p \wedge q$ ('p and q')

Logical disjunction $p \vee q$ ('p or q')

Logical implication $p \Rightarrow q$ ('p implies q' or 'if p, then q')

Logical equivalence $p \Leftrightarrow q$ ('p is logically equivalent to q')

Proof by contradiction

This formalism allows us to understand the notion of a proof by contradiction, which is summed up by the logical equivalence:

$$(p \Rightarrow q) \qquad \Leftrightarrow \qquad (p \wedge \neg q \Rightarrow \neg p)$$

1.2.2 Predicates

A **predicate** is a formula that depends on one or more variables. In fact, we have seen predicates before, when we called them 'characteristic properties'. Here are some more examples:

$$
\begin{aligned}
p(x) &= \text{`}x\text{ is a prime number'} \\
q(y) &= \text{`}y\text{ is the square of a natural number'} \\
r(x, y) &= \text{`}x\text{ is divisible by }y\text{'}
\end{aligned}
$$

1.2.3 Quantifiers

In a set X, for a given predicate $p(x)$ with $x \in X$, we can ask whether p is always true, or perhaps only sometimes. This is expressed mathematically as follows:

Universal quantifier: $\forall x,\ p(x)$ (we say 'for every x, $p(x)$ holds')

Existential quantifier: $\exists x,\ p(x)$ (we say 'there exists x, such that $p(x)$ holds')

Unique existential quantifier: $\exists! x,\ p(x)$ (we say 'there exists one and *only one* x, such that $p(x)$ holds')

Example 1.1: Suppose that, as above, $p(x) =$ 'x is a prime number'. If $X = \mathbb{N}$, then it is true that *there exists* $x \in X$ that is a prime number, i.e. $\exists x,\ p(x)$. However, it is *not true* that *every* $x \in X$ is a prime number. That is, $\neg(\forall x,\ p(x))$.

Example 1.2: Consider the predicate $p(x) =$ '$x^2 = x$'. If $X = \{1, 2, 3, \dots\}$, then $1 \in X$ is the unique element in X for which $p(x)$ is true. That is, $\exists! x,\ p(x)$. On the set $Y = \{2, 3, 4, \dots\}$ the predicate $p(x)$ is never true, i.e. $\neg(\exists x,\ p(x))$.

The notions of predicates and quantifiers allow us to formalize the idea of induction:

Theorem 1.3 (Principle of Induction)**:** Let $N \in \mathbb{N}$ and denote by $p(n)$ a predicate defined for every $n \geq N$, $n \in \mathbb{N}$. Suppose that the following hold:

1. $p(N)$ is true,
2. $\forall n \geq N$, $p(n) \Rightarrow p(n+1)$.

Then $p(n)$ is true for all integers $n \geq N$.

Proof. By contradiction, assume that $\exists n \geq N$ for which $p(n)$ is false. Then the set

$$F = \{n \in \mathbb{N} \mid n \geq N \text{ and } p(n) \text{ is false}\}$$

is not empty. Define $m \in F$ to be the smallest number in F. Then $p(m)$ is false. Therefore $m \neq N$ (recall that we know that $p(N)$ is true). So necessarily $m > N$, and it follows that $m - 1 \geq N$. By our definition of the number m, $p(m-1)$ must be true (otherwise $m-1$ would have been the smallest number in F). But we know that $\forall n \geq N$, $p(n) \Rightarrow p(n+1)$. Taking $n = m - 1$ we get that $p(m-1) \Rightarrow p(m)$. But this is not true, since $p(m-1)$ is true while $p(m)$ is false. We have therefore reached a contradiction, so that $\neg(\exists n \geq N$ for which $p(n)$ is false), i.e. $\forall n \geq N$, $p(n)$ is true. □

Example 1.3 (Bernoulli inequality)**:** We claim that $\forall r \geq -1$, the *Bernoulli inequality*

$$(1+r)^n \geq 1 + nr, \qquad \forall n \in \mathbb{N},$$

holds. We prove this by induction. Here

$$p(n) = \text{`}(1+r)^n \geq 1 + nr\text{'}.$$

1. For $n = 0$, we have $(1+r)^0 = 1$ and $1 + 0 \cdot r = 1$ so that $(1+r)^0 \geq 1 + 0 \cdot r$ and therefore $p(0)$ is true.

2. Now assume that $p(n)$ is true. This is called the **induction assumption**. Let us show that $p(n+1)$ is true. Using the fact that $1 + r \geq 0$, we have

$$\begin{aligned}
(1+r)^{n+1} &= (1+r)(1+r)^n \\
&\geq (1+r)(1+nr) \qquad \textit{(here we use the induction assumption and that } 1+r \geq 0\textit{)} \\
&= 1 + (n+1)r + nr^2 \\
&\geq 1 + (n+1)r. \qquad \textit{(since } nr^2 \geq 0\textit{)}
\end{aligned}$$

Hence $p(n+1)$ is true, and by the Principle of Induction (we usually just say '*by induction*') the Bernoulli inequality holds for all $n \in \mathbb{N}$.

Proof by induction

To prove that a statement $p(n)$ is true for all $n \geq N$ by induction, we need to demonstrate two things:

1. that $p(N)$ (the **base case**) is true,
2. that $p(n)$ being true (the **induction assumption**) implies $p(n+1)$ being true for all $n \geq N$.

1.3 Sets of numbers

We have already seen the definitions of the set $\mathbb{N}$ of natural numbers and the set $\mathbb{Z}$ of integers. We further define

$$\mathbb{N}_+ = \mathbb{N} \setminus \{0\}$$

to be the set of positive integers.

The set $\mathbb{Q}$ of rational numbers

We now have the tools to define the set $\mathbb{Q}$ as:

$$\mathbb{Q} = \left\{ r = \frac{p}{q} \;\middle|\; p \in \mathbb{Z},\ q \in \mathbb{N}_+ \right\}$$

Any rational number r has infinitely many representations. For example

$$\frac{1}{2} = \frac{2}{4} = \frac{3}{6} = \cdots$$

Therefore, we normally choose the (unique) representative such that p and q have no common divisors.

We can also write fractions in base 10 (which is the standard base for us), as

$$r = \pm(c_k 10^k + c_{k-1} 10^{k-1} + \cdots c_1 10 + c_0 + c_{-1} 10^{-1} + c_{-2} 10^{-2} + \cdots)$$

where all coefficients $c_k, c_{k-1}, \ldots, c_0, c_{-1}, \ldots$ can assume any of the values $0, 1, 2, \ldots, 9$. Here are some examples:

$$\begin{aligned}
\frac{1}{2} &= +(0 \cdot 10^k + \cdots + 0 \cdot 10 + 0 + 5 \cdot 10^{-1} + 0 \cdot 10^{-2} + \cdots) = 0.5, \\
-\frac{29}{2} = -14\frac{1}{2} &= -(0 \cdot 10^k + \cdots + 1 \cdot 10 + 4 + 5 \cdot 10^{-1} + 0 \cdot 10^{-2} + \cdots) = -14.5, \\
\frac{1}{3} &= +(0 \cdot 10^k + \cdots + 0 \cdot 10 + 0 + 3 \cdot 10^{-1} + 3 \cdot 10^{-2} + \cdots) = 0.\overline{33},
\end{aligned}$$

where the overline means that the expression repeats itself infinitely:

$$\overline{a_1 a_2 \cdots a_m} = (a_1 a_2 \cdots a_m)(a_1 a_2 \cdots a_m)(a_1 a_2 \cdots a_m) \cdots$$

Decimal representation of rational numbers

Using long division is can be verified that the decimal representation of any rational number is either finite (as in the example for $\frac{1}{2}$) or infinitely repeating (as in the example for $\frac{1}{3}$). The converse is also true: every number that has a finite or repeating representation is rational.

The set $\mathbb{R}$ of real numbers

The world around us is made up of three space dimensions and one time dimension which are all continuous. We move through space and time in a continuous fashion. So, a natural question is whether the rational numbers are enough to describe our world? For instance, as time elapses from 1pm to 2pm, can all intermediate moments be described by rational numbers?

The answer turns out to be *no*. Between 1 and 2, for instance, there exists a number which is the solution of the equation $x^2 = 2$ (we call this number *the square root of* 2 and denote it by $\sqrt{2}$) which is not rational; we normally say that it is **irrational**.

Lemma 1.4 ($\sqrt{2}$ is irrational)**:** The number x satisfying the equation $x^2 = 2$ is irrational.

Proof. We prove this by contradiction. Suppose that there is a rational solution $x = \frac{p}{q}$ (with p and q having no common divisors). Then $\frac{p^2}{q^2} = 2$ so that

$$p^2 = 2q^2.$$

Hence p^2 is even (it is divisible by 2). If p^2 is even then so is p itself. Hence $\exists k \in \mathbb{N}$ such that $p = 2k$ (observe that, in fact, k must be positive, otherwise $x = 0$). We therefore have $(2k)^2 = 2q^2$ which implies that

$$2k^2 = q^2.$$

It follows that q^2 is even, and consequently so is q itself. But this means that p and q have the number 2 as a common divisor, a contradiction. Therefore there is no rational representation for x. $\square$

The real number line. We can write all rational numbers along a single line, placing bigger numbers to the right. Since $1^2 = 1$ and $2^2 = 4$, it seems obvious that $\sqrt{2}$ must lie between 1 and 2. We have thus found an irrational number between 1 and 2,

$$1 < \sqrt{2} < 2.$$

Since it is irrational, it has an infinite representation that *never* repeats. Therefore it is *impossible* to exactly write its value. The beginning of its expansion is:

$$\sqrt{2} = 1.41421356\cdots$$

It follows that

$$\begin{aligned} 1.4 <& \sqrt{2} < 1.5 \\ 1.41 <& \sqrt{2} < 1.42 \\ 1.414 <& \sqrt{2} < 1.415 \\ 1.4142 <& \sqrt{2} < 1.4143 \\ &\vdots \end{aligned}$$

These are sequences of rational numbers to the left and to the right of $\sqrt{2}$, both of which tend to $\sqrt{2}$, but never reach it. So $\sqrt{2}$ fills a certain gap. Indeed, with the irrational numbers added, we get a continuous line of increasing numbers, called the *real number line*. This property of having a *continuous* line is called *completeness* and we say that the real number line is **complete** (intuitively, it means that there are no gaps).

Actually, there are *many* irrationals. It turns out that there are infinitely many rationals and infinitely many irrationals. They are all intertwined:

- between any two rationals $r_1 < r_2$ there are infinitely many irrationals,
- between any two irrationals $y_1 < y_2$ there are infinitely many rationals.

However, there are *more* irrationals than there are rationals. If we write the set of real numbers (the real number line) as the (disjoint) union of the rationals and the irrationals

$$\mathbb{R} = \mathbb{Q} \sqcup (\mathbb{R} \setminus \mathbb{Q})$$

then both $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are infinite, but $\mathbb{R} \setminus \mathbb{Q}$ is 'bigger' in a certain sense. This statement is made rigorous using a mathematical theory called *measure theory*. It can be understood intuitively as follows: if a number is chosen randomly between, say, 0 and 1, then it is *almost surely* an irrational number (in other words, the probability of choosing a rational number is zero). So within the real number line there are 'more' irrationals. However there are also rationals *everywhere*. We say that **the rational numbers $\mathbb{Q}$ are dense within the reals $\mathbb{R}$.**

1.3.1 The ordering of real numbers

The real numbers are a **totally ordered set**: $\forall x, y \in \mathbb{R}$, one (and *only* one) of the following properties holds:

$$x = y \quad \text{or} \quad x < y \quad \text{or} \quad x > y.$$

We often use the following symbols for these important subsets of $\mathbb{R}$:

$$\begin{aligned} \mathbb{R}_+ &= \{x \in \mathbb{R} \ \mid \ x > 0\} \\ \mathbb{R}_- &= \{x \in \mathbb{R} \ \mid \ x < 0\} \\ \mathbb{R}_* &= \{x \in \mathbb{R} \ \mid \ x \geq 0\} = \mathbb{R}_+ \cup \{0\} \end{aligned}$$

Infinity. It is convenient to introduce a symbol to help us express the fact that there is no greatest number. We thus introduce the symbol for **plus infinity**

$$+\infty,$$

which is thought of as an object that is greater than any real number. Similarly, **minus infinity**

$$-\infty$$

symbolizes an object that is smaller than any real number. Note that these are *not* numbers.

Intervals. The notion of an *interval* – a part of the real number line – will be very important:

Intervals

Let $a, b \in \mathbb{R}$ with $a \leq b$. Then the **closed interval** between a and b is defined as

$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}.$$

Let $a, b \in \mathbb{R}$ with $a < b$. Then the **open interval** between a and b is defined as

$$(a, b) = \{x \in \mathbb{R} \mid a < x < b\}.$$

That is, closed intervals include all points between a and b, including a and b themselves. Open intervals do not include the endpoints. The points between a and b (excluding a and b themselves) are called **interior points**.

We can also define intervals that are closed on one end and open on the other. Let $a < b$, then

$$[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\},$$
$$(a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}.$$

Other important sets are **half-lines**, which are sets that have a lower/upper limit only on one side. Here the symbols for plus or minus infinity come in handy:

$$[a, +\infty) = \{x \in \mathbb{R} \mid a \leq x\},$$
$$(a, +\infty) = \{x \in \mathbb{R} \mid a < x\},$$
$$(-\infty, b] = \{x \in \mathbb{R} \mid x \leq b\},$$
$$(-\infty, b) = \{x \in \mathbb{R} \mid x < b\}.$$

The entire real line is often represented as the set of all points that are greater than $-\infty$ and less than $+\infty$:

$$\mathbb{R} = (-\infty, +\infty).$$

Absolute value. An important operation that we will frequently encounter is the **absolute value** of a real number x:

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

That is, $|x|$ measures the distance of x from 0, regardless of the sign of x. Consequently, we can define the **distance** between $x, y \in \mathbb{R}$ as:

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

1.3.2 Bounded sets

The notion of a bounded set is an extension of the notion of an interval:

Bounded sets

- A subset $A \subseteq \mathbb{R}$ is said to be **bounded from above** if $\exists b \in \mathbb{R}$ such that

  $$x \leq b, \qquad \text{for all } x \in A.$$

  Such a real number b is called an **upper bound** for A.

- $A \subseteq \mathbb{R}$ is said to be **bounded from below** if $\exists a \in \mathbb{R}$ such that

  $$x \geq a, \qquad \text{for all } x \in A.$$

  Such a real number a is called a **lower bound** for A.

- A is called **bounded** if it is bounded from below and from above.

- If A is not bounded from above and not bounded from below, we say that A is **unbounded**.

Lower and upper bounds are not unique

It is very important to observe that both lower and upper bounds are not unique: if a is a lower bound for a subset $A \subseteq \mathbb{R}$, then any $y \leq a$ is also a lower bound. Similarly, if b is an upper bound for A, then any $z \geq b$ is also an upper bound.

Example 1.4:

1. The subset $A_1 = \{-2, 0.5, 7\}$ is bounded. Any number $a \leq -2$ is a lower bound and any number $b \geq 7$ is an upper bound.

2. The subset $A_2 = \mathbb{N}$ is bounded from below but not from above. Any number $a \leq 0$ is a lower bound.

3. The subset $A_3 = \mathbb{Z}$ is unbounded.

4. The subset $A_4 = \{\frac{1}{n} \mid n \in \mathbb{N}_+\} = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$ is bounded. Why? We see that a possible upper bound is 1, which belongs to A_4. What about a lower bound? The number 0 is a lower bound, however 0 does not belong to A_4. Below we will see that there is no lower bound that belongs to A_4.

5. The subset $A_5 = \{x \mid |x| > 100\}$ is unbounded.

Let us take a closer look at the lower bound of $A_4 = \{\frac{1}{n} \mid n \in \mathbb{N}_+\}$:

Lemma 1.5: Any lower bound of the set $A_4 = \{\frac{1}{n} \mid n \in \mathbb{N}_+\}$ does not belong to A_4.

Proof. By contradiction, suppose that $\exists a \in A_4$ such that $a \le x$ for all $x \in A_4$. Since $a \in A_4$, there exists $N \in \mathbb{N}$ such that $a = \frac{1}{N}$. Observe that $\frac{1}{N+1}$ is also an element of A_4 and $\frac{1}{N+1} < \frac{1}{N} = a$. But this contradicts our assumption that a is a lower bound for A_4. Hence no element in A_4 can be a lower bound of A_4. $\square$

In fact, we can prove a stronger statement:

Lemma 1.6: The set $A_4 = \{\frac{1}{n} \mid n \in \mathbb{N}_+\}$ does not have a positive lower bound.

Proof. By contradiction, assume that there exists $r > 0$ that is a lower bound for A_4. Define $N = \lceil \frac{1}{r} \rceil$ to be the first integer greater than or equal to $\frac{1}{r}$. Then $N + 1 > \frac{1}{r}$ and consequently $\frac{1}{N+1} < r$. But $\frac{1}{N+1} \in A_4$, in contradiction to the assumption that r is a lower bound. Hence there is no positive lower bound, and 0 is the *greatest lower bound.* $\square$

Supremum and infimum

Let $A \subset \mathbb{R}$ be a subset.

- The **supremum** (if exists) of A (also called the **least upper bound, l.u.b.**) is the smallest of all upper bounds of A. It is denoted

$$s = \sup A$$

  and it fulfils the following two conditions:

  1. $\forall x \in A,\ x \le s$,
  2. $\forall r < s\ \exists x \in A$ s.t. $x > r$.

  If there is no such number, we define $\sup A = +\infty$.

- The **infimum** (if exists) of A (also called the **greatest lower bound, g.l.b.**) is the largest of all lower bounds of A. It is denoted

$$\ell = \inf A$$

  and it fulfils the following two conditions:

  1. $\forall x \in A,\ x \ge \ell$,
  2. $\forall r > \ell\ \exists x \in A$ s.t. $x < r$.

  If there is no such number, we define $\inf A = -\infty$.

Note that in the above definition, we defined *the* supremum and *the* infimum, implying that these two numbers are unique. *A priori*, that is not obvious (even though it is true). It requires a proof.

Proposition 1.7: For any subset $A \subseteq \mathbb{R}$, there are unique elements $\ell, s \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$ such that $\ell = \inf A$ and $s = \sup A$.

Proof. Exercise. *Hint: prove it by contradiction.* □

The supremum and the infimum might not belong to the set!

It is very important to remember that for a subset $A \subseteq \mathbb{R}$, its supremum and its infimum might *not* belong to it. We have seen it with A_4 above: its infimum is 0, yet $0 \notin A_4$. For $A_2 = \mathbb{N}$, the supremum is $+\infty$, which isn't a number, and in particular isn't an element of A_2.

Keeping in mind the preceding comment, in the case that the supremum and/or infimum *do* belong to the set we give them another name:

Maximum and minimum

Let $A \subset \mathbb{R}$ be a subset. If $\sup A \in A$ then we say that the supremum is *attained*, and it is called the **maximum** of A and denoted

$$\max A.$$

Similarly, if $\inf A \in A$ then we say that the infimum is *attained*, and it is called the **minimum** of A and denoted

$$\min A.$$

1.3.3 The cardinality of subsets of $\mathbb{R}$

The *cardinalty* of a set A is a measure of its size. The cardinality of a set containing finitely many elements is simply the number of elements: the cardinality of $A = \{-11, 600, \sqrt{17}\}$ is 3. The cardinality of the set $B = \{\text{Giulia, Sam, Amelia}\}$ is also 3. The concept of cardinality becomes more delicate when dealing with sets containing infinitely many elements.

Countable sets

A subset $A \subseteq \mathbb{R}$ is said to be **countable** if it is possible to enumerate all its elements.

Example 1.5: 1. Any set with finitely many elements is countable.

2. The set $\mathbb{N}_+ = \{1, 2, 3, \dots\}$ is trivially countable: we assign to the element $1 \in \mathbb{N}_+$ the number 1, to the element $2 \in \mathbb{N}_+$ the number 2, and so on.

3. The sets $\mathbb{N}, \mathbb{Z}$ are also countable. The set of even natural numbers is also countable. *Can you prove it?*

Proposition 1.8: The set $\mathbb{Q}$ of rational numbers is countable.

Proof. To prove that a set is countable we need to demonstrate that we can enumerate its elements. All non-zero rational numbers are of the form $\pm\frac{p}{q}$ where $p, q \in \mathbb{N}_+$. These can be organized in a table as follows:

$$\begin{tikzcd}
\pm\frac{1}{1} \arrow[r] & \pm\frac{2}{1} \arrow[dl] & \pm\frac{3}{1} \arrow[r] & \pm\frac{4}{1} \arrow[dl] & \cdots \\
\pm\frac{1}{2} \arrow[d] & \pm\frac{2}{2} \arrow[ur] & \pm\frac{3}{2} \arrow[dl] & \pm\frac{4}{2} & \cdots \\
\pm\frac{1}{3} \arrow[ur] & \pm\frac{2}{3} \arrow[dl] & \pm\frac{3}{3} & \pm\frac{4}{3} & \cdots \\
\pm\frac{1}{4} & \pm\frac{2}{4} & \pm\frac{3}{4} & \pm\frac{4}{4} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{tikzcd}$$

The first row contains all rationals with $q = 1$, the second row contains all rationals with $q = 2$, and so on. At the same time, the first column contains all rationals with $p = 1$, the second column contains all rationals with $p = 2$, and so on. Thus every non-zero rational appears within this table. In fact, every non-zero rational appears infinitely many times in this table, because for every $\pm\frac{p}{q}$, the table also contains $\pm\frac{2p}{2q}$, $\pm\frac{3p}{3q}$, and so on.

The arrows in the table demonstrate a strategy for enumerating the rationals. Since 0 doesn't appear in the table, we start by counting it: we assign to it the number 1. Then we enumerate the elements $\pm\frac{1}{1}$ to which the numbers 2 and 3 are assigned. Next we move on to $\pm\frac{2}{1}$ to which the numbers 4 and 5 are assigned. Then $\pm\frac{1}{2}$, to which 6 and 7 are assigned, $\pm\frac{1}{3}$ to which 8 and 9 are assigned, $\pm\frac{2}{2}$ to which 10 and 11 are assigned, and so on. Eventually to every rational will be assigned a natural number, completing the proof. □

Theorem 1.9: The set $\mathbb{R}$ of real numbers if *not* countable.

Proof. The following proof, *by contradiction*, due to Georg Cantor, goes back to the late 19th century. It's enough to just look at the real numbers between 0 and 1. Each $x \in (0, 1)$ has a decimal representation:

$$x = 0.\ a_1\ a_2\ a_3\ a_4\ a_5\ \ldots\ldots \qquad \text{where } a_i \in \{0, 1, \ldots, 9\} \text{ for every } i \in \mathbb{N}_+.$$

Suppose, by contradiction, that the real numbers in $(0, 1)$ were countable. Then we can enumerate all $x \in (0, 1)$. Let's enumerate them as $r_1, r_2, \ldots, r_n, \ldots$. Each of these has a decimal representation: $r_n = 0\ .\ c_{n,1}\ c_{n,2}\ c_{n,3}\ c_{n,4}\ \ldots\ldots$ Let's write the following table, with r_1 on the first line, r_2 on the second, and so on:

$$
\begin{array}{rlllllll}
r_1 = & 0\,. & \boxed{c_{1,1}} & c_{1,2} & c_{1,3} & c_{1,4} & c_{1,5} & \cdots \\
r_2 = & 0\,. & c_{2,1} & \boxed{c_{2,2}} & c_{2,3} & c_{2,4} & c_{2,5} & \cdots \\
r_3 = & 0\,. & c_{3,1} & c_{3,2} & \boxed{c_{3,3}} & c_{3,4} & c_{3,5} & \cdots \\
r_4 = & 0\,. & c_{4,1} & c_{4,2} & c_{4,3} & \boxed{c_{4,4}} & c_{4,5} & \cdots \\
r_5 = & 0\,. & c_{5,1} & c_{5,2} & c_{5,3} & c_{5,4} & \boxed{c_{5,5}} & \cdots \\
 & & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
$$

Remember that each digit $c_{i,j}$ is an integer between 0 and 9. Let's now look at the numbers $c_{i,i}$ on the *diagonal*. Consider a real number between 0 and 1 defined as:

$$r = 0.\ d_1\ d_2\ d_3\ d_4\ d_5\ \ldots\ldots \qquad \text{where } d_i \neq c_{i,i} \quad \forall i \in \mathbb{N}_+.$$

Then

$$
\begin{array}{rl}
r \neq r_1 & \text{since the } \textit{first} \text{ digit in their expansions is different} \\
r \neq r_2 & \text{since the } \textit{second} \text{ digit in their expansions is different} \\
r \neq r_3 & \text{since the } \textit{third} \text{ digit in their expansions is different} \\
\vdots & \\
r \neq r_n & \text{since the } n\text{th digit in their expansions is different} \\
\vdots &
\end{array}
$$

Hence r isn't equal to any of the reals that have been enumerated. However, that is a contradiction to the assumption that all reals in $(0, 1)$ have been enumerated. □

This method of proof is called a **diagonal argument**, and has become an important technique for proving various results in the years since Cantor introduced it.

1.4 Cartesian product

Ordered pairs and Cartesian product

Let X and Y be two nonempty sets. Then we define an **ordered pair** to be

$$(x, y)$$

where $x \in X$ and $y \in Y$. The set of all ordered pairs from X and Y is called the **Cartesian product of** X **and** Y and is denoted $X \times Y$:

$$X \times Y = \{(x, y) \mid x \in X,\ y \in Y\}.$$

When $X = Y$ we often write X^2 rather than $X \times X$.

The order of elements in an ordered pair is important: the first *component* belongs to X and the second *component* belongs to Y. Thus, the ordered pair (x, y) is *fundamentally different* from the set $\{x, y\}$.

Example 1.6:

1. One of the most important examples of a Cartesian product, which we will often use, is that of the **plane** $\mathbb{R}^2$. This is the set of all ordered pairs of real numbers:

   $$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) \mid x, y \in \mathbb{R}\}.$$

   This is called the **Cartesian plane**.

2. A practical example is that of latitude and longitude: we can describe points on Earth in terms of their angle north/south of the equator, and east/west relative to a chosen Prime Meridian (a big circle connecting the North and South poles). Normally this is chosen to be the Greenwich Meridian, passing through the Royal Observatory in Greenwich, England (this choice dates back to 1851). In this system, the Colosseum, for instance, is located at $(41.8902°\mathrm{N}, 12.4922°\mathrm{E})$.

An ordered *pair* need not be limited to two components. One can have n elements from n sets, where $n \in \mathbb{N}$:

$$(x_1, x_2, \ldots, x_n), \qquad x_i \in X_i, \; i = 1, \ldots, n.$$

This is called an **n-tuple**. Then the Cartesian product involves the sets $X_1, \ldots, X_n$:

$$X_1 \times X_2 \times \cdots \times X_n.$$

If $X_i = X$ for all $i = 1, \ldots, n$, then we simply write $X \times X \times \cdots \times X = X^n$.

1.5 Relations in the Cartesian plane

In the Cartesian plane (we will simply call it *the plane*) $\mathbb{R}^2$ we typically denote the first coordinate by x and the second by y. We can describe subsets of $\mathbb{R}^2$ by equations and inequalities involving x and y. This is best understood using some examples:

Example 1.7: Sets described by equations:

1. The equation $y = 0$ describes the x-axis.
2. The equation $x = 0$ describes the y-axis.
3. The equation $x = y$ describes the line through the origin with slope 1.
4. The equation $x^2 + y^2 = 1$ describes the circle of radius 1 around the origin.

Example 1.8: Sets described by inequalities:

1. The inequality $y < 0$ describes the *lower half plane* excluding the x axis.
2. The inequality $x \geq 0$ describes the *right half plane* including the y axis.
3. The inequality $x < y$ describes the half plane to the left of the line through the origin with slope 1.
4. The inequality $x^2 + y^2 \leq 1$ describes the interior of the circle of radius 1 around the origin, including its boundary.

1.6 Factorials and binomial coefficients

Here we briefly discuss notions that arise in mathematical fields such as *Discrete Mathematics*, *Combinatorics* and *Probability*. The most basic problem is as follows. Consider the set X of n students

$$X = \{\underbrace{\text{Andrea}, \text{Marta}, \text{Jim}, \text{Victoria}, \dots, \text{Caterina}}_{n \text{ students}}\}$$

As this is a *set*, there is no importance for the ordering of the students. We could express X also as

$$X = \{\text{Caterina}, \text{Marta}, \text{Victoria}, \text{Andrea}, \dots, \text{Jim.}\}$$

However, the elements of X (the n students) are all distinct, and we often care about their ordering. can ask two natural questions:

1. In how many ways can we order the n students? Above we have seen two examples of how to order (or **permute**) the students. There is a simple formula that gives us the number of *permutations*:

- We start by choosing the first student. There are n students in total, hence we have n students to choose from.
- Next, we want to choose the second student to appear in our ordering. We have already chosen one student, so there are only $n-1$ students left to choose from.
- For the third student we have $n-2$ students to choose from.
- And so on.....
- For the $(n-1)st$ student we have two students to choose from.
- For the nth student we no longer have a choice.

Hence, we find that the number of possible permutations of n elements is

$$n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot 2 \cdot 1.$$

n factorial

Since this is an important formula, it has its own special symbol:

$$n! = n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot 2 \cdot 1$$

which is called n **factorial**. We also define

$$0! = 1.$$

2. In how many ways can we choose k students of the n students? Repeating the same argument as above, we have

- n options for the first student,
- $n-1$ options for the second student,
- $n-2$ options for the third student,
- and so on....
- $n-(k-1) = n-k+1$ options for the kth student.

So we get

$$n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot (n-k+2) \cdot (n-k+1).$$

We observe that this expression can be simplified:

$$n \cdot (n-1) \cdot (n-2) \cdots (n-k+1) = \frac{n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1}{(n-k) \cdot (n-k-1) \cdot (n-k-2) \cdots 2 \cdot 1} = \frac{n!}{(n-k)!}$$

Now we make an important observation: the ordering of the k students *does not matter for us.* So, for example if $k=2$, there is no difference for us between {Jim, Victoria} and {Victoria, Jim}. So we need to eliminate such repetitions. But these repetitions are precisely the number of possible permutations of k elements, which we have seen: it is $k!$. Hence we need to divide the above formula by this number.

n choose k

The number of possible ways to choose k elements from a total of n elements (where $0 \leq k \leq n$) is called n choose k and is denoted

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

This is also called the **binomial coefficient**.

It turns out that these coefficient satisfy certain recursive relations (that is, one can compute $\binom{n}{k}$ from knowledge of $\binom{n-1}{j}$ for all $0 \leq j \leq n-1$). The simple way to visualize this is through *Pascal's triangle*, whose first eight lines look as follows:

$$\begin{array}{ccccccccccccccc}
 & & & & & & & 1 & & & & & & & \\
 & & & & & & 1 & & 1 & & & & & & \\
 & & & & & 1 & & 2 & & 1 & & & & & \\
 & & & & 1 & & 3 & & 3 & & 1 & & & & \\
 & & & 1 & & 4 & & 6 & & 4 & & 1 & & & \\
 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & \\
 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & \\
1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1
\end{array}$$

If we let n denote the line number (starting from 0) and k denote the position of a given number (also, starting from 0) within a line, then the number appearing in the triangle is precisely $\binom{n}{k}$. Hence, for instance,

$$1 = \binom{0}{0} = \binom{n}{0} = \binom{n}{n}$$

for all $n \in \mathbb{N}$, and

$$n = \binom{n}{1} = \binom{n}{n-1}$$

for all $n \in \mathbb{N}_+$. Some other specific examples are

$$6 = \binom{4}{2}, \qquad 10 = \binom{5}{2} = \binom{5}{3}, \qquad 35 = \binom{7}{3} = \binom{7}{4}.$$

Exercise 1.1: Can you see the pattern?

Newton's binomial formula

For any $n \in \mathbb{N}$ and $a, b \in \mathbb{R}$, there holds **Newton's binomial formula**:

$$(a+b)^n = a^n + na^{n-1}b + \cdots + \binom{n}{k}a^{n-k}b^k + \cdots + nab^{n-1} + b^n$$

Exercise 1.2: Prove Newton's binomial formula.

Summation notation

Let $n, N \in \mathbb{Z}$ with $N \geq n$. Let $c_n, \ldots, c_N \in \mathbb{R}$. Then the expression

$$\sum_{k=n}^{N} c_k$$

is a concise way to write the sum

$$c_n + \cdots + c_N.$$

The integers n and N are known as the **lower** and **upper limits of summation**, respectively. The **subscripts** $n, \ldots, N$ are called the **indices**. The numbers c_k $(k = n, \ldots, N)$ are called the **summands**. The summation can also be denoted:

$$\sum_{k=n}^{N} = \sum_{k \in \{n,\ldots,N\}}$$

indicating that the summation is over the set of integer indices $\{n, \ldots, N\}$. The symbol k is known as the **summation index** and it is a **dummy variable**: this means that it can be replaced by any other symbol without changing the meaning of the expression:

$$\sum_{k=n}^{N} c_k = \sum_{j=n}^{N} c_j = \sum_{m=n}^{N} c_m = \sum_{\star=n}^{N} c_\star = \sum_{\clubsuit=n}^{N} c_\clubsuit$$

The summation notation allows us to simplify the expression for Newton's binomial formula:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

Chapter 2

Functions

2.1 Definitions and examples

Definition

Let X and Y be two sets. A **function** f **from** X **to** Y is a rule that associates to any element $x \in X$ at most one element $y \in Y$. The subset of elements in X to which f associates an element in Y is called the domain of f and is denoted $\mathrm{dom}(f)$. We write

$$f : \mathrm{dom}(f) \subseteq X \to Y.$$

For $x \in \mathrm{dom}(f)$, the element $y \in Y$ associated to it by f is called the **image of** x **under** f and is denoted $y = f(x)$. We often write

$$f : x \mapsto f(x).$$

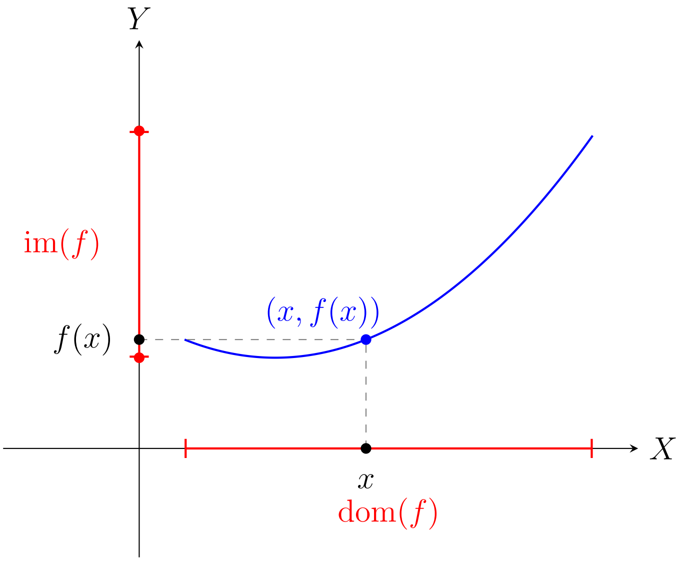
The subset of Y of all images of elements in X is called the **range of** f and is denoted:

$$\mathrm{im}(f) = \{y \in Y \mid \exists x \in \mathrm{dom}(f),\ y = f(x)\}.$$

If $Y = \mathbb{R}$ we say that the function f is **real-valued**. Finally, the **graph of** f is the following subset of the Cartesian product $X \times Y$:

$$\Gamma(f) = \{(x, f(x)) \in X \times Y \mid x \in \mathrm{dom}(f)\}.$$

Here is an example of how we can visualize these properties:



Example 2.1: Some notable examples for $f : \mathbb{R} \to \mathbb{R}$ include:

1. *Linear functions*: $f(x) = ax$ where $a \in \mathbb{R}$, $a \neq 0$. The graph is a straight line through the origin with slope a (the line cannot be vertical).

2. *Affine functions*: $f(x) = ax + b$ where $a, b \in \mathbb{R}$, $a \neq 0$. The graph is a straight line through the point $(0, b)$ with slope a (the line cannot be vertical).

3. *Quadratic functions*: $f(x) = ax^2 + bx + c$ where $a, b, c \in \mathbb{R}$, $a \neq 0$. The graph is a parabola.

4. *Square root*: $f(x) = \sqrt{x}$. This is the first function mentioned here whose domain is not $\mathbb{R}$: $\mathrm{dom}(\sqrt{\ }) = \{x \in \mathbb{R} \mid x \geq 0\}$.

5. *Absolute value*:
$$f : \mathbb{R} \to \mathbb{R}, \qquad f(x) = |x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

6. *Sign function*:
$$f : \mathbb{R} \to \mathbb{Z}, \qquad f(x) = \mathrm{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

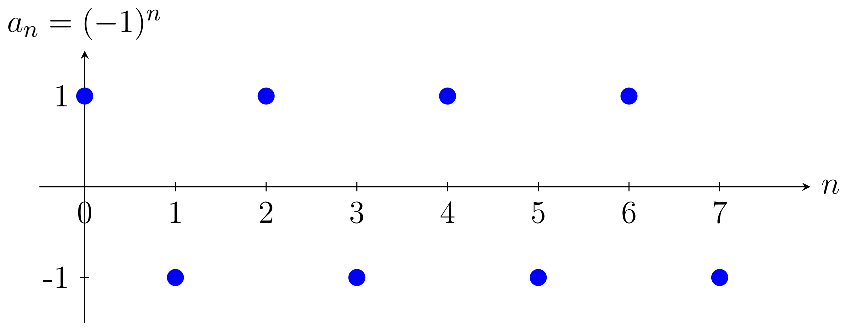
7. *Ceiling* ('rounding up'):
$$f : \mathbb{R} \to \mathbb{Z}, \qquad f(x) = \lceil x \rceil = \text{smallest } n \in \mathbb{Z} \text{ s.t. } n \geq x.$$

8. *Floor* ('rounding down'):
$$f : \mathbb{R} \to \mathbb{Z}, \qquad f(x) = \lfloor x \rfloor = \text{greatest } n \in \mathbb{Z} \text{ s.t. } n \leq x.$$

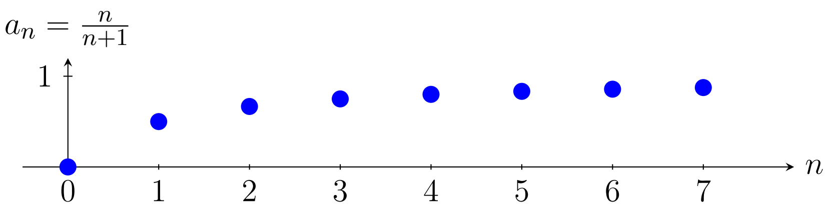
Sequences

A sequence of real numbers $a_0, a_1, a_2, \ldots$ can be viewed as a function $f : \mathrm{dom}(f) \subseteq \mathbb{N} \to \mathbb{R}$, where $f(n) = a_n$ for all $n \in \mathrm{dom}(f)$.

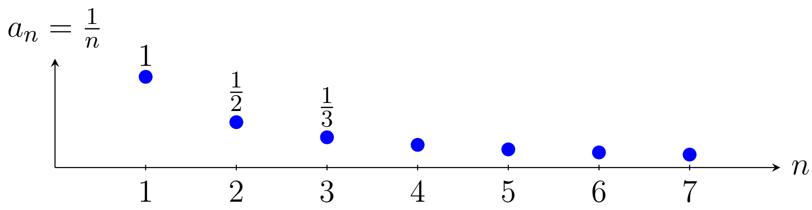
We start with the graph of the sequence $a_n = (-1)^n$, $n \in \mathbb{N}$. This sequence is simply given by $(-1)^n = \begin{cases} 1 & \text{when } n \text{ is even} \\ -1 & \text{when } n \text{ is odd} \end{cases}$ and looks as follows:



Here is the graph of $a_n = \frac{n}{n+1}$, $n \in \mathbb{N}$:



Here is the graph of $a_n = \frac{1}{n}$, with the smaller domain $n \in \mathbb{N}_+$:



We normally denote a sequence $a_0, a_1, \ldots$ as

$$\{a_n\}_{n=0}^{\infty} = a_0, a_1, \ldots$$

If the sequence has a lower index n_1 and an upper index $n_2 > n_1$, this becomes

$$\{a_n\}_{n=n_1}^{n_2} = a_{n_1}, a_{n_1+1}, \ldots, a_{n_2-1}, a_{n_2}$$

2.2 Range and pre-image

Definitions

Let $f : X \to Y$ and let $A \subseteq X$. The **image of A under f** is the subset of Y

$$f(A) = \{f(x) \mid x \in A\} \subseteq \mathrm{im}(f) \subseteq Y.$$

Let $y \in Y$. The **pre-image of y under f** is the subset of X

$$f^{-1}(y) = \{x \in X \mid f(x) = y\} \subseteq \mathrm{dom}(f) \subseteq X.$$

Let $B \subseteq Y$. The **pre-image of B under f** is the subset of X

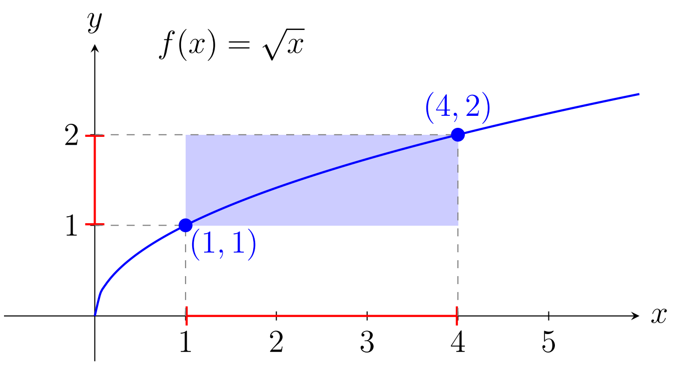
$$f^{-1}(B) = \{x \in X \mid f(x) \in B\} \subseteq \mathrm{dom}(f) \subseteq X.$$

Notice that

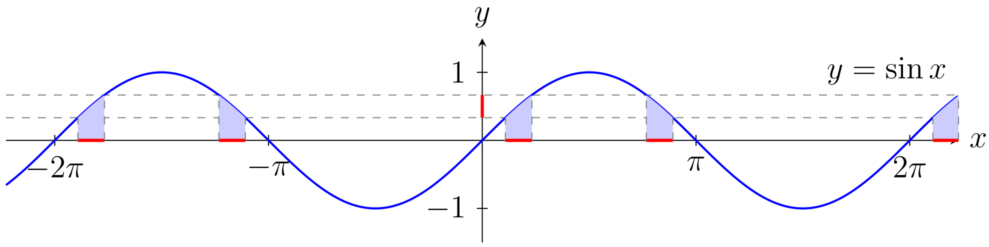
- $f(X) = \mathrm{im}(f)$.
- It is possible that $f^{-1}(y)$ or that $f^{-1}(B)$ are empty. For example, for $f(x) = x^2$, $f^{-1}(-5) = \emptyset$ and $f^{-1}([-4, -2]) = \emptyset$.

Example 2.2: Here are some examples of functions $\mathbb{R} \to \mathbb{R}$:

1. Let f be given by $f(x) = 2x$. Let $A = (a, b)$, where $a < b$. Then $f(A) = (2a, 2b)$. For any $y \in \mathbb{R}$, $f^{-1}(y) = \frac{y}{2}$.
2. Let f be given by $f(x) = 4$. Then for any non-empty $A \subseteq \mathbb{R}$, $f(A) = \{4\}$. Moreover, $f^{-1}(4) = \mathbb{R}$, while $f^{-1}(y) = \emptyset$ for any $y \neq 4$.
3. Let $f(x) = \mathrm{sign}(x)$. Then $f([0, 1]) = \{0, 1\}$, and $f^{-1}(-1) = \mathbb{R}_{-}$. Note that $f(0) = 0$, and $f(\{0\}) = \{0\}$.
4. Let $f(x) = \sqrt{x}$. Then $f((1, 4)) = (1, 2)$, $f^{-1}([1, 2]) = [1, 4]$, $f^{-1}(-1) = \emptyset$.



5. For $f(x) = \sin x$, we can see that $f^{-1}([\frac{1}{3}, \frac{2}{3}])$ is the union of infinitely many intervals.



We can now talk about the supremum, infimum, maximum and minimum of the image of various sets under a real-valued function f:

Supremum and infimum of a real-valued function

Let $f : X \to \mathbb{R}$ be a real-valued function. Let $A \subseteq \mathrm{dom}(f)$. The **supremum of f on** A is the supremum of the image of A under f:

$$\sup_A f = \sup_{x\in A} f(x) = \sup\{f(x) \mid x \in A\}.$$

Similarly, the **infimum of f on** A is the infimum of the image of A under f:

$$\inf_A f = \inf_{x\in A} f(x) = \inf\{f(x) \mid x \in A\}.$$

As we have already seen, the supremum can be an element of $\mathbb{R} \cup \{+\infty\}$ and the infimum can be an element of $\{-\infty\} \cup \mathbb{R}$.

Boundedness of a real-valued function

If $\sup_{x\in A} f(x) < +\infty$ (i.e. it is a real number), we say that f is **bounded from above on** A. If $\inf_{x\in A} f(x) > -\infty$ (i.e. it is a real number), we say that f is **bounded from below on** A. If f is bounded from above and below on A, we say that it is **bounded on** A.

Maximum and minimum of a real-valued function

If $\sup_{x\in A} f(x) < +\infty$ and it belongs to $f(A)$ then it is the **maximum of f on** A. It is denoted

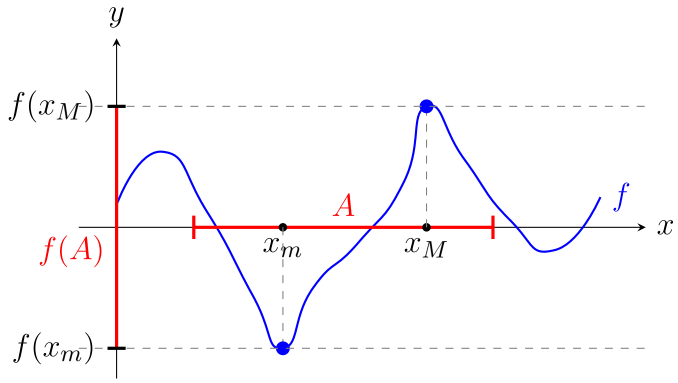
$$\max_A f \qquad \text{or} \qquad \max_{x\in A} f(x).$$

If $\inf_{x\in A} f(x) > -\infty$ and it belongs to $f(A)$ then it is the **minimum of f on** A. It is denoted

$$\min_A f \qquad \text{or} \qquad \min_{x\in A} f(x).$$

Since the minimum and the maximum of f on A belong to $f(A)$, there exist $x_m \in A$ and $x_M \in A$ such that

$$f(x_M) = \max_A f \qquad \text{and} \qquad f(x_m) = \min_A f.$$



Example 2.3: 1. For $\sin x$, we have

$$\max_{x\in\mathbb{R}}(\sin x) = 1 \qquad \text{and} \qquad \min_{x\in\mathbb{R}}(\sin x) = -1.$$

2. For x^2 we have

$$\sup_{x\in\mathbb{R}}(x^2) = +\infty \qquad \text{and} \qquad \min_{x\in\mathbb{R}}(x^2) = 0.$$

If $A = [-10, -3)$ then

$$\max_{x\in A}(x^2) = 100 \qquad \text{and} \qquad \inf_{x\in A}(x^2) = 9.$$

Note that in this last case, the infimum is not achieved, so there's no minimum on A.

2.3 Surjectivity, injectivity, and invertibility

Let us define some important *global* properties of functions, i.e. properties that tell us something about the function as a whole.

Definitions

Let $f : X \to Y$.

- We say that f is **surjective** (or **onto**) if $\mathrm{im}(f) = Y$.
- We say that f is **injective** (or $1-1$, **one-to-one**) if for every $y \in Y$, the subset $f^{-1}(y) \subseteq X$ contains *at most* one element.
- If f is both surjective and injective, it is called a **bijection** (or a **bijective function**).

Let us try to understand these concepts. Surjectivity means that for every $y \in Y$ there exists (at least) one $x \in X$ such that $f(x) = y$. See Figure 2.1.

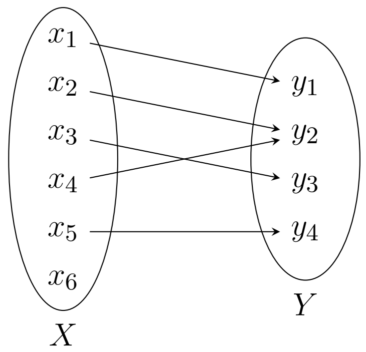


Figure 2.1: A surjective (*onto*) function

Injectivity means that every $y \in Y$ has at most one pre-image. So for any $y \in Y$, either there is no $x \in X$ such that $f(x) = y$ or there is one (and no more) $x \in X$ such that $f(x) = y$. Equivalently, if $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$. See Figure 2.2.

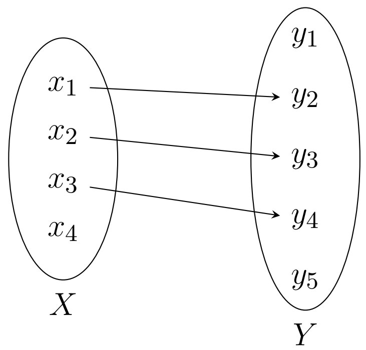


Figure 2.2: An injective (*1-1*) function

Finally, a bijective function f associates to every $x \in X$ exactly one $y \in Y$, and vice versa. See Figure 2.3. In this case we say that the sets X **and** Y **are in one-to-one (*1-1*) correspondence** under f.

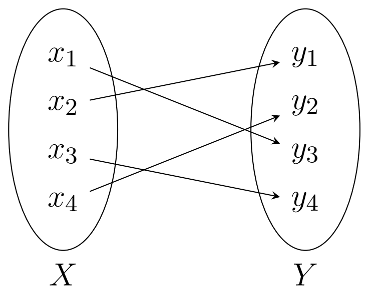


Figure 2.3: A bijective function

The case of a function $f : \mathbb{R} \to \mathbb{R}$

When $f : \mathbb{R} \to \mathbb{R}$ the aforementioned properties can be seen by looking at its graph in $\mathbb{R}^2$:

- f is *surjective* if its graph intersects any horizontal line at least once;
- f is *injective* if its graph intersects any horizontal line at most once;
- f is *bijective* if its graph intersects every horizontal line exactly once.

We observe that the arrows of an injective (1-1) function can be reversed to obtain a function from Y to X. So we define:

Inverse function

Let $f : X \to Y$ be one-to-one. We define its **inverse** $f^{-1} : \mathrm{im}(f) \subseteq Y \to X$ as follows:

$$f^{-1}(y) = x \qquad \text{where } x \text{ is the unique element in } X \text{ satisfying } f(x) = y.$$

Therefore any one-to-one function is also **invertible**. Observe that:

$$\mathrm{dom}(f) = \mathrm{im}(f^{-1}) \qquad \text{and} \qquad \mathrm{im}(f) = \mathrm{dom}(f^{-1})$$

Note that there is some abuse of notation[1] here: $f^{-1}(y)$ is defined to be the *element* $x \in X$, while before (see Section 2.2) we defined $f^{-1}(y)$ to be the *subset* $\{x\} \subseteq X$. In the case of a one-to-one function, this can be forgiven, because there is little practical difference between x and $\{x\}$.

The graph of the inverse function

Recall that the graph of $f : \mathrm{dom}(f) \subseteq X \to Y$ is defined as

$$\Gamma(f) = \{(x, f(x) \in X \times Y \ \mid\ x \in \mathrm{dom}(f)\}.$$

This immediately implies that

$$\begin{aligned}\Gamma(f^{-1}) &= \{(y, f^{-1}(y)) \in Y \times X \ \mid\ y \in \mathrm{dom}(f^{-1})\} \\ &= \{(f(x), x) \in Y \times X \ \mid\ x \in \mathrm{dom}(f)\}.\end{aligned}$$

Comparing the two expressions for $\Gamma(f)$ and for $\Gamma(f^{-1})$, one can see that the graph of f^{-1} is obtained by mirroring the graph of f along the $x = y$ line. We can clearly see this for $f(x) = x^2$, $x \geq 0$, and its inverse $f^{-1}(x) = \sqrt{x}$, see Figure 2.4.

[1] "*Abuse of notation*" means that our notation is not entirely consistent.

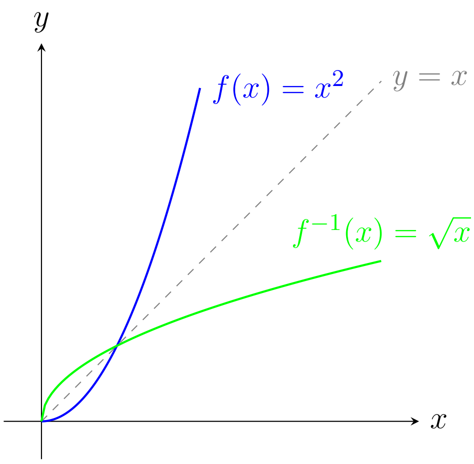


Figure 2.4: $f(x) = x^2$ and its inverse $f^{-1}(x) = \sqrt{x}$, defined on $x \geq 0$, are mirror images with respect to the line $x = y$

Sometimes we want to only look at part of the domain of a function. For example, in the example above, we looked at $f(x) = x^2$ only for $x \geq 0$, so that we could look at its inverse. Otherwise, if we had looked at $x \in \mathbb{R}$, then the preimage of any $y \geq 0$ is $\{+\sqrt{y}, -\sqrt{y}\}$ – i.e., there is no inverse function. What we did was to *restrict* $f(x) = x^2$ to $x \geq 0$:

Restriction of a function

Let $f : X \to Y$ be a function. Let $A \subseteq \mathrm{dom}(f)$ be a subset of the domain of f. The restriction of f to A is a 'new' function $f|_A$ that is defined only on A, where it is identical to f:

$$f|_A : A \to Y \qquad \text{defined as} \qquad f|_A(x) = f(x), \quad \forall x \in A.$$

In Figure 2.4,the blue graph is the graph of the restriction of x^2 to $A = \{x \in \mathbb{R} \ | \ x \geq 0\}$.

2.4 Monotone functions and sequences

Functions that always increase/decrease are of particular interest because they might have important applications. For example, when you study *thermodynamics* you will see that **entropy** is a monotone increasing function of time, meaning that our world always become more disorganized.

Monotone functions $f : \mathbb{R} \to \mathbb{R}$

Monotone functions $f : \mathbb{R} \to \mathbb{R}$

A function $f : \mathbb{R} \to \mathbb{R}$ is said to be **(monotonically) increasing on** $I \subseteq \mathrm{dom}(f)$ if for every $x_1, x_2 \in I$ with $x_1 < x_2$ we have $f(x_1) \leq f(x_2)$:

$$\forall x_1, x_2 \in I, \qquad x_1 < x_2 \quad \Rightarrow \quad f(x_1) \leq f(x_2).$$

The function f is said to be **strictly increasing on** I if

$$\forall x_1, x_2 \in I, \qquad x_1 < x_2 \quad \Rightarrow \quad f(x_1) < f(x_2).$$

Similarly, f is said to be **(monotonically) decreasing on** I if for every $x_1, x_2 \in I$ with $x_1 < x_2$ we have $f(x_1) \geq f(x_2)$:

$$\forall x_1, x_2 \in I, \qquad x_1 < x_2 \quad \Rightarrow \quad f(x_1) \geq f(x_2).$$

The function f is said to be **strictly decreasing on** I if

$$\forall x_1, x_2 \in I, \qquad x_1 < x_2 \quad \Rightarrow \quad f(x_1) > f(x_2).$$

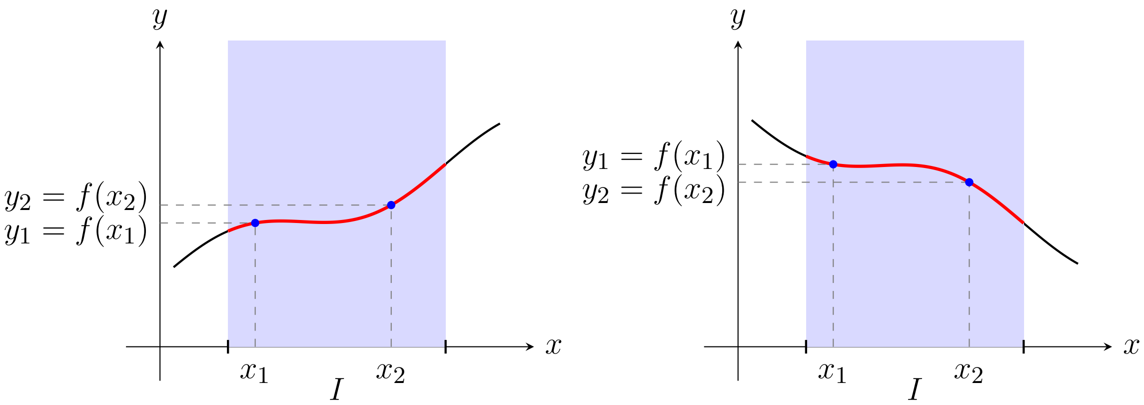


Figure 2.5: An increasing function (left) and a decreasing function (right)

Note that monotonically increasing/decreasing functions are allowed to have slope 0 and even to remain constant. So a constant function $f(x) = c$ is (trivially) both monotone increasing and monotone decreasing. Step functions such as $f(x) = \lceil x \rceil$ or $f(x) = \lfloor x \rfloor$ are monotonically increasing (but not decreasing).

Example 2.4:

1. $f(x) = c$ is monotonically increasing and decreasing.
2. $f(x) = x^2$ is neither increasing nor decreasing on $\mathbb{R}$.
3. $f(x) = x^2$ is strictly increasing on $[0, +\infty)$.
4. $f(x) = x^2$ is strictly decreasing on $(-\infty, 0]$.
5. $f(x) = x^3$ is strictly increasing on $\mathbb{R}$.

6. $f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x^2 & \text{for } x \geq 0 \end{cases}$ is monotonically increasing on $\mathbb{R}$, and it is *strictly* increasing on $[0, +\infty)$.

Proposition 2.1: A function that is strictly increasing/decreasing on its domain is injective (one-to-one).

Proof. Consider first the case that $f : \mathrm{dom}(f) \subseteq \mathbb{R} \to \mathbb{R}$ is strictly increasing on $\mathrm{dom}(f)$. Let $x_1, x_2 \in \mathrm{dom}(f)$ with $x_1 \neq x_2$. Without loss of generality $x_1 < x_2$, so that $f(x_1) < f(x_2)$. In particular, $f(x_1) \neq f(x_2)$, so that f is injective. The case of a strictly decreasing function follows the same idea of proof. □

★★★ The converse statement – i.e. that an injective function is strictly monotone – is not true, see Figure 2.6 for a counterexample.

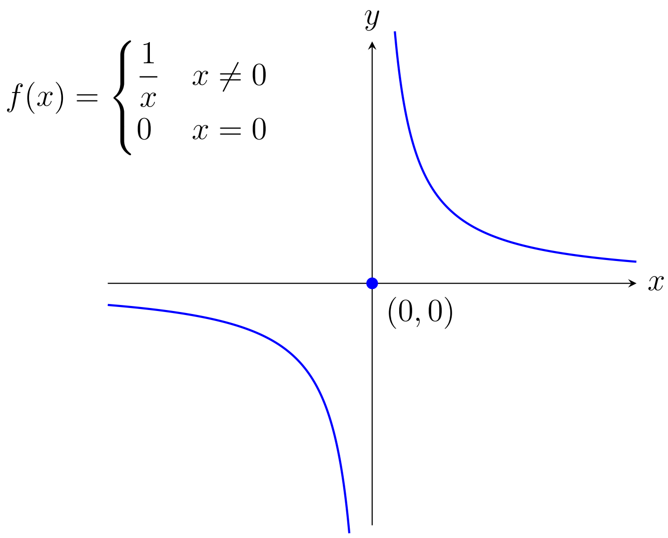


Figure 2.6: A one-to-one function that is neither increasing nor decreasing

Lemma 2.2: Let $f, g : \mathbb{R} \to \mathbb{R}$ be monotonically increasing on some $A \subseteq \mathbb{R}$. Then $f + g$ is also monotonically increasing on A. If either f or g are *strictly* increasing on A, then so is $f + g$. The same statements hold if we replace everywhere the word 'increasing' with the word 'decreasing'.

Proof. Exercise. □

Monotone sequences

A sequence a_n is said to be **(monotonically) increasing on** $\{N, N+1, \dots\}$ if

$$\forall n \geq N, \qquad a_n \leq a_{n+1}.$$

A sequence a_n is said to be **strictly increasing on** $\{N, N+1, \dots\}$ if

$$\forall n \geq N, \qquad a_n < a_{n+1}.$$

A sequence a_n is said to be **(monotonically) decreasing on** $\{N, N+1, \dots\}$ if

$$\forall n \geq N, \qquad a_n \geq a_{n+1}.$$

A sequence a_n is said to be **strictly decreasing on** $\{N, N+1, \dots\}$ if

$$\forall n \geq N, \qquad a_n > a_{n+1}.$$

Example 2.5:

1. The sequence $a_n = \frac{1}{n}$, $n \in \mathbb{N}_+$, is strictly decreasing.
2. The sequence $a_n = \frac{n}{n+1}$, $n \in \mathbb{N}$ is strictly increasing.
3. The sequence $a_n = (-1)^n$, $n \in \mathbb{N}$ is neither increasing nor decreasing.

2.5 Composition of functions

The composition of functions – i.e. the application of two (or more) functions successively – is something that often comes up in mathematics and its applications.

Example 2.6 (Taxi fare)**:** Suppose that the fare for riding a taxi is made of a flat fee of 3 Euros plus twice the distance travelled (in kilometers). So the fee for riding x kilometers is:

$$f(x) = 2x + 3.$$

Now, suppose that a card payment carries a 5% surcharge of the total fare:

$$g(y) = 1.05y.$$

So, if we travel x kilometers, and want to pay by card, the total amount to pay is:

$$g(f(x)) = 1.05(2x + 3) = 2.1x + 3.15$$

This is a composition of functions.

Let X, Y, Z be sets and let $f : X \to Y$ and $g : Y \to Z$ be two functions. The **composition of f and g** is a new function $h : X \to Z$ defined as

$$h(x) = g(f(x)).$$

It is denoted by $h = g \circ f$ so we can also write $h(x) = (g \circ f)(x)$.

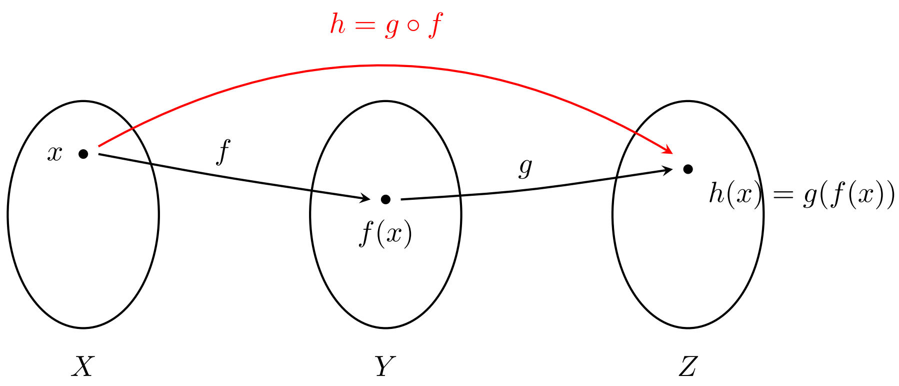


Figure 2.7: Composition of functions $h(x) = g(f(x))$

The domain of h is defined as follows:

$$x \in \mathrm{dom}(h) \qquad \Leftrightarrow \qquad x \in \mathrm{dom}(f) \quad \text{and} \quad f(x) \in \mathrm{dom}(g).$$

Example 2.7:

1. If $f(x) = x^2$, $g(y) = y - 3$, then $h(x) = g(f(x)) = x^2 - 3$, and $\mathrm{dom}(h) = \mathrm{dom}(f)$.
2. If $f(x) = e^x$, $g(y) = -y$, then $h(x) = g(f(x)) = -e^x$, and $\mathrm{dom}(h) = \mathrm{dom}(f)$.
3. If $f(x) = -x$, $g(y) = e^y$, then $h(x) = g(f(x)) = e^{-x}$, and $\mathrm{dom}(h) = \mathrm{dom}(f)$.
4. If $f(x) = \sqrt{x}$, $g(y) = y^2$, then $h(x) = g(f(x)) = (\sqrt{x})^2 = x$, and $\mathrm{dom}(h) = \mathrm{dom}(f) = [0, +\infty)$.
5. If $f(x) = x^2$, $g(y) = \sqrt{y}$, then $h(x) = g(f(x)) = \sqrt{x^2} = |x|$, and $\mathrm{dom}(h) = \mathrm{dom}(f) = \mathbb{R}$.
6. If $f(x) = \frac{1}{x}$, $g(y) = \sin y$, then $h(x) = g(f(x)) = \sin\frac{1}{x}$, and $\mathrm{dom}(h) = \mathrm{dom}(f) = \mathbb{R} \setminus \{0\}$.
7. If $f(x) = \sin x$, $g(y) = \frac{1}{y}$, then $h(x) = g(f(x)) = \frac{1}{\sin x}$, and $\mathrm{dom}(h) \neq \mathrm{dom}(f)$. In this case $\mathrm{dom}(h)$ is all $x \in \mathbb{R}$ s.t. $\sin x \neq 0$.

These examples show us that the **composition of functions is not a commutative operation**:

$$f \circ g \neq g \circ f.$$

We can also see that a function and its inverse 'cancel' one another. More precisely, if f is one-to-one (and therefore f^{-1} exists) then

$$f \circ f^{-1} = \mathrm{Id}_{\mathrm{dom}(f^{-1})} = \mathrm{Id}_{\mathrm{im}(f)} \qquad \text{and} \qquad f^{-1} \circ f = \mathrm{Id}_{\mathrm{dom}(f)}$$

where the function Id is defined as follows:

The identity function

Id is the identity function that can be defined on any subset:

$$\mathrm{Id}_A(x) = x, \qquad \forall x \in A.$$

Here A could be a subset of any set.

Proposition 2.3: Let $f : X \to Y$ and $g : Y \to Z$ both be one-to-one functions. Then

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

Proof. First, we show that $g \circ f$ is one-to-one, so its inverse exists. Let $x_1, x_2 \in X$ with $(g \circ f)(x_1) = (g \circ f)(x_2)$. Then $g(f(x_1)) = g(f(x_2))$. Since g is one-to-one, $f(x_1) = f(x_2)$. Since f is one-to-one, $x_1 = x_2$. Thus $g \circ f$ is one-to-one.

Now we show $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. Let $z \in \mathrm{dom}((g \circ f)^{-1}) = \mathrm{im}(g \circ f)$. Then there exists a unique $x \in X$ such that $(g \circ f)(x) = z$, and $(g \circ f)^{-1}(z) = x$. We have $g(f(x)) = z$, so $f(x) = g^{-1}(z)$, and thus $x = f^{-1}(g^{-1}(z)) = (f^{-1} \circ g^{-1})(z)$. Therefore $(g \circ f)^{-1}(z) = (f^{-1} \circ g^{-1})(z)$ for all $z \in \mathrm{dom}((g \circ f)^{-1})$. □

Lemma 2.4: Let $f, g : \mathbb{R} \to \mathbb{R}$. Then:

$$\begin{aligned}
f, g \text{ are both monotone increasing} \quad &\Rightarrow \quad g \circ f \text{ is monotone increasing.}\\
f, g \text{ are both monotone decreasing} \quad &\Rightarrow \quad g \circ f \text{ is monotone increasing.}\\
f, g \text{ are monotone of different kinds} \quad &\Rightarrow \quad g \circ f \text{ is monotone decreasing.}
\end{aligned}$$

Proof. Exercise. □

Translations, rescalings and reflections

Here we discuss three types of simple functions that might often appear as part of a composition of functions.

Translations are an important family of functions: they simply 'move' the variable x by some fixed amount. Here is the simple definition. Given a fixed $c \in \mathbb{R}$, define the **translation by** c, the function $t_c : \mathbb{R} \to \mathbb{R}$, as

$$t_c(x) = x + c.$$

Here x is the variable, as always, and the number c is normally called a *parameter*. Given some function $f : \mathbb{R} \to \mathbb{R}$, we see that

$$\begin{aligned}
(f \circ t_c)(x) &= f(t_c(x)) = f(x + c),\\
(t_c \circ f)(x) &= t_c(f(x)) = f(x) + c.
\end{aligned}$$

The first function is a shift of the graph of f to the left by c (if $c < 0$ then the shift is to the right). The second function is a shift of the graph of f up by c (if $c < 0$ then the shift is down).

Rescalings 'squeeze' or 'stretch' a function horizontally. Here's the definition. Given a fixed $c > 0$, define the **scaling by** c, the function $s_c : \mathbb{R} \to \mathbb{R}$, as

$$s_c(x) = cx.$$

Composition with some function $f : \mathbb{R} \to \mathbb{R}$ gives

$$(f \circ s_c)(x) = f(s_c(x)) = f(cx).$$

If $0 < c < 1$ then the graph of f is 'stretched' by a factor of $\frac{1}{c}$, whereas if $c > 1$ then the graph of f is 'squeezed' by a factor of c, see Figure 2.8.

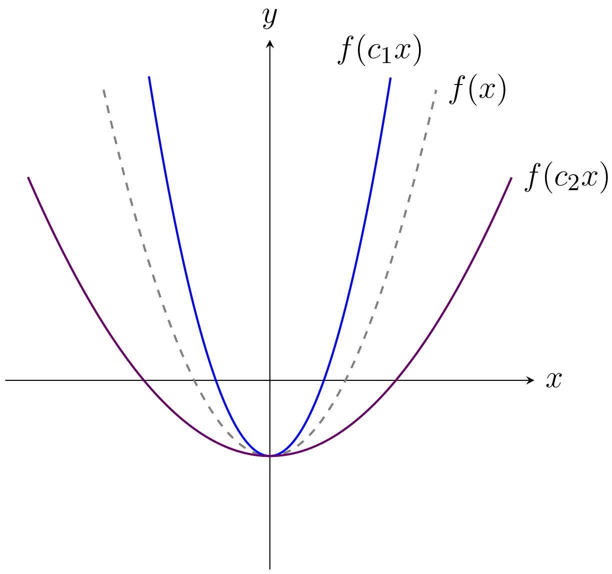


Figure 2.8: Rescalings of a function $f(x)$ by $c_1 > 1$ and $0 < c_2 < 1$.

We can also rescale a function vertically. Given a fixed $c > 0$, define

$$(s_c \circ f)(x) = s_c(f(x)) = cf(x).$$

A constant $c > 1$ will 'stretch' the graph of the function along the y-axis, and $0 < c < 1$ will 'squeeze' the graph of the function.

Reflection of a function 'flips' the graph of a function along the y-axis. Define the **reflection function** $r : \mathbb{R} \to \mathbb{R}$ as

$$r(x) = -x.$$

Then

$$(f \circ r)(x) = f(r(x)) = f(-x).$$

Switching the order of composition of the two functions will result in flipping the graph along the x-axis:

$$(r \circ f)(x) = r(f(x)) = -f(x).$$

2.6 Elementary functions

Even, odd and periodic functions

Even and odd functions

Let $f : \mathrm{dom}(f) \subseteq \mathbb{R} \to \mathbb{R}$ and assume that if $x \in \mathrm{dom}(f)$ then also $-x \in \mathrm{dom}(f)$. We say that f is **even** if

$$f(x) = f(-x), \qquad \forall x \in \mathrm{dom}(f).$$

We say that f is **odd** if

$$f(x) = -f(-x), \qquad \forall x \in \mathrm{dom}(f).$$

Periodic functions

A function $f : \mathrm{dom}(f) \subseteq \mathbb{R} \to \mathbb{R}$ is said to have **period** $p > 0$ if $\forall x \in \mathrm{dom}(f)$, $\{x + np \mid n \in \mathbb{Z}\} \subseteq \mathrm{dom}(f)$ and if

$$f(x) = f(x + np), \qquad \forall x \in \mathrm{dom}(f),\ \forall n \in \mathbb{Z}.$$

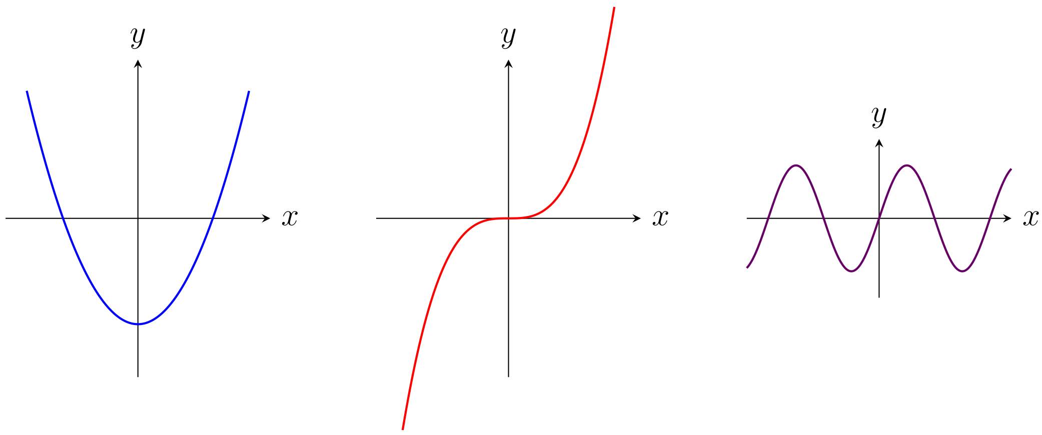


Figure 2.9: Examples of an even function (left), an odd function (middle), and a periodic function (right).

2.6.1 Powers

Non-negative integer powers

Power functions

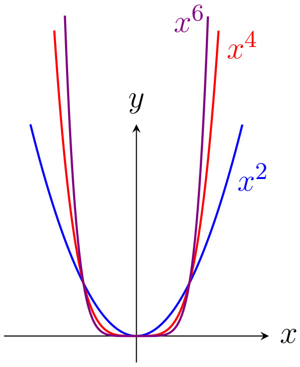
A (non-negative integer) **power function** is a function of the form

$$f(x) = x^n, \qquad n \in \mathbb{N}.$$

The domain is $\mathrm{dom}(f) = \mathbb{R}$ for all $n \in \mathbb{N}$.

Properties

- n even $\Rightarrow f(x) = x^n$ is an even function: $f(-x) = (-x)^n = x^n = f(x)$.
- n odd $\Rightarrow f(x) = x^n$ is an odd function: $f(-x) = (-x)^n = -x^n = -f(x)$.
- For $n \geq 1$, $f(x) = x^n$ is strictly increasing on $[0, +\infty)$.
- For even $n \geq 2$, $f(x) = x^n$ is strictly decreasing on $(-\infty, 0]$ and strictly increasing on $[0, +\infty)$.
- For odd $n \geq 1$, $f(x) = x^n$ is strictly increasing on $\mathbb{R}$.



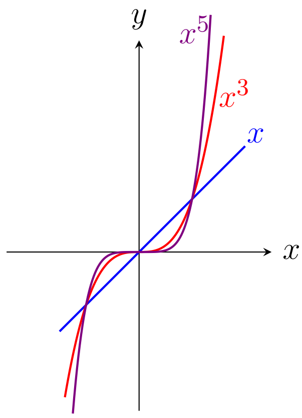


Figure 2.10: Even powers (left) and odd powers (right)

Positive rational powers

qth root

The power function $f(x) = x^{1/q}$ where $q \in \{2, 3, \dots\}$ is called the **qth root of** x and is denoted

$$f(x) = \sqrt[q]{x}, \qquad \operatorname{dom}(f) = \begin{cases} \mathbb{R} & q \text{ is odd} \\ [0, +\infty) & q \text{ is even} \end{cases}$$

It is the inverse function of $y = x^q$.

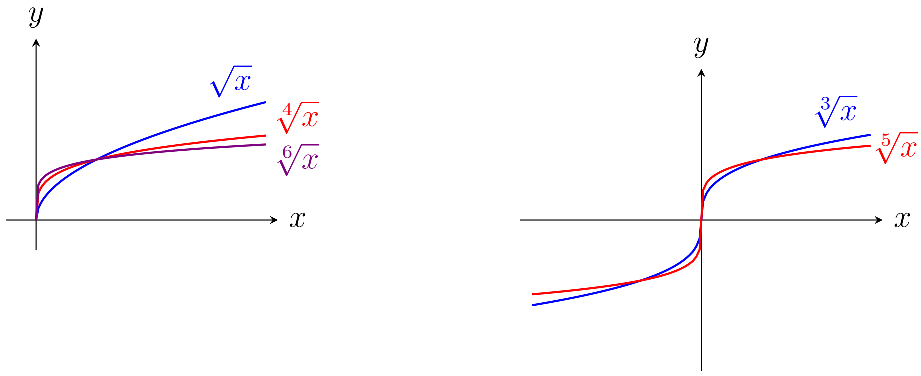


Figure 2.11: Roots with even index (left) and odd index (right)

Positive rational power

Let $p, q \in \mathbb{N}$ have no common divisors. A power function x^r with $r = \frac{p}{q}$ is defined as follows:

$$f(x) = x^r = x^{p/q} = (x^p)^{1/q} = \sqrt[q]{x^p}, \qquad \mathrm{dom}(f) = \begin{cases} \mathbb{R} & q \text{ is odd} \\ [0, +\infty) & q \text{ is even} \end{cases}$$

Positive irrational powers

It is not obvious how to define an irrational power. For instance, what is the meaning of

$$\pi^{\sqrt{2}} = ?$$

Since we know that $\sqrt{2} = 1.414213562\cdots$ we can define the number $\pi^{\sqrt{2}}$ to be the number that we approach by taking better and better approximations of $\sqrt{2}$:

$$\begin{aligned} &\pi^{1.4} \\ &\pi^{1.41} \\ &\pi^{1.414} \\ &\pi^{1.4142} \\ &\pi^{1.41421} \\ &\pi^{1.414213} \\ &\vdots \\ &\pi^{\sqrt{2}} \end{aligned}$$

However this is not *a priori* simple. We skip this important problem in this course.

Positive irrational power

A power function x^s with $s \in \mathbb{R}_+ \setminus \mathbb{Q}$ is defined as the 'limit' of x^r where $r \in \mathbb{Q}$ and r approaches s. The **domain is** $[0, +\infty)$.

Negative powers

Negative power functions

- For $n \in \mathbb{N}_+$ the negative power function is

$$f(x) = x^{-n} = \frac{1}{x^n}, \qquad \mathrm{dom}(f) = \mathbb{R} \setminus \{0\}.$$

- For $r = \frac{p}{q} \in \mathbb{Q}$ (with $p, q \in \mathbb{N}_+$ having no common divisors)

$$f(x) = x^{-r} = \frac{1}{\sqrt[q]{x^p}}, \qquad \mathrm{dom}(f) = \begin{cases} \mathbb{R} \setminus \{0\} & q \text{ is odd} \\ (0, +\infty) & q \text{ is even} \end{cases}$$

- For $s \in \mathbb{R}_+ \setminus \mathbb{Q}$ the negative power function is

$$f(x) = x^{-s} = \frac{1}{x^s}, \qquad \mathrm{dom}(f) = (0, +\infty).$$

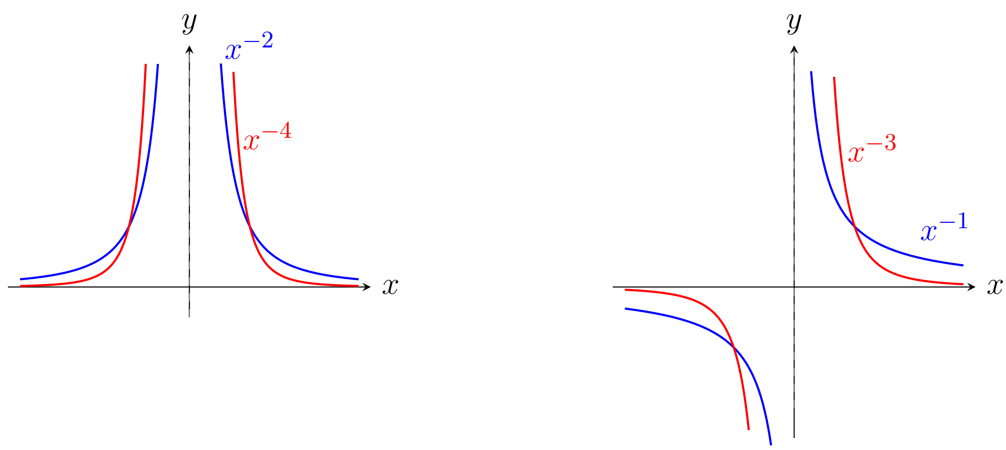


Figure 2.12: Negative powers with even exponent (left) and odd exponent (right)

Properties of negative integer powers

- For $n \in \mathbb{N}_+$, $f(x) = x^{-n}$ is strictly decreasing on $(0, +\infty)$ and on $(-\infty, 0)$.
- If n is even, $f(x) = x^{-n}$ is an even function.
- If n is odd, $f(x) = x^{-n}$ is an odd function.

Properties of real powers

For $x > 0$ and $r, s \in \mathbb{R}$:

- $f(x) = x^r$ is strictly increasing on $(0, +\infty)$ if $r > 0$.
- $x^r \cdot x^s = x^{r+s}$
- $\frac{x^r}{x^s} = x^{r-s}$
- $(x^r)^s = x^{rs}$
- For $0 < r < 1$, the function x^r is concave (bends downward).
- For $r > 1$, the function x^r is convex (bends upward).

2.6.2 Polynomials and rational functions

Polynomial functions

A **polynomial function** is a function of the form

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

where $a_0, a_1, \ldots, a_n \in \mathbb{R}$ are constants and $a_n \neq 0$. The number n is called the **degree** of the polynomial, denoted $\deg(p) = n$. The domain is $\mathrm{dom}(p) = \mathbb{R}$.

Example 2.8:

1. Constant functions: $p(x) = c$ (degree 0).
2. Linear functions: $p(x) = ax + b$ with $a \neq 0$ (degree 1).
3. Quadratic functions: $p(x) = ax^2 + bx + c$ with $a \neq 0$ (degree 2).
4. Cubic functions: $p(x) = ax^3 + bx^2 + cx + d$ with $a \neq 0$ (degree 3).

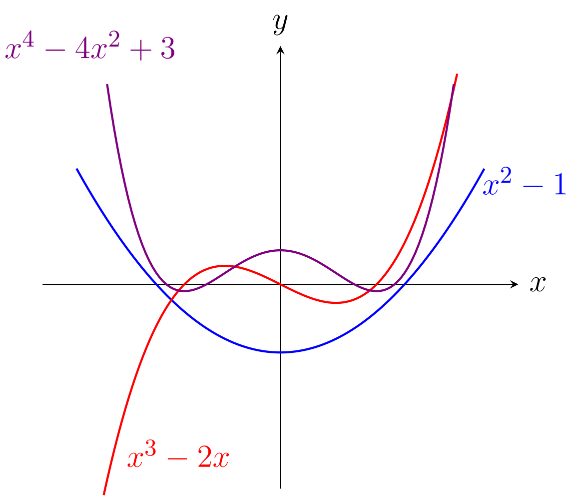


Figure 2.13: Examples of polynomial functions

Properties of polynomials

- The sum and product of two polynomials is a polynomial.
- If p has degree n, then p has at most n real roots (zeros).

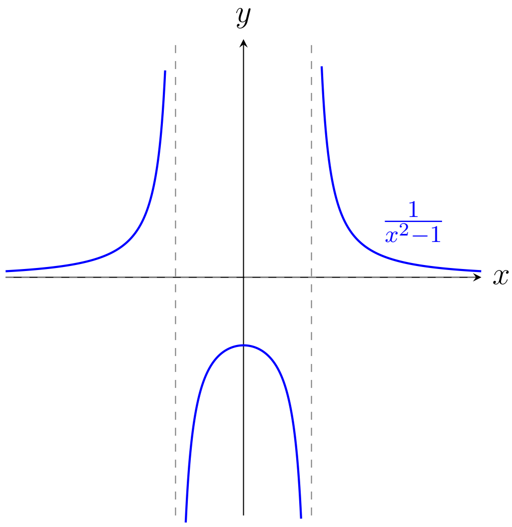
Rational functions

A **rational function** is a quotient of two polynomials:

$$r(x) = \frac{p(x)}{q(x)}$$

where p and q are polynomials with $q \not\equiv 0$. If p and q have no common factors then the domain is

$$\mathrm{dom}(r) = \{x \in \mathbb{R} \mid q(x) \neq 0\}.$$



2.6.3 Exponential and logarithmic functions

The exponential function

For $a > 0$, $a \neq 1$, the **exponential function with base** a is

$$f(x) = a^x$$

with $\mathrm{dom}(f) = \mathbb{R}$ and $\mathrm{im}(f) = (0, +\infty)$. The most important case is $a = e$, where $e \approx 2.71828\ldots$ is Euler's number. We write

$$\exp(x) = e^x.$$

Properties of exponential functions

For $a > 0$, $a \neq 1$:

- a^x is strictly increasing if $a > 1$ and strictly decreasing if $0 < a < 1$.
- $a^0 = 1$ for all $a > 0$.
- $a^x > 0$ for all $x \in \mathbb{R}$.

Algebraic rules for exponentials

For $a, b > 0$ and $x, y \in \mathbb{R}$:

$$
\begin{aligned}
a^x \cdot a^y &= a^{x+y} \\
\frac{a^x}{a^y} &= a^{x-y} \\
(a^x)^y &= a^{xy} \\
a^x b^x &= (ab)^x \\
a^{-x} &= \frac{1}{a^x}
\end{aligned}
$$

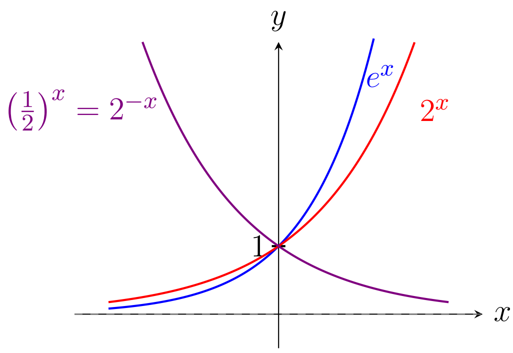


Figure 2.14: Exponential functions with different bases

The logarithmic function

For $a > 0$, $a \neq 1$, the **logarithm with base** a is the inverse of a^x:

$$f(x) = \log_a(x)$$

with $\mathrm{dom}(f) = (0, +\infty)$ and $\mathrm{im}(f) = \mathbb{R}$. The most important case is $a = e$, called the **natural logarithm**:

$$\ln(x) = \log_e(x).$$

Properties of logarithms

For $a > 0$, $a \neq 1$:

- $\log_a(x)$ is strictly increasing if $a > 1$ and strictly decreasing if $0 < a < 1$.
- $\log_a(1) = 0$ and $\log_a(a) = 1$ for all $a > 0$, $a \neq 1$.
- $\log_a(a^x) = x$ for all $x \in \mathbb{R}$ and $a^{\log_a(x)} = x$ for all $x > 0$.

Algebraic rules for logarithms

For $a > 0$, $a \neq 1$, and $x, y > 0$:

$$\begin{aligned}
\log_a(xy) &= \log_a(x) + \log_a(y) \\
\log_a\left(\frac{x}{y}\right) &= \log_a(x) - \log_a(y) \\
\log_a(x^r) &= r\log_a(x) \qquad \text{for } r \in \mathbb{R} \\
\log_a(x) &= \frac{\log_b(x)}{\log_b(a)} \qquad \text{(change of base, } b > 0,\ b \neq 1)
\end{aligned}$$

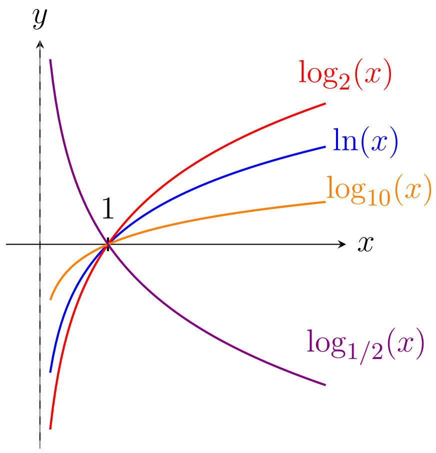


Figure 2.15: Logarithmic functions with different bases

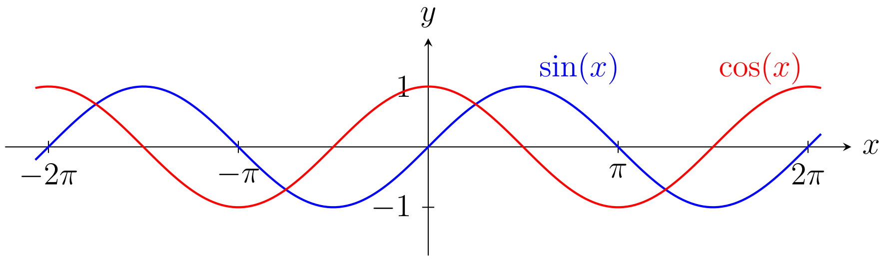
2.6.4 Trigonometric functions and their inverses

Sine and cosine

The **sine** and **cosine** functions are defined as

$$f(x) = \sin(x), \qquad g(x) = \cos(x)$$

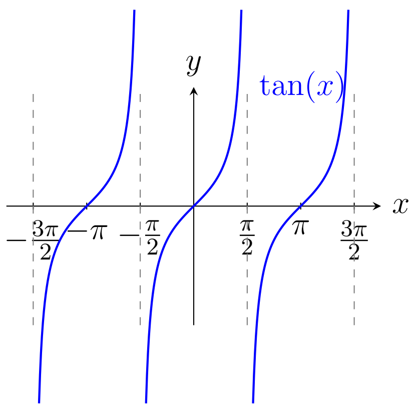
with $\mathrm{dom}(f) = \mathrm{dom}(g) = \mathbb{R}$ and $\mathrm{im}(f) = \mathrm{im}(g) = [-1, 1]$. Both functions are periodic with period 2π.

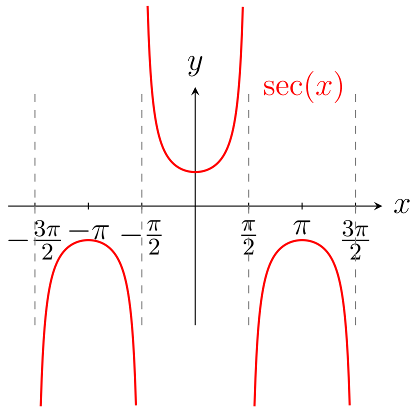


Other trigonometric functions

The other trigonometric functions are defined in terms of sine and cosine:

$$\begin{aligned}
\tan(x) &= \frac{\sin(x)}{\cos(x)} & \quad \operatorname{dom}(\tan) &= \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \;\middle|\; k \in \mathbb{Z} \right\} \\
\cot(x) &= \frac{\cos(x)}{\sin(x)} & \quad \operatorname{dom}(\cot) &= \mathbb{R} \setminus \{ k\pi \mid k \in \mathbb{Z} \} \\
\sec(x) &= \frac{1}{\cos(x)} & \quad \operatorname{dom}(\sec) &= \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \;\middle|\; k \in \mathbb{Z} \right\} \\
\csc(x) &= \frac{1}{\sin(x)} & \quad \operatorname{dom}(\csc) &= \mathbb{R} \setminus \{ k\pi \mid k \in \mathbb{Z} \}
\end{aligned}$$





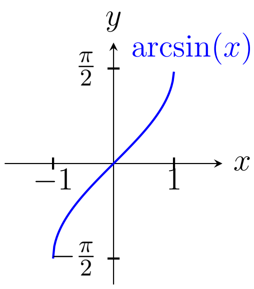
Fundamental trigonometric identities

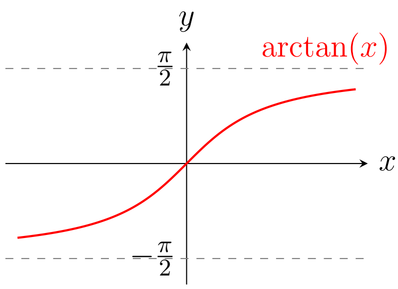
$$\begin{aligned}
\sin^2(\alpha) + \cos^2(\alpha) &= 1 \\
\tan^2(\alpha) + 1 &= \sec^2(\alpha) \\
1 + \cot^2(\alpha) &= \csc^2(\alpha) \\
\sin(\alpha \pm \beta) &= \sin(\alpha)\cos(\beta) \pm \cos(\alpha)\sin(\beta) \\
\cos(\alpha \pm \beta) &= \cos(\alpha)\cos(\beta) \mp \sin(\alpha)\sin(\beta) \\
\sin(2\alpha) &= 2\sin(\alpha)\cos(\alpha) \\
\cos(2\alpha) &= \cos^2(\alpha) - \sin^2(\alpha) = 2\cos^2(\alpha) - 1 = 1 - 2\sin^2(\alpha) \\
\sin(-\alpha) &= -\sin(\alpha) \qquad \text{(sine is odd)} \\
\cos(-\alpha) &= \cos(\alpha) \qquad \text{(cosine is even)} \\
\sin(\alpha) - \sin(\beta) &= 2\cos\left(\frac{\alpha+\beta}{2}\right)\sin\left(\frac{\alpha-\beta}{2}\right) \\
\cos(\alpha) - \cos(\beta) &= -2\sin\left(\frac{\alpha+\beta}{2}\right)\sin\left(\frac{\alpha-\beta}{2}\right)
\end{aligned}$$

Inverse trigonometric functions

The inverse trigonometric functions are defined by restricting the domains appropriately:

- $\arcsin : [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$ is the inverse of sin restricted to $[-\frac{\pi}{2}, \frac{\pi}{2}]$.
- $\arccos : [-1, 1] \to [0, \pi]$ is the inverse of cos restricted to $[0, \pi]$.
- $\arctan : \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$ is the inverse of tan restricted to $(-\frac{\pi}{2}, \frac{\pi}{2})$.





Here are some notable values of the trigonometric functions:

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	π
$\sin\theta$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1	0
$\cos\theta$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0	-1
$\tan\theta$	0	$\frac{\sqrt{3}}{3}$	1	$\sqrt{3}$	-	0

Chapter 3

Complex Numbers

Complex numbers extend the real number system to solve equations that have no real solutions, such as $x^2 + 1 = 0$. They are essential in electrical engineering (AC circuit analysis, signal processing), mechanical engineering (vibrations, control systems), and many other fields. The most important mathematical property of complex numbers is that they provide a complete system where every non-constant polynomial equation has a solution. This is known as the **Fundamental Theorem of Algebra**. For example, the equation $x^2 + 1 = 0$ has no real solution, but has two complex solutions: $x = i$ and $x = -i$.

Definition

A **complex number** is an expression of the form $z = a + bi$, where $a, b \in \mathbb{R}$ and i is the **imaginary unit** satisfying $i^2 = -1$. The set of all complex numbers is denoted by $\mathbb{C}$.

- a is called the **real part** of z, denoted $\mathrm{Re}(z)$
- b is called the **imaginary part** of z, denoted $\mathrm{Im}(z)$
- If $b = 0$, then z is a real number
- If $a = 0$, then z is called a **purely imaginary** number

3.1 Algebraic operations

Algebraic operations

Let $z = a + bi$ and $w = c + di$ be complex numbers.

$$\begin{aligned}
\textbf{Addition:}\quad & z + w = (a + c) + (b + d)i \\
\textbf{Subtraction:}\quad & z - w = (a - c) + (b - d)i \\
\textbf{Multiplication:}\quad & z \cdot w = (ac - bd) + (ad + bc)i \\
\textbf{Division:}\quad & \frac{z}{w} = \frac{(ac + bd) + (bc - ad)i}{c^2 + d^2}, \quad w \neq 0 \\
\textbf{Inverse:}\quad & z^{-1} = \frac{1}{z} = \frac{a - bi}{a^2 + b^2}, \quad z \neq 0
\end{aligned}$$

Example 3.1: Let $z = 2 + 3i$ and $w = 1 - 2i$. Then:

$$\begin{aligned}
z + w &= (2 + 1) + (3 - 2)i = 3 + i \\
z \cdot w &= (2 \cdot 1 - 3 \cdot (-2)) + (2 \cdot (-2) + 3 \cdot 1)i = (2 + 6) + (-4 + 3)i = 8 - i \\
\frac{z}{w} &= \frac{(2 \cdot 1 + 3 \cdot (-2)) + (3 \cdot 1 - 2 \cdot (-2))i}{1^2 + (-2)^2} = \frac{(2 - 6) + (3 + 4)i}{5} = \frac{-4 + 7i}{5} \\
z^{-1} &= \frac{1}{2 + 3i} = \frac{2 - 3i}{2^2 + 3^2} = \frac{2 - 3i}{13} = \frac{2}{13} - \frac{3}{13}i
\end{aligned}$$

Example 3.2: Let $z = 4 - i$ and $w = 2 + 3i$. Compute z^2 and $\frac{w}{z}$:

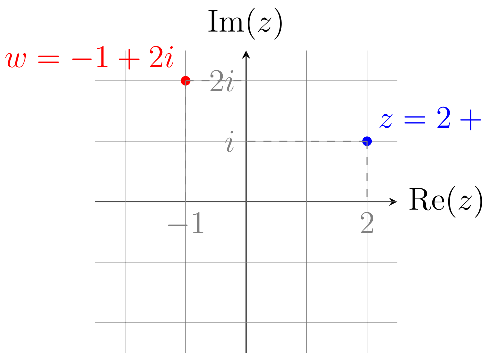
$$\begin{aligned}
z^2 &= (4 - i)^2 = 16 - 8i + i^2 = 16 - 8i - 1 = 15 - 8i \\
\frac{w}{z} &= \frac{2 + 3i}{4 - i} = \frac{(2 + 3i)(4 + i)}{(4 - i)(4 + i)} = \frac{(8 - 3) + (2 + 12)i}{4^2 - i^2} = \frac{5 + 14i}{17} = \frac{5}{17} + \frac{14}{17}i
\end{aligned}$$

3.2 Cartesian coordinates

Complex plane

The complex number $z = a + bi$ can be represented as a point (a, b) in the **complex plane**:

- The horizontal axis represents the real part ($\mathrm{Re}(z)$)
- The vertical axis represents the imaginary part ($\mathrm{Im}(z)$)



Key Properties of Complex Numbers

For $z = a + bi$:

- The **modulus of** z (absolute value) is
$$|z| = \sqrt{a^2 + b^2}.$$

- The **argument of** z is
$$\arg(z) = \theta$$
where $\theta = \arctan \frac{b}{a}$.

- The **complex conjugate of** z is
$$\bar{z} = a - bi.$$

- The real and imaginary parts can be expressed as:
$$\mathrm{Re}(z) = \frac{z + \bar{z}}{2}, \qquad \mathrm{Im}(z) = \frac{z - \bar{z}}{2i}$$

Example 3.3: For $z = 3 + 4i$:

$$\begin{aligned} |z| &= \sqrt{3^2 + 4^2} = \sqrt{9 + 16} = \sqrt{25} = 5 \\ \arg(z) &= \arctan\left(\frac{4}{3}\right) \approx 0.9273 \text{ radians} \\ \bar{z} &= 3 - 4i \end{aligned}$$

We can also verify the real and imaginary parts (though we already know they are 3 and 4, respectively:

$$\begin{aligned} \mathrm{Re}(z) &= \frac{(3 + 4i) + (3 - 4i)}{2} = \frac{6}{2} = 3 \\ \mathrm{Im}(z) &= \frac{(3 + 4i) - (3 - 4i)}{2i} = \frac{8i}{2i} = 4 \end{aligned}$$

3.3 Trigonometric and exponential form

Polar form

A complex number $z = a + bi$ can be written in **trigonometric form**:

$$z = r(\cos\theta + i\sin\theta)$$

where $r = |z|$ and $\theta = \arctan\frac{b}{a}$. Observe that this expression is 2π**-periodic as a function of** θ.

Euler's formula and identity

Euler's formula

$$e^{i\theta} = \cos\theta + i\sin\theta$$

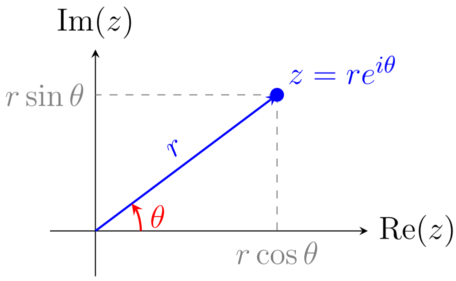
converts the trigonometric form of z into **exponential form**

$$z = re^{i\theta}.$$

For $z = -1$ this gives **Euler's identity**

$$e^{i\pi} + 1 = 0,$$

which is the only simple formula in mathematics that contains the five major constants $0, 1, i, \pi, e$.



Operations in exponential form

Let $z = re^{i\theta}$ and $w = se^{i\phi}$ be complex numbers in exponential form.

Multiplication: $z \cdot w = rs\, e^{i(\theta+\phi)}$

Division: $\dfrac{z}{w} = \dfrac{r}{s}\, e^{i(\theta-\phi)}, \quad w \neq 0$

Inverse: $z^{-1} = \dfrac{1}{r}\, e^{-i\theta}, \quad z \neq 0$

Conjugate: $\bar{z} = r\, e^{-i\theta}$

Example 3.4: Let $z = 2e^{i\pi/3}$ and $w = 3e^{i\pi/6}$. Then:

$$\begin{aligned}
z \cdot w &= (2 \cdot 3)e^{i(\pi/3+\pi/6)} = 6e^{i\pi/2} = 6i \\
\frac{z}{w} &= \frac{2}{3}e^{i(\pi/3-\pi/6)} = \frac{2}{3}e^{i\pi/6} = \frac{2}{3}(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}) = \frac{\sqrt{3}}{3} + \frac{1}{3}i \\
z^{-1} &= \frac{1}{2}e^{-i\pi/3} = \frac{1}{2}(\cos\frac{\pi}{3} - i\sin\frac{\pi}{3}) = \frac{1}{4} - i\frac{\sqrt{3}}{4} \\
\bar{z} &= 2e^{-i\pi/3} = 2(\cos\frac{\pi}{3} - i\sin\frac{\pi}{3}) = 1 - i\sqrt{3}
\end{aligned}$$

Example 3.5: Convert $z = 1 + i$ to exponential form:

$$\begin{aligned}
|z| &= \sqrt{1^2 + 1^2} = \sqrt{2} \\
\arg(z) &= \frac{\pi}{4} \\
z &= \sqrt{2}(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}) = \sqrt{2}e^{i\pi/4}
\end{aligned}$$

Example 3.6: Convert $z = -2 + 2i\sqrt{3}$ to exponential form:

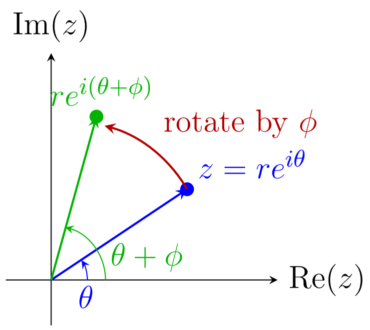
$$\begin{aligned}
|z| &= \sqrt{(-2)^2 + (2\sqrt{3})^2} = \sqrt{4 + 12} = \sqrt{16} = 4 \\
\arg(z) &= \pi - \arctan\left(\frac{2\sqrt{3}}{2}\right) = \pi - \frac{\pi}{3} = \frac{2\pi}{3} \\
z &= 4e^{i2\pi/3} = 4(\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}) = 4(-\frac{1}{2} + i\frac{\sqrt{3}}{2}) = -2 + 2i\sqrt{3}
\end{aligned}$$

Rotation by multiplication

Multiplying a complex number by $e^{i\phi}$ rotates it counterclockwise by angle ϕ around the origin. Let $z = re^{i\theta}$ and consider multiplication by $e^{i\phi}$:

$$z \cdot e^{i\phi} = re^{i\theta} \cdot e^{i\phi} = re^{i(\theta+\phi)}$$

The number z with argument θ is rotated to a new complex number with argument $\theta+\phi$, while the modulus r remains unchanged.



This geometric interpretation makes complex multiplication particularly useful in applications involving rotations, such as electrical engineering and computer graphics.

3.4 Powers and nth roots

De Moivre's Formula

For any $n \in \mathbb{Z}$ and complex number $z = r(\cos\theta + i\sin\theta) = re^{i\theta}$:

$$z^n = r^n e^{in\theta} = r^n(\cos(n\theta) + i\sin(n\theta))$$

This allows us to calculate the nth root of a complex number (now we restrict to $n \in \mathbb{N}$). Given $w = \rho e^{i\phi}$, we seek z such that $z^n = w$. Writing $z = re^{i\theta}$, we have:

$$r^n e^{in\theta} = \rho e^{i\phi} \quad \Rightarrow \quad r = \sqrt[n]{\rho} \quad \text{and} \quad n\theta = \phi + 2\pi k,\ k \in \mathbb{Z}$$

Since the complex exponential is 2π-periodic, we obtain n distinct roots with different arguments indexed as:

$$\theta_k = \frac{\phi + 2\pi k}{n}, \qquad k = 0, 1, \ldots, n-1.$$

nth roots

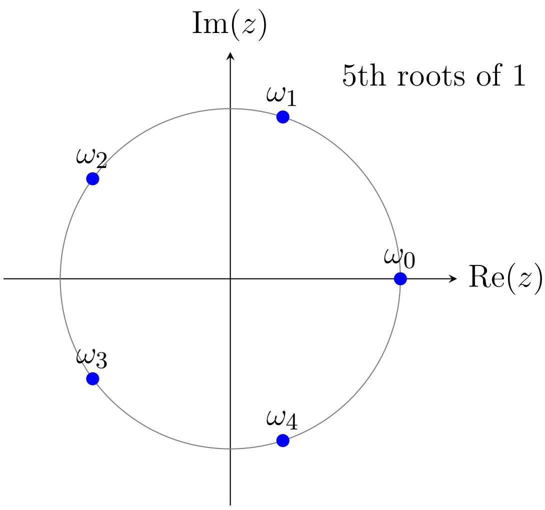
Given $w = \rho e^{i\phi}$, the n distinct nth roots are:

$$\sqrt[n]{w} = \sqrt[n]{\rho}\left(\cos\left(\frac{\phi + 2\pi k}{n}\right) + i\sin\left(\frac{\phi + 2\pi k}{n}\right)\right), \quad k = 0, 1, \ldots, n-1$$

In exponential form:

$$\sqrt[n]{w} = \sqrt[n]{\rho}\, e^{i\frac{\phi + 2\pi k}{n}}, \quad k = 0, 1, \ldots, n-1$$

The roots are equally spaced on a circle of radius $\sqrt[n]{\rho}$ in the complex plane.



Example 3.7: The cube roots of 1 are:

$$\omega_0 = \cos 0 + i \sin 0 = 1$$
$$\omega_1 = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$$
$$\omega_2 = \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$$

Example 3.8: The fifth roots of 1 are:

$$\omega_0 = \cos 0 + i \sin 0 = 1$$
$$\omega_1 = \cos\frac{2\pi}{5} + i\sin\frac{2\pi}{5}$$
$$\omega_2 = \cos\frac{4\pi}{5} + i\sin\frac{4\pi}{5}$$
$$\omega_3 = \cos\frac{6\pi}{5} + i\sin\frac{6\pi}{5}$$
$$\omega_4 = \cos\frac{8\pi}{5} + i\sin\frac{8\pi}{5}$$

Example 3.9: Find all cube roots of $z = 8i$. First write z in exponential form:

$$|z| = 8, \quad \arg(z) = \frac{\pi}{2}$$
$$z = 8e^{i\pi/2}$$

The cube roots are:

$$\omega_0 = \sqrt[3]{8}e^{i(\pi/2)/3} = 2e^{i\pi/6} = 2(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}) = \sqrt{3} + i$$
$$\omega_1 = 2e^{i(\pi/2+2\pi)/3} = 2e^{i5\pi/6} = 2(\cos\frac{5\pi}{6} + i\sin\frac{5\pi}{6}) = -\sqrt{3} + i$$
$$\omega_2 = 2e^{i(\pi/2+4\pi)/3} = 2e^{i3\pi/2} = 2(\cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2}) = -2i$$

Example 3.10: Compute $(1+i)^6$ using De Moivre's theorem:

$$|1+i| = \sqrt{2}, \quad \arg(1+i) = \frac{\pi}{4}$$
$$(1+i)^6 = (\sqrt{2})^6(\cos\frac{6\pi}{4} + i\sin\frac{6\pi}{4}) = 8(\cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2}) = 8(0-i) = -8i$$

Example 3.11: Solve $z^4 + 16 = 0$:

$$z^4 = -16 = 16e^{i\pi}$$
$$z = \sqrt[4]{16}e^{i(\pi+2\pi k)/4} = 2e^{i(\pi+2\pi k)/4}, \quad k = 0, 1, 2, 3$$
$$z_0 = 2e^{i\pi/4} = \sqrt{2} + i\sqrt{2}$$
$$z_1 = 2e^{i3\pi/4} = -\sqrt{2} + i\sqrt{2}$$
$$z_2 = 2e^{i5\pi/4} = -\sqrt{2} - i\sqrt{2}$$
$$z_3 = 2e^{i7\pi/4} = \sqrt{2} - i\sqrt{2}$$

3.5 Algebraic equations and the Fundamental Theorem of Algebra

The study of polynomial equations leads to one of the most important results in mathematics: the Fundamental Theorem of Algebra. This theorem explains why complex numbers are not just a mathematical curiosity, but a necessary extension of the real number system that provides complete solutions to polynomial equations.

Fundamental Theorem of Algebra

Every non-constant polynomial of degree n with complex coefficients has exactly n complex roots (counting multiplicity).

Solving quadratic equations

For a quadratic equation $az^2 + bz + c = 0$ with real coefficients:

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

If the discriminant $\Delta = b^2 - 4ac < 0$, the roots are complex conjugates.

Example 3.12: Solve $z^2 + 4z + 13 = 0$:

$$\begin{aligned} z &= \frac{-4 \pm \sqrt{16 - 52}}{2} = \frac{-4 \pm \sqrt{-36}}{2} \\ &= \frac{-4 \pm 6i}{2} = -2 \pm 3i \end{aligned}$$

The roots are complex conjugates: $-2 + 3i$ and $-2 - 3i$.

Complex conjugate root theorem

If a polynomial with real coefficients has a complex root $a + bi$ ($b \neq 0$), then its complex conjugate $a - bi$ is also a root.

Example 3.13: The polynomial $z^3 - z^2 + z - 1 = 0$ has roots:

$$z = 1, \quad z = i, \quad z = -i.$$

Note that the complex roots i and $-i$ are conjugates.

Summary of Key Formulas

Property	Formula
Modulus	$\lvert z\rvert = \sqrt{a^2 + b^2}$
Argument	$\arg(z) = \arctan(b/a)$
Conjugate	$\bar{z} = a - bi$
Real part	$\mathrm{Re}(z) = \frac{z+\bar{z}}{2}$
Imaginary part	$\mathrm{Im}(z) = \frac{z-\bar{z}}{2i}$
Polar form	$z = r(\cos\theta + i\sin\theta)$
Exponential form	$z = re^{i\theta}$
Multiplication	$z \cdot w = rs\, e^{i(\theta+\phi)}$
Division	$\frac{z}{w} = \frac{r}{s}\, e^{i(\theta-\phi)}$
De Moivre's	$z^n = r^n e^{in\theta}$
nth roots	$\sqrt[n]{z} = \sqrt[n]{r}\, e^{i(\theta+2\pi k)/n}$

Chapter 4

Limits and continuity

The concepts of limits and continuity lie at the very foundation of mathematical analysis. They formalize the intuitive ideas of *approaching* and *connectedness*, and provide the rigorous language needed to define derivatives and integrals.

Historically, these ideas evolved over centuries. Ancient Greek mathematicians like Eudoxus and Archimedes used rudimentary forms of limits in the *method of exhaustion* to compute areas and volumes. However, it was not until the 17th century, with the development of calculus by Newton and Leibniz, that the notion of a limit became central—though it remained informal and often controversial.

In the 19th century, mathematicians such as Cauchy and Weierstrass finally placed limits on a firm logical foundation using the famous ε-δ definitions. The key insight was to quantify the idea of "arbitrarily close" using two positive numbers: ε for the error tolerance in the function's output, and δ for the corresponding distance in the input. This rigorous framework eliminated the paradoxes and ambiguities that had plagued early calculus and remains the standard today.

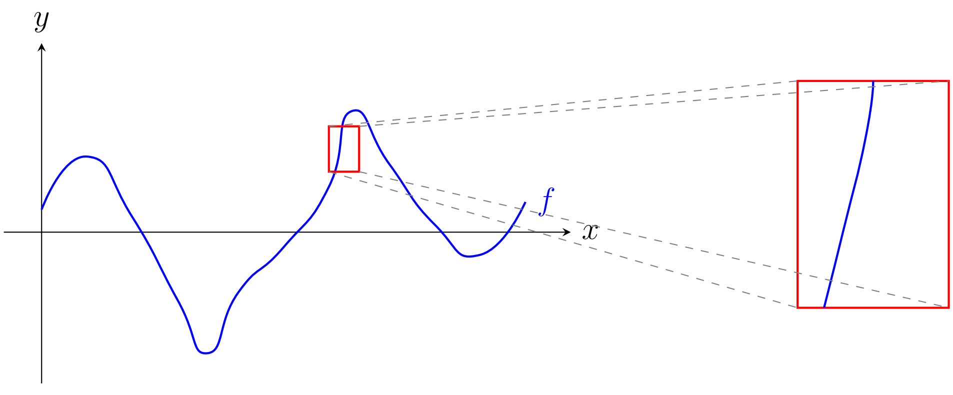


Figure 4.1: We will develop rigorous tools for "zooming-in"

4.1 Neighborhoods

To discuss continuity and limits, we need to be able to "zoom" into the graph of a function. To do this, we need a definition of a small subset around a given point that we want to investigate. This is called a *neighborhood*:

Neighborhood

Let $x_0 \in \mathbb{R}$ be a point on the real line and let $r > 0$ be a positive parameter. The **neighborhood of** x_0 **of radius** r is the *open* and bounded interval

$$I_r(x_0) = (x_0 - r, x_0 + r).$$

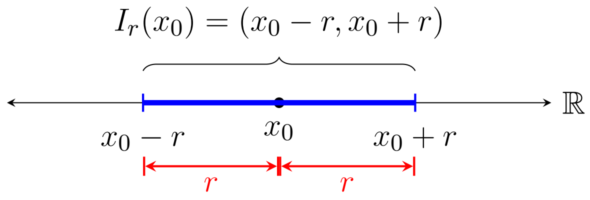


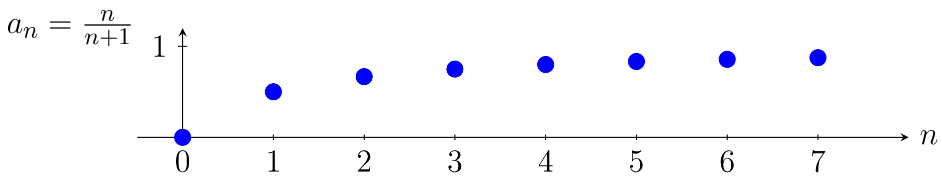
Figure 4.2: Neighborhood of x_0 with radius r

The idea will be that to understand what is happening at x_0, we shall take smaller and smaller neighborhoods of it, so that we can zoom in more and more.

4.2 Limits of sequences

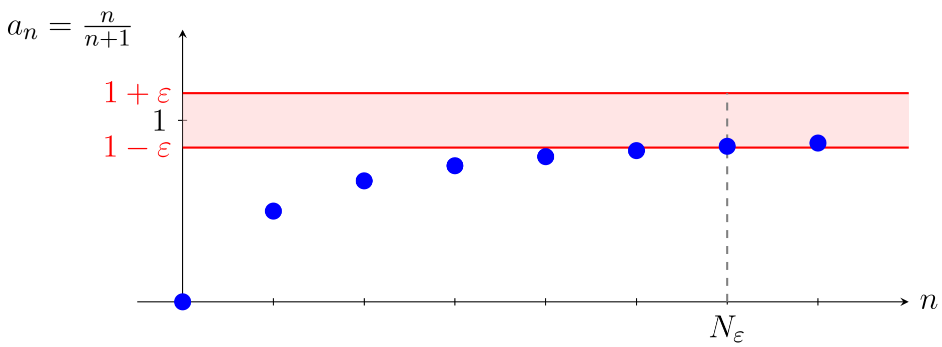
We start by considering sequences, which are functions $f : \mathbb{N} \to \mathbb{R}$. Consider the following two examples which we've seen before:

Example 4.1: Consider the sequence $a_n = \frac{n}{n+1}$.



It is evident that this sequence 'approaches' the number 1. However, we want to make this rigorous. To do this, we look at a small neighborhood of 1 and ask whether the sequence eventually enters this neighborhood (and never leaves!). So our question is this:

> **Q:** *For any* $\varepsilon > 0$*, does there exist* $N_\varepsilon \in \mathbb{N}$ *such that for all* $n > N_\varepsilon$*,* $a_n \in (1 - \varepsilon, 1 + \varepsilon)$*? Or, equivalently,*
>
> $$|1 - a_n| < \varepsilon, \qquad \forall n > N_\varepsilon?$$



So here's the process:

- Fix $\varepsilon > 0$ ***(think of ε as being positive but VERY small!)***
- Compute $1 - a_n$:

$$1 - a_n = 1 - \frac{n}{n+1} = \frac{n+1-n}{n+1} = \frac{1}{n+1}.$$

- Apply the condition $|1 - a_n| < \varepsilon$:

$$\frac{1}{n+1} < \varepsilon \quad \Leftrightarrow \quad 1 < \varepsilon(n+1) \quad \Leftrightarrow \quad \frac{1}{\varepsilon} < n+1$$

- Find N_ε: define

$$N_\varepsilon = \left\lceil \frac{1}{\varepsilon} \right\rceil.$$

- Verify the condition for $\forall n > N_\varepsilon$:

$$\begin{aligned} \forall n > N_\varepsilon, \qquad n+1 > N_\varepsilon + 1 = \left\lceil \frac{1}{\varepsilon} \right\rceil + 1 \geq \frac{1}{\varepsilon} + 1 > \frac{1}{\varepsilon} \\ \Leftrightarrow \qquad \frac{1}{n+1} < \varepsilon \end{aligned}$$

We have therefore answered the previous question affirmatively:

A: *Yes,*

$$\forall \varepsilon > 0, \ \exists N_\varepsilon \in \mathbb{N}, \ \text{s.t.} \ \forall n > N_\varepsilon, \ |1 - a_n| < \varepsilon.$$

In fact, there is an explicit choice for N_ε, namely: $N_\varepsilon = \left\lceil \frac{1}{\varepsilon} \right\rceil$.

Since this is true *for every* $\varepsilon > 0$, the sequence $a_n = \frac{n}{n+1}$ gets arbitrarily close to the number 1. We say that 1 is the *limit* of this sequence.

Finite limit of a sequence

Let $\{a_n\}_{n\in\mathbb{N}}$ be a real-valued sequence. We say that the sequence **converges to** $\ell \in \mathbb{R}$ if for any $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ such that for any $n > N_\varepsilon$, we have $|\ell - a_n| < \varepsilon$. In this case we write

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

The condition for convergence can be written symbolically as:

$$\forall \varepsilon > 0, \ \exists N_\varepsilon \in \mathbb{N}, \ \text{s.t.} \ \forall n > N_\varepsilon, \ |\ell - a_n| < \varepsilon.$$

Returning to Example 4.1, we conclude that:

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

A sequence might also grow indefinitely and 'tend' to $+\infty$ (for example, $a_n = e^n$), or decrease and 'tend' to $-\infty$ (for example, $a_n = -n^8$). This can be formalized as follows:

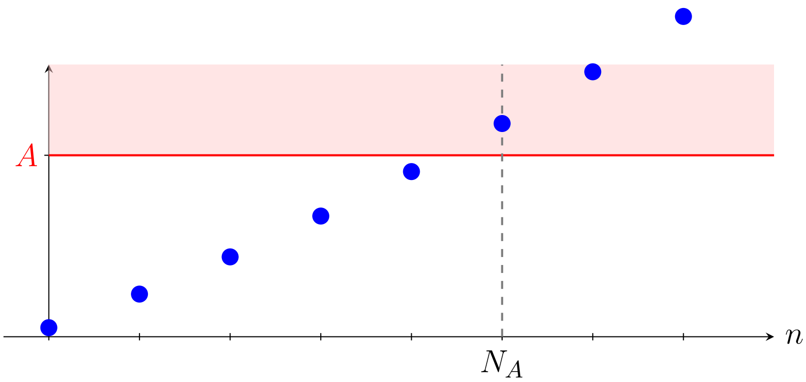
Divergent sequence (to $+\infty$)

Let $\{a_n\}_{n\in\mathbb{N}}$ be a real-valued sequence. We say that the sequence **diverges to** $+\infty$ if for any $A>0$ there exists $N_A\in\mathbb{N}$ such that for any $n>N_A$, we have $a_n>A$. In this case we write

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

The condition for convergence can be written symbolically as:

$$\forall A>0,\ \exists N_A\in\mathbb{N},\ \text{s.t. } \forall n>N_A,\ a_n>A.$$



Here we ***think of*** A ***as being VERY big!***

We can make a similar definition for a sequence divergent to $-\infty$:

Divergent sequence (to $-\infty$)

Let $\{a_n\}_{n\in\mathbb{N}}$ be a real-valued sequence. We say that the sequence **diverges to** $-\infty$ if for any $A<0$ there exists $N_A\in\mathbb{N}$ such that for any $n>N_A$, we have $a_n<A$. In this case we write

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

The condition for convergence can be written symbolically as:

$$\forall A<0,\ \exists N_A\in\mathbb{N},\ \text{s.t. } \forall n>N_A,\ a_n<A.$$

Here we ***think of*** A ***as being VERY big and negative!***

Example 4.2: Consider the sequence $a_n=\frac{n!}{n^{100}}$. Does $\{a_n\}_{n\in\mathbb{N}}$ have a limit as $n\to\infty$? If so, what is it? Does it diverge?

Solution. Let us rewrite the nth term as follows:

$$\begin{aligned}
a_n &= \frac{n!}{n^{100}} \\
&= \frac{n\cdot(n-1)\cdot(n-2)\cdots(n-98)\cdot(n-99)\cdot(n-100)\cdots 3\cdot 2\cdot 1}{n\cdot n\cdots n} \\
&= \frac{n\cdot(n-1)\cdot(n-2)\cdots(n-98)\cdot(n-99)}{n\cdot n\cdots n}\cdot(n-100)\cdots 3\cdot 2\cdot 1.
\end{aligned}$$

Observe that

$$\forall n > 198, \qquad \frac{n-99}{n} > \frac{1}{2}.$$

Hence, for all $n > 198$:

$$\begin{aligned}
a_n &= \frac{n}{n}\cdot\frac{n-1}{n}\cdots\frac{n-98}{n}\cdot\frac{n-99}{n}\cdot(n-100)! \\
&> \frac{1}{2}\cdot\frac{1}{2}\cdots\frac{1}{2}\cdot\frac{1}{2}\cdot(n-100)! \\
&= \frac{1}{2^{100}}\cdot(n-100)!
\end{aligned}$$

Now, $\frac{1}{2^{100}}$ is some (*very* small) positive number, however $(n-100)!$ diverges as $n\to\infty$, so that eventually, for any $A>0$, there exists $N_A\in\mathbb{N}$ such that $\frac{1}{2^{100}}\cdot(n-100)! > A$ for all $n > N_A$. Hence

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

Indeterminate sequences

Some sequences do not converge but also do not diverge. These are called **indeterminate sequences**. Examples include:

- The sequence $a_n = (-1)^n$ is *bounded* but does *not* converge.
- The sequence $a_n = n(-1)^n$ is *not* bounded (not from below and not from above) *and* does *not* converge.

A crucial feature of indeterminate sequences is that their values are not monotone: they increase and decrease repeatedly. The following theorem shows that this is indeed a crucial feature.

Theorem 4.1: Let $\{a_n\}_{n\in\mathbb{N}}$ be a real-valued sequence. Assume that there exists $N\in\mathbb{N}$ such that for all $n > N$, the sequence is monotone. Then the sequence *cannot* be indeterminate. More precisely:

- In the case that the sequence is monotone increasing:
  - ⋆ If $\sup\{a_n \mid n > N\} < +\infty$, then

$$\lim_{n\to\infty} a_n = \sup\{a_n \mid n > N\}.$$

⋆ If $\sup\{a_n \mid n > N\} = +\infty$, then the sequence diverges to $+\infty$.

- In the case that the sequence is monotone decreasing:

  ⋆ If $\inf\{a_n \mid n > N\} > -\infty$, then
  $$\lim_{n\to\infty} a_n = \inf\{a_n \mid n > N\}.$$

  ⋆ If $\inf\{a_n \mid n > N\} = -\infty$, then the sequence diverges to $-\infty$.

Proof. We only prove for the monotone *increasing* case (the decreasing case follows the same proof). For brevity we shall write

$$\sup_{n>N} a_n = \sup\{a_n \mid n > N\}$$

⋆ Suppose that $\sup_{n>N} a_n = \ell < +\infty$. Fix some $\varepsilon > 0$. By the definition of the supremum,

1. there exists some index $N_\varepsilon > N$ such that $\ell - a_{N_\varepsilon} < \varepsilon$;
2. for all $n > N$, $a_n \leq \ell$.

Combining these with the fact that the sequence is monotone increasing for $n > N$, we have the following sequence of inequalities

$$a_{N+1} \leq a_{N+2} \leq \cdots \leq \underbrace{a_{N_\varepsilon}}_{>\ell-\varepsilon} \leq a_{N_\varepsilon+1} \leq \cdots \leq \ell$$

Neglecting the terms up to a_{N_ε}, this can be written as

$$\ell - \varepsilon < a_{N_\varepsilon} \leq a_{N_\varepsilon+1} \leq \cdots \leq \ell$$

This means that for all $n \geq N_\varepsilon$, $|\ell - a_n| < \varepsilon$. By the definition of the limit of a sequence, this means that
$$\lim_{n\to\infty} a_n = \ell.$$

⋆ Now, suppose that $\sup_{n>N} a_n = +\infty$. Then (by definition) for every $A > 0$, there exists $N_A > N$ such that $a_{N_A} > A$. So we have

$$A < a_{N_A} \leq a_{N_A+1} \leq \cdots$$

By definition, this precisely means that

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

□

Example 4.3:

1. The sequence $a_n = \frac{n}{n+1}$ $(n \in \mathbb{N})$ is monotonically increasing, and its supremum is 1. Hence its limit exists (and it is 1).
2. The sequence $a_n = \frac{1}{n}$ $(n \in \mathbb{N}_+)$ is monotonically decreasing and its infimum is 0. Hence its limit exists, and it is also 0.

Proposition 4.2 (The number e): The sequence $a_n = \left(1 + \frac{1}{n}\right)^n$, $n \in \mathbb{N}$, is monotonically increasing and bounded from above. Hence it has a limit, which is denoted e (this is the famous Euler's number, and this is how it is defined):

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

Proof. We need to show that the sequence a_n is bounded and monotonically increasing.

The sequence is monotonically increasing. Actually, it is strictly increasing. We write

$$\begin{aligned} a_n &= \left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\frac{1}{n^k} = \sum_{k=0}^{n} \frac{1}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k} \\ &= \sum_{k=0}^{n} \frac{1}{k!}\frac{n}{n}\cdot\frac{n-1}{n}\cdots\frac{n-k+1}{n} \\ &= \sum_{k=0}^{n} \frac{1}{k!}1\cdot\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right). \end{aligned}$$

Similarly, we can express a_{n+1} as:

$$a_{n+1} = \sum_{k=0}^{n+1} \frac{1}{k!}1\cdot\left(1-\frac{1}{n+1}\right)\cdots\left(1-\frac{k-1}{n+1}\right).$$

Comparing a_n and a_{n+1} we see that:

$$\begin{array}{rl} a_n = \displaystyle\sum_{k=0}^{n} \frac{1}{k!}1\cdot & \left(1-\dfrac{1}{n}\right)\cdots\cdots\left(1-\dfrac{k-1}{n}\right) \\ & \quad\wedge \qquad\qquad\quad \wedge \\ a_{n+1} = \displaystyle\sum_{k=0}^{n+1} \frac{1}{k!}1\cdot & \left(1-\dfrac{1}{n+1}\right)\cdots\left(1-\dfrac{k-1}{n+1}\right) \end{array}$$

each term in the product is bigger in the expression for a_{n+1} (and, moreover, a_{n+1} has an additional positive summand $k = n+1$). Therefore $a_{n+1} > a_n$ (strict inequality).

The sequence is bounded. Observe that $a_1 = 2$, so that 2 is a lower bound (the sequence is increasing). We will now show that 3 is an upper bound. We shall use the inequality

$$k! = \underbrace{k(k-1)(k-2)\cdots 2}_{k-1 \text{ terms}}\cdot 1 \geq \underbrace{2\cdot 2\cdots 2}_{k-1 \text{ times}} = 2^{k-1}.$$

We write a_n as before:

$$\begin{aligned}
a_n &= \sum_{k=0}^{n} \frac{1}{k!} 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \dots \cdot \left(1 - \frac{k-1}{n}\right) \\
&< \sum_{k=0}^{n} \frac{1}{k!} = 1 + \sum_{k=1}^{n} \frac{1}{k!} \\
&\le 1 + \sum_{k=1}^{n} \frac{1}{2^{k-1}} \\
&= 1 + \sum_{k=0}^{n-1} \frac{1}{2^k}.
\end{aligned}$$

We know the formula for the partial sum of a geometric series:

$$\sum_{k=0}^{n-1} \frac{1}{2^k} = \frac{1 - \frac{1}{2^n}}{1 - \frac{1}{2}} = 2\left(1 - \frac{1}{2^n}\right) < 2.$$

So we find that

$$a_n < 1 + \sum_{k=0}^{n-1} \frac{1}{2^k} < 1 + 2 = 3.$$

□

4.3 Limits of functions

Limits at infinity ($x \to +\infty$)

Our first few definitions are very similar to definitions we've already seen for sequences:

Finite limit at infinity (horizontal asymptote)

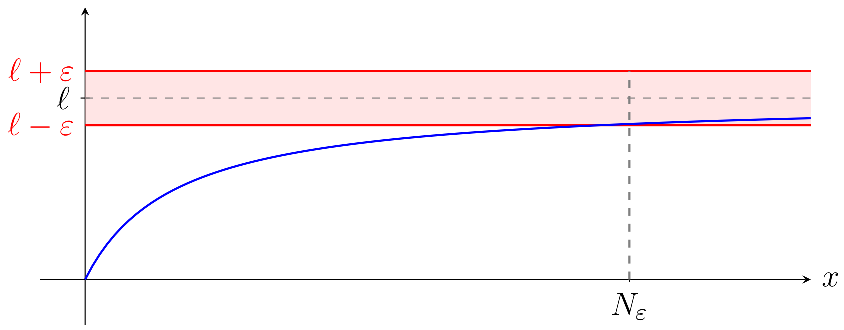
Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. If there exists $\ell \in \mathbb{R}$ such that for any $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{R}$ such that for all $x > N_\varepsilon$, $|\ell - f(x)| < \varepsilon$, we say that f tends to ℓ as $x \to +\infty$, and we write

$$\lim_{x \to +\infty} f(x) = \ell.$$

The condition for convergence can be written symbolically as:

$$\forall \varepsilon > 0,\ \exists N_\varepsilon \in \mathbb{R},\ \text{s.t. } \forall x > N_\varepsilon,\ |\ell - f(x)| < \varepsilon.$$

In this case we say that the line $y = \ell$ is a **right horizontal asymptote** of $f(x)$.



Example 4.4: Let us show that the function $f(x) = \frac{1}{x}$ tends to 0 as $x \to +\infty$.

Fix $\varepsilon > 0$. We need to find $N_\varepsilon \in \mathbb{R}$ such that for all $x > N_\varepsilon$, we have $\left|0 - \frac{1}{x}\right| < \varepsilon$. Note that for $x > 0$, $\left|\frac{1}{x}\right| < \varepsilon$ is equivalent to $x > \frac{1}{\varepsilon}$. Take $N_\varepsilon = \frac{1}{\varepsilon}$. Then for any $x > N_\varepsilon$, we have:

$$\left|0 - \frac{1}{x}\right| = \left|\frac{1}{x}\right| = \frac{1}{x} < \frac{1}{N_\varepsilon} = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this shows that for every $\varepsilon > 0$, there exists N_ε such that for all $x > N_\varepsilon$, $\left|0 - \frac{1}{x}\right| < \varepsilon$. Therefore,

$$\lim_{x \to +\infty} \frac{1}{x} = 0.$$

Example 4.5: Let us show that the function $f(x) = \frac{2x^2+3x-1}{x^2+1}$ tends to 2 as $x \to +\infty$.

Fix $\varepsilon > 0$. We need to find $N_\varepsilon \in \mathbb{R}$ such that for all $x > N_\varepsilon$, we have $\left|2 - \frac{2x^2+3x-1}{x^2+1}\right| < \varepsilon$.

First, simplify the expression:

$$\begin{aligned}\left|2 - \frac{2x^2+3x-1}{x^2+1}\right| &= \left|\frac{2(x^2+1) - (2x^2+3x-1)}{x^2+1}\right| \\ &= \left|\frac{2x^2+2-2x^2-3x+1}{x^2+1}\right| \\ &= \left|\frac{-3x+3}{x^2+1}\right| \\ &= \frac{3|x-1|}{x^2+1}.\end{aligned}$$

For $x > 1$, we have $|x-1| = x-1 < x$, so:

$$\frac{3|x-1|}{x^2+1} < \frac{3x}{x^2+1} < \frac{3x}{x^2} = \frac{3}{x}.$$

We want $\frac{3}{x} < \varepsilon$, which is equivalent to $x > \frac{3}{\varepsilon}$.

Take $N_\varepsilon = \max\left\{1, \frac{3}{\varepsilon}\right\}$ *(The* max *here is to ensure that N_ε is at least* 1*, which is a requirement from before)*. Then for any $x > N_\varepsilon$, we have:

$$\left|2 - \frac{2x^2+3x-1}{x^2+1}\right| < \frac{3}{x} < \frac{3}{N_\varepsilon} \le \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this shows that for every $\varepsilon > 0$, there exists N_ε such that for all $x > N_\varepsilon$, $|2 - f(x)| < \varepsilon$. Therefore,

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

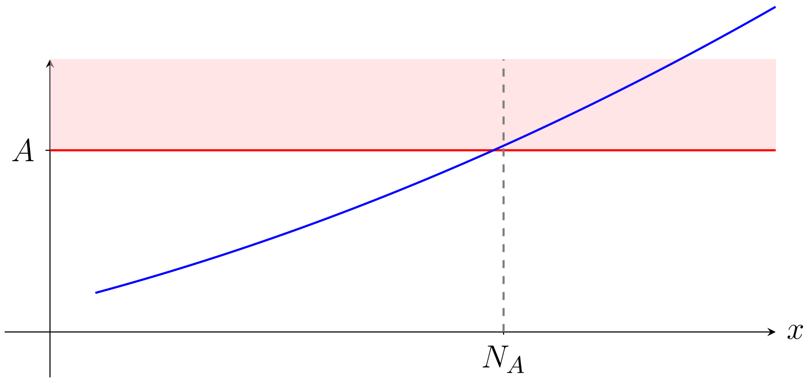
Positive infinite limit at infinity

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. If for any $A > 0$ there exists $N_A \in \mathbb{R}$ such that for all $x > N_A$, $f(x) > A$, we say that f tends to $+\infty$ as $x \to +\infty$, and we write

$$\lim_{x \to +\infty} f(x) = +\infty.$$

The condition for convergence can be written symbolically as:

$$\forall A > 0,\ \exists N_A \in \mathbb{R},\ \text{s.t. } \forall x > N_A,\ \ f(x) > A.$$



Example 4.6: We show that $f(x) = \ln(x)$ tends to $+\infty$ as $x \to +\infty$.

Fix $A > 0$. We need to find $N_A \in \mathbb{R}$ such that for all $x > N_A$, we have $\ln(x) > A$. Take $N_A = e^A$. Then for any $x > e^A$, we have:

$$\ln(x) > \ln(e^A) = A.$$

Since $A > 0$ was arbitrary, this shows that for every $A > 0$, there exists N_A such that for all $x > N_A$, $\ln(x) > A$. Therefore,

$$\lim_{x \to +\infty} \ln(x) = +\infty.$$

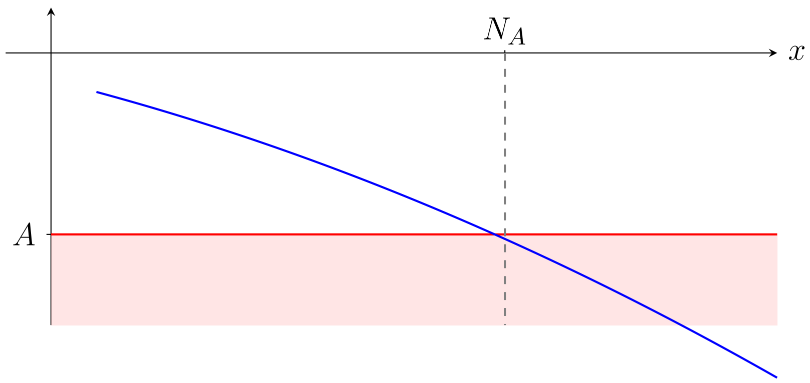
Negative infinite limit at infinity

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. If for any $A < 0$ there exists $N_A \in \mathbb{R}$ such that for all $x > N_A$, $f(x) < A$, we say that f tends to $-\infty$ as $x \to +\infty$, and we write

$$\lim_{x \to +\infty} f(x) = -\infty.$$

The condition for convergence can be written symbolically as:

$$\forall A < 0,\ \exists N_A \in \mathbb{R},\ \text{s.t. } \forall x > N_A,\ \ f(x) < A.$$



Limits at negative infinity ($x \to -\infty$)

The previous definitions can all be modified to consider limits of functions as x tends to $-\infty$. We omit these here, but show an example:

Example 4.7: The function $f(x) = \frac{3x+2}{x-1}$ tends to 3 as $x \to -\infty$.

Fix $\varepsilon > 0$. We need to find $N_\varepsilon \in \mathbb{R}$ such that for all $x < N_\varepsilon$, we have $\left|3 - \frac{3x+2}{x-1}\right| < \varepsilon$. First, simplify the expression:

$$\begin{aligned}\left|3 - \frac{3x+2}{x-1}\right| &= \left|\frac{3(x-1)-(3x+2)}{x-1}\right| \\ &= \left|\frac{3x-3-3x-2}{x-1}\right| \\ &= \left|\frac{-5}{x-1}\right| \\ &= \frac{5}{|x-1|}.\end{aligned}$$

For $x < 0$, we have $|x-1| = 1 - x > -x > 0$, so:

$$\frac{5}{|x-1|} < \frac{5}{-x} = -\frac{5}{x}.$$

We want $-\frac{5}{x} < \varepsilon$, which for $x < 0$ is equivalent to $-x > \frac{5}{\varepsilon}$, or $x < -\frac{5}{\varepsilon}$. Take $N_\varepsilon = -\frac{5}{\varepsilon}$. Then for any $x < N_\varepsilon$, we have:

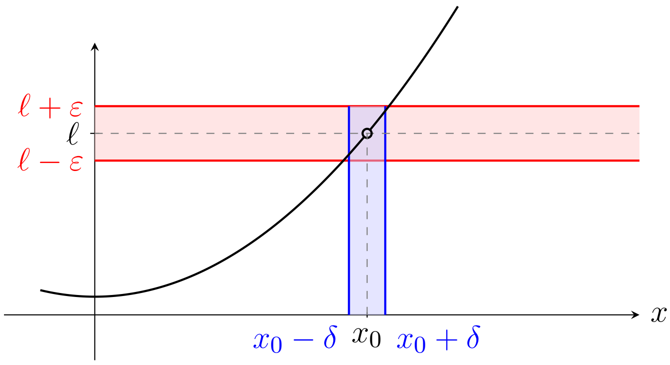
$$\left|3 - \frac{3x+2}{x-1}\right| < -\frac{5}{x} < -\frac{5}{N_\varepsilon} = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this shows that for every $\varepsilon > 0$, there exists N_ε such that for all $x < N_\varepsilon$, $|3 - f(x)| < \varepsilon$. Therefore,

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

Finite limits and continuity

When we want to study the properties of a function $f : \mathbb{R} \to \mathbb{R}$ at a given point $x_0 \in \mathbb{R}$, we want to understand how it behaves for other points $x \neq x_0$ that are close to x_0. Our method is similar to what we've seen before. We call $\ell \in \mathbb{R}$ the 'suspected' value of f at x_0, and consider an ε-neighborhood of ℓ. We then ask whether a small neighborhood of x_0 is within the pre-image of this neighborhood. We note that it isn't necessarily the case that $\ell = f(x_0)$ (we will see several such scenarios). However, if $\ell = f(x_0)$ then the function is continuous at x_0.



$\varepsilon - \delta$ formulation

The so-called $\varepsilon - \delta$ formulation is a staple of modern analysis: $\varepsilon > 0$ measures a small permissible margin of error along the y-axis, and, corresponding to it is $\delta > 0$ (*which depends upon* ε), measuring the corresponding allowable points of input along the x-axis.

Point of continuity

In the case that $\ell = f(x_0)$ and the function approaches $f(x_0)$ for points x near x_0, then we say that the function is continuous at x_0:

Point of continuity

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. We say that f **is continuous at** $x_0 \in \mathbb{R}$ if for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ (depending on ε) such that for all $x \in (x_0 - \delta, x_0 + \delta)$, we have that $f(x) \in (f(x_0) - \varepsilon, f(x_0) + \varepsilon)$.

The condition for continuity can be written symbolically as:

$$\forall \varepsilon > 0,\ \exists \delta = \delta(\varepsilon) > 0,\ \text{s.t. } |x - x_0| < \delta \ \Rightarrow\ |f(x) - f(x_0)| < \varepsilon.$$

In this case we write

$$\lim_{x \to x_0} f(x) = f(x_0).$$

Example 4.8: We prove that $f(x) = x^2$ is continuous at $x_0 = 2$.

Let $\varepsilon > 0$ be given. We want to find $\delta > 0$ such that:

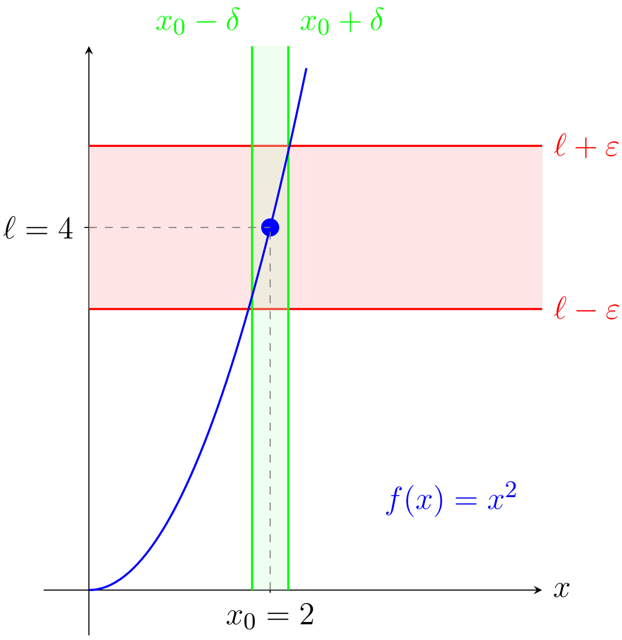
$$|x - 2| < \delta \Rightarrow |x^2 - 4| < \varepsilon$$

Note that $|x^2 - 4| = |x-2||x+2|$. If we restrict $|x-2| < 1$, then $1 < x < 3$, so $|x+2| < 5$.

Choose $\delta = \min\left\{1, \frac{\varepsilon}{5}\right\}$. Then if $|x-2| < \delta$:

$$|x^2 - 4| = |x-2||x+2| < \frac{\varepsilon}{5} \cdot 5 = \varepsilon$$

Thus, $f(x) = x^2$ is continuous at $x_0 = 2$ with $\delta = \min\left\{1, \frac{\varepsilon}{5}\right\}$.



Example 4.9: We prove that $f(x) = \cos x$ is continuous at every point $x_0 \in \mathbb{R}$.

Let $x_0 \in \mathbb{R}$ and $\varepsilon > 0$ be given. We want to find $\delta > 0$ such that:

$$|x - x_0| < \delta \Rightarrow |\cos x - \cos x_0| < \varepsilon$$

Using the trigonometric identity:

$$\cos x - \cos x_0 = -2 \sin\left(\frac{x + x_0}{2}\right) \sin\left(\frac{x - x_0}{2}\right)$$

Taking absolute values and using the fact that $|\sin\theta| \leq |\theta|$ (this is a simple geometric fact):

$$|\cos x - \cos x_0| = 2\left|\sin\left(\frac{x + x_0}{2}\right)\right|\left|\sin\left(\frac{x - x_0}{2}\right)\right| \leq 2 \cdot 1 \cdot \left|\frac{x - x_0}{2}\right| = |x - x_0|$$

Thus, if we choose $\delta = \varepsilon$, then:

$$|x - x_0| < \delta = \varepsilon \Rightarrow |\cos x - \cos x_0| \leq |x - x_0| < \varepsilon$$

Therefore, $f(x) = \cos x$ is continuous at x_0 with $\delta = \varepsilon$.

Continuity on a set

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function, and let $I \subseteq \mathrm{dom}(f)$. If f is continuous at every point $x \in I$, then we say that f **is continuous on the set** I.

Proposition 4.3: All the elementary functions that we've seen in Section 2.6 are continuous on their entire domains.

Proof. We have just shown this for the cosine. The sine function is proved in similar fashion. For the other elementary functions (powers, polynomials, rational functions, the other trigonometric functions and their inverses, exponential and logarithms) we postpone the proof until a later time. □

Removable discontinuity

Another scenario, is either if $f(x_0) \neq \ell$ or if $f(x_0)$ is not defined (i.e. x_0 isn't in the domain of f). In this case we can only say that f has a limit as $x \to x_0$, but that limit does not equal $f(x_0)$ (either because $f(x_0)$ is a different value, or because it is not defined).

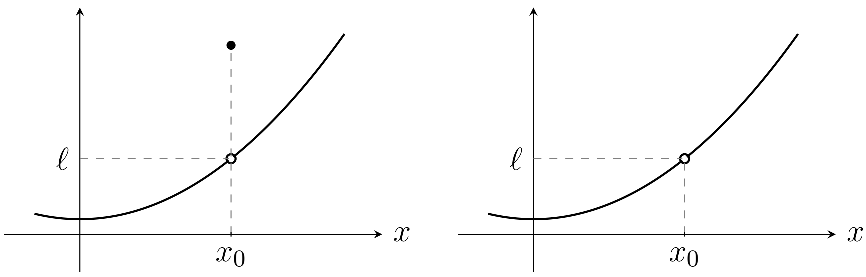


Figure 4.3: Left: $f(x_0) \neq \ell$, and right: f not defined at x_0.

To write the $\varepsilon - \delta$ definition we must exclude the point x_0 itself from consideration:

Removable discontinuity

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. We say that $f(x)$ **tends to** ℓ **as** $x \to x_0$ if for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for all $x \in (x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$, we have that $f(x) \in (\ell - \varepsilon, \ell + \varepsilon)$.

The condition for convergence can be written symbolically as:

$$\forall \varepsilon > 0,\ \exists \delta = \delta(\varepsilon) > 0,\ \text{s.t. } 0 < |x - x_0| < \delta\ \Rightarrow\ |f(x) - \ell| < \varepsilon.$$

In this case we write

$$\lim_{x \to x_0} f(x) = \ell.$$

If $f(x_0) = \ell$, then f is continuous at x_0, as we have previously defined. However, if $f(x_0) \neq \ell$ or if f is not defined at x_0, then we say that f **has a removable discontinuity at** x_0.

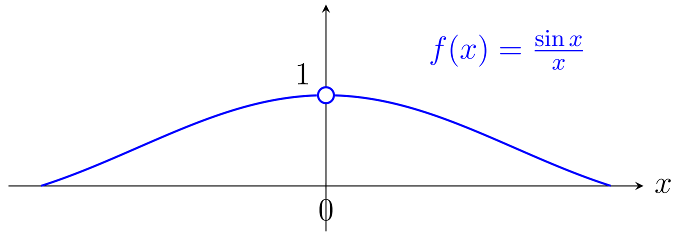
Example 4.10: We will show later that

$$\lim_{x\to 0} \frac{\sin x}{x} = 1$$

even though this function is not defined at $x = 0$. That is, $x = 0$ is a removable discontinuity of the function $\frac{\sin x}{x}$. This means that the function

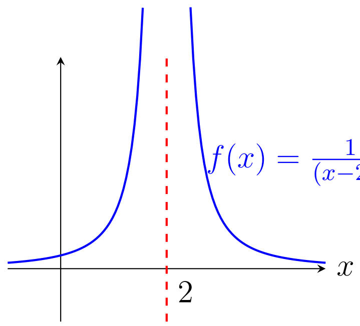
$$f(x) = \begin{cases} \frac{\sin x}{x} & x \neq 0 \\ 1 & x = 0 \end{cases}$$

is a continuous function on $\mathbb{R}$ (we're using here the fact that $\frac{\sin x}{x}$ is continuous at all $x \neq 0$, though we have not proved that). For this reason, when mathematicians normally think about the function $\frac{\sin x}{x}$ they actually think about the function $f(x)$ which is defined at $x = 0$ as well.



Infinite limits (vertical asymptotes)

The limit of a function $f(x)$ as $x \to x_0$ need not be finite. For example, the function $f(x) = \frac{1}{(x-2)^2}$ is not defined at $x = 2$, and as $x \to 2$ the values can become arbitrarily large:



In this case we say that the function has a *vertical asymptote* at $x = 2$.

Positive infinite limit at x_0 (vertical asymptote)

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function defined in a neighborhood of some point $x_0 \in \mathbb{R}$, but possibly not at x_0 itself. We say that f **tends to** $+\infty$ **as** $x \to x_0$ if for every $A > 0$ there exists $\delta = \delta(A) > 0$ such that for all $0 < |x - x_0| < \delta$, we have $f(x) > A$, and we write

$$\lim_{x \to x_0} f(x) = +\infty.$$

The condition for convergence can be written symbolically as:

$$\forall A > 0, \ \exists \delta > 0, \ \text{s.t. } \forall 0 < |x - x_0| < \delta, \ f(x) > A.$$

In this case we say that the line $x = x_0$ is a **vertical asymptote** of $f(x)$.

Let us prove the indeed this is the case for $f(x) = \frac{1}{(x-2)^2}$:

Example 4.11: Show that $f(x) = \frac{1}{(x-2)^2}$ tends to $+\infty$ as $x \to 2$.

We want to show that for every $A > 0$, there exists $\delta > 0$ such that:

$$0 < |x - 2| < \delta \Rightarrow \frac{1}{(x-2)^2} > A.$$

Fix $A > 0$. We need to find $\delta > 0$ such that:

$$\frac{1}{(x-2)^2} > A$$

This inequality is equivalent to:

$$(x-2)^2 < \frac{1}{A}$$

Taking square roots (and noting both sides are positive):

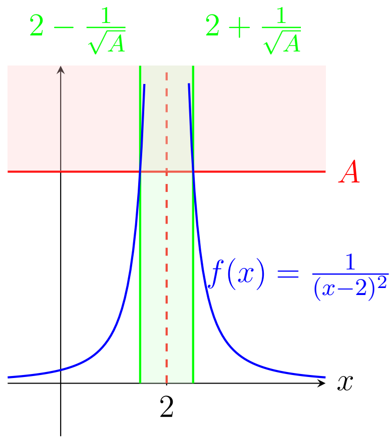
$$|x - 2| < \frac{1}{\sqrt{A}}$$

Therefore, if we choose $\delta = \frac{1}{\sqrt{A}}$, then for $0 < |x - 2| < \delta$:

$$\frac{1}{(x-2)^2} > \frac{1}{\delta^2} = A.$$

Since $A > 0$ was arbitrary, we conclude that

$$\lim_{x \to 2} \frac{1}{(x-2)^2} = +\infty.$$



Negative infinite limit at x_0 (vertical asymptote)

Let $f:\mathbb{R}\to\mathbb{R}$ be a real-valued function defined in a neighborhood of some point $x_0\in\mathbb{R}$, but possibly not at x_0 itself. We say that f **tends to** $-\infty$ **as** $x\to x_0$ if for every $A<0$ there exists $\delta=\delta(A)>0$ such that for all $0<|x-x_0|<\delta$, we have $f(x)<A$, and we write

$$\lim_{x\to x_0} f(x)=-\infty.$$

The condition for convergence can be written symbolically as:

$$\forall A<0,\ \exists\delta>0,\ \text{s.t. } \forall 0<|x-x_0|<\delta,\ f(x)<A.$$

In this case we also say that the line $x=x_0$ is a **vertical asymptote** of $f(x)$.

Note that as defined here, the infinite limit must be the same whether x tends to x_0 from the right or from the left. Hence, the function $f(x)=\frac{1}{x}$ does *not* have a limit as $x\to 0$.

Left and right limits and discontinuity points

We have already seen some examples of functions that have points where we tend to different values if we approach from the left or from the right. Two simple examples at the functions $f(x)=\frac{1}{x}$ and the ceiling function $g(x)=\lceil x\rceil$, sketched below.

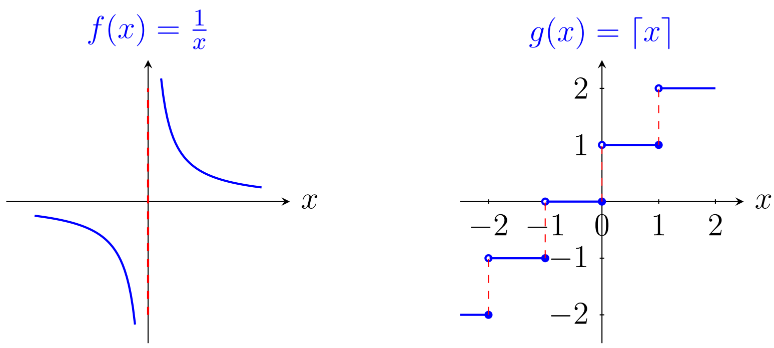


Figure 4.4: Examples of functions that have points with different left and right limits.

Hence we want to repeat the ideas that we've seen above, with the only difference being the neighborhoods around x_0: we'll want a *right* neighborhood and a *left* neighborhood.

Finite right limit

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. We say that f **has a right limit** ℓ **at** $x_0 \in \mathbb{R}$ if for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for all $x \in (x_0, x_0 + \delta)$, we have that $f(x) \in (\ell - \varepsilon, \ell + \varepsilon)$ and we write

$$\lim_{x \to x_0^+} f(x) = \ell.$$

The condition for a right limit can be written symbolically as:

$$\forall \varepsilon > 0,\ \exists \delta = \delta(\varepsilon) > 0,\ \text{s.t. } 0 < x - x_0 < \delta \ \Rightarrow\ |f(x) - \ell| < \varepsilon.$$

Right-continuous

A function $f : \mathbb{R} \to \mathbb{R}$ is said to be **right-continuous at** x_0 if

$$\lim_{x \to x_0^+} f(x) = f(x_0).$$

$\pm\infty$ right limit

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. We say that f **tends to** $\pm\infty$ **from the right at** $x_0 \in \mathbb{R}$ if for every $A \gtrless 0$ there exists $\delta = \delta(A) > 0$ such that for all $x \in (x_0, x_0 + \delta)$, we have that $f(x) \gtrless A$ and we write

$$\lim_{x \to x_0^+} f(x) = \pm\infty.$$

The condition for an infinite right limit can be written symbolically as:

$$\forall A \gtrless 0,\ \exists \delta = \delta(A) > 0,\ \text{s.t. } 0 < x - x_0 < \delta \ \Rightarrow\ f(x) \gtrless A.$$

Finite left limit

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. We say that f **has a left limit** ℓ **at** $x_0 \in \mathbb{R}$ if for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for all $x \in (x_0 - \delta, x_0)$, we have that $f(x) \in (\ell - \varepsilon, \ell + \varepsilon)$ and we write

$$\lim_{x \to x_0^-} f(x) = \ell.$$

The condition for a left limit can be written symbolically as:

$$\forall \varepsilon > 0,\ \exists \delta = \delta(\varepsilon) > 0,\ \text{s.t. } 0 < x_0 - x < \delta \ \Rightarrow\ |f(x) - \ell| < \varepsilon.$$

Left-continuous

A function $f : \mathbb{R} \to \mathbb{R}$ is said to be **left-continuous at** x_0 if

$$\lim_{x \to x_0^-} f(x) = f(x_0).$$

$\pm\infty$ left limit

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. We say that f **tends to** $\pm\infty$ **from the left at** $x_0 \in \mathbb{R}$ if for every $A \gtrless 0$ there exists $\delta = \delta(A) > 0$ such that for all $x \in (x_0 - \delta, x_0)$, we have that $f(x) \gtrless A$ and we write

$$\lim_{x \to x_0^-} f(x) = \pm\infty.$$

The condition for an infinite left limit can be written symbolically as:

$$\forall A \gtrless 0,\ \exists \delta = \delta(A) > 0,\ \text{s.t. } 0 < x_0 - x < \delta \ \Rightarrow\ f(x) \gtrless A.$$

Looking back at Figure 4.4 it appears that the function $f(x) = \frac{1}{x}$ has limits $\pm\infty$ left/right limits, and that the function $g(x) = \lceil x \rceil$ has different left/right limits at all integer points, though it appears that it is always left-continuous.

Jump discontinuity

Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function. If at some point x_0 the left and right limits exist, yet

$$\lim_{x \to x_0^+} f(x) \neq \lim_{x \to x_0^-} f(x)$$

then we say that f has a **jump discontinuity** at x_0.

Conversely, we have the following simple proposition:

Proposition 4.4: Let $f : \mathbb{R} \to \mathbb{R}$ be defined in a neighborhood of x_0 (possibly not at x_0 itself). Then

$$\lim_{x \to x_0} f(x) = L \qquad \Leftrightarrow \qquad \lim_{x \to x_0^+} f(x) = L \quad \text{and} \quad \lim_{x \to x_0^-} f(x) = L$$

where L can be any number or $\pm\infty$. Moreover, the function is continuous at x_0 if and only if it is both right- and left-continuous at x_0.

Proof. Exercise. □

Now we have the tools to consider the removable discontinuity of $f(x) = \frac{\sin x}{x}$ at $x = 0$:

Example 4.12: Consider the function $f(x) = \frac{\sin x}{x}$ for $x \neq 0$. Show that f has a removable discontinuity at $x = 0$ and find the limit.

We want to show that $\lim_{x\to 0} \frac{\sin x}{x}$ exists. We guess that as $x \to 0$, the value of $f(x)$ should tend to 1 (there are good reasons for this particular guess, which we will see later in the course). We want to show that $\lim_{x\to 0} \frac{\sin x}{x} = 1$. Consider the unit circle and let x be a small positive angle. From geometric considerations, for $0 < x < \frac{\pi}{2}$ we have:

$$\sin x < x < \tan x$$

(the first inequality is relatively simple, the second requires a bit more work). Dividing by $\sin x$ (which is positive, so inequalities don't change direction):

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x}.$$

Taking reciprocals (which reverses inequalities):

$$\cos x < \frac{\sin x}{x} < 1.$$

Since $\cos x$ is continuous and $\cos 0 = 1$, the only possible option (we will prove this later, it is called *the Squeeze Theorem*) is that:

$$\lim_{x\to 0^+} \frac{\sin x}{x} = 1$$

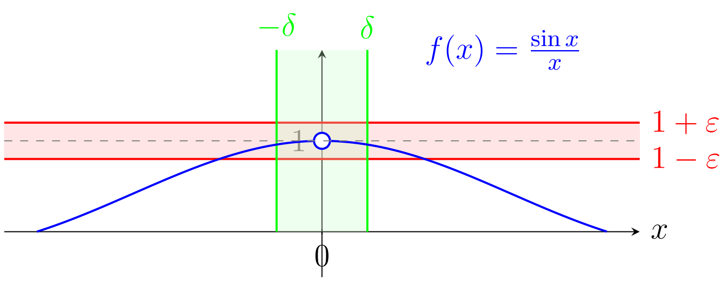
For $x < 0$, let $y = -x > 0$, then:

$$\frac{\sin x}{x} = \frac{\sin(-y)}{-y} = \frac{-\sin y}{-y} = \frac{\sin y}{y}$$

So the left limit equals the right limit:

$$\lim_{x\to 0^-} \frac{\sin x}{x} = \lim_{x\to 0^+} \frac{\sin x}{x} = 1$$

Therefore, using Proposition 4.4, $\lim_{x\to 0} \frac{\sin x}{x} = 1$, and the discontinuity at $x = 0$ is removable.



Finally, some really bad discontinuities!

We have seen removable discontinuities and jump discontinuities. Perhaps it would be wise to define what is a discontinuity is:

Discontinuity

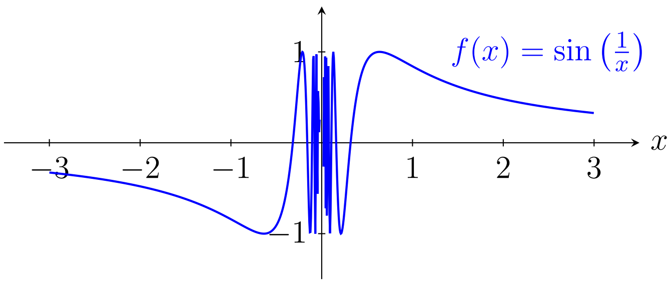
Let $f:\mathbb{R}\to\mathbb{R}$ be a real-valued function. If f is not continuous at x_0 then we say that it is **discontinuous at** x_0 and x_0 is called a **point of discontintinuity**.

There are points of discontinuity that are neither removable nor jump discontinuities:

Consider the function

$$f(x)=\sin\left(\frac{1}{x}\right)$$

whose domain is $\mathbb{R}\setminus\{0\}$. It is discontinuous at $x_0=0$ because the limit does not exist: indeed, as x approaches 0, the argument $\frac{1}{x}$ grows without bound, causing the sine function to oscillate infinitely rapidly between -1 and 1. No matter how small a δ-neighborhood around $x_0=0$ we choose, the function takes all values between -1 and 1 infinitely many times, preventing convergence to any particular limit value.



Discontinuity of the second type

A discontinuity point that is neither removable nor jump, is called a **discontinuity of the second type**.

Limits of monotone functions

The situation is better for monotone functions, just as it was for monotone sequences:

Theorem 4.5: A monotone (increasing or decreasing) function $f:\mathbb{R}\to\mathbb{R}$ cannot have a discontinuity of the second type. That is, a monotone function could only have removable discontinuities, jump discontinuities, or have asymptotes (vertical or horizontal).

Proof. We prove the theorem for a monotone increasing function. The same ideas will carry over for a monotone decreasing function. We split the proof into two claims:

(1) Claim: for any $x_0\in\{-\infty\}\cup\mathbb{R}$,

$$\lim_{x\to x_0^+}f(x)=\inf_{x>x_0}f(x).$$

Let $L_+=\inf_{x>x_0}f(x)$ and suppose that $L_+\in\mathbb{R}$. By the definition of the infimum, for any $\varepsilon>0$, there exists $x_1>x_0$ such that $f(x_1)<L_++\varepsilon$. Since f is monotone increasing, for all $x\in(x_0,x_1)$, we have $L_+\le f(x)\le f(x_1)<L_++\varepsilon$. Thus, $|f(x)-L_+|<\varepsilon$

whenever $0 < x - x_0 < x_1 - x_0$, proving the right-hand limit exists and equals L_+. A similar idea proves the claim for $L_+ = -\infty$.

(2) Claim: for any $x_0 \in \mathbb{R} \cup \{+\infty\}$,

$$\lim_{x \to x_0^-} f(x) = \sup_{x < x_0} f(x).$$

Let $L_- = \sup_{x<x_0} f(x)$. By the definition of the supremum, for any $\varepsilon > 0$, there exists $x_1 < x_0$ such that $f(x_1) > L_- - \varepsilon$. Since f is increasing, for all $x \in (x_1, x_0)$, we have $L_- - \varepsilon < f(x_1) \leq f(x) \leq M$. Thus, $|f(x) - L_-| < \varepsilon$ whenever $0 < x_0 - x < x_0 - x_1$, proving the left-hand limit exists and equals L_-. A similar idea proves the claim for $L_- = +\infty$.

Hence, at any point $x_0 \in \mathbb{R}$, both one-sided limits exist (though they may be infinite). The only possible discontinuities are:

- **Removable discontinuity**: when $L_- = \lim_{x \to x_0^-} f(x) = \lim_{x \to x_0^+} f(x) = L_+$.
- **Jump discontinuity**: when $L_- = \lim_{x \to x_0^-} f(x) < \lim_{x \to x_0^+} f(x) = L_+$
- **Vertical asymptote**: when one of the one-sided limits is infinite (then the other one will not exist because of monotonicity): $L_+ = -\infty$ or $L_- = +\infty$.

A discontinuity of the second type cannot occur. □

Corollary 4.6: Let $f : \mathbb{R} \to \mathbb{R}$ be monotone increasing. Then for any $x_0 \in \mathbb{R}$, if f is defined in a neighborhood of x_0 (but not necessarily at x_0),

$$\lim_{x \to x_0^-} f(x) \leq \lim_{x \to x_0^+} f(x)$$

If f is defined at x_0, then

$$\lim_{x \to x_0^-} f(x) \leq f(x_0) \leq \lim_{x \to x_0^+} f(x).$$

An analogous statement holds for a monotone decreasing function.

Proof. This is an immediate consequence of Theorem 4.5. □

Chapter 5

Properties and computation of limits

5.1 Uniqueness of the limit and local sign of a function

Uniqueness

We always write *the* limit, not *a* limit. Implicitly, we say that it is *unique*. This is true, however it requires proof. Here is the formal statement (an analogous statement could be made for sequences):

Theorem 5.1 (Uniqueness of limits): Let $f : \mathbb{R} \to \mathbb{R}$ and let $x_0 \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$. Suppose that $\lim_{x\to x_0} f(x) = \ell$, where ℓ could be finite of infinite. Then there can be no limit other than ℓ as $x \to x_0$.

Proof. Exercise. *Hint: by contradiction.* □

Local sign

It is intuitively clear that if a function has a positive limit (or $+\infty$), then as we approach this limit the values of the function must also be positive. Analogously, if a limit is negative (or $-\infty$), then the values nearby should be negative. This is stated as follows:

Theorem 5.2 (Local sign): Let $f : \mathbb{R} \to \mathbb{R}$ and let $x_0 \in \mathbb{R}$.

- If $\lim_{x\to x_0} f(x) > 0$ or $\lim_{x\to x_0} f(x) = +\infty$
  then $f > 0$ on a neighborhood of x_0 (potentially excluding x_0 itself).
- If $\lim_{x\to +\infty} f(x) > 0$ or $\lim_{x\to +\infty} f(x) = +\infty$
  then there exists $M > 0$ s.t. $f > 0$ on $\{x > M\}$.
- If $\lim_{x\to -\infty} f(x) > 0$ or $\lim_{x\to -\infty} f(x) = +\infty$
  then there exists $M < 0$ s.t. $f > 0$ on $\{x < M\}$.

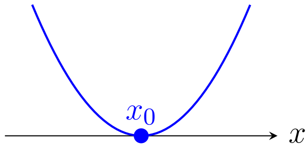
Analogous statements hold if these limits are negative.

Proof. We prove the first claim: $\lim_{x\to x_0} f(x) > 0 \quad \Rightarrow \quad f > 0$ on a neighborhood of x_0. Let $\ell = \lim_{x\to x_0} f(x) > 0$. Let $\varepsilon = \frac{\ell}{2} > 0$. By the definition of the limit, there exists $\delta = \delta(\varepsilon) > 0$ such that for $0 < |x - x_0| < \delta$

$$f(x) \in (\ell - \varepsilon, \ell + \varepsilon) = \left(\ell - \frac{\ell}{2}, \ell + \frac{\ell}{2}\right) = \left(\frac{\ell}{2}, \frac{3\ell}{2}\right) \subset (0, +\infty).$$

Hence $f > 0$ on this neighborhood of x_0 (potentially excluding x_0 itself), which completes the proof. The other claims in the theorem are proved in a similar way. $\square$

The converse of this theorem is *almost* true. As the figure below shows, we can have situations where for all x satisfying $0 < |x - x_0| < \delta$ (for some $\delta > 0$ small), $f(x) > 0$, and yet $f(x_0) = 0$.



Hence we can prove the following statement, which is not quite the converse of the previous theorem:

Theorem 5.3: Let $f : \mathbb{R} \to \mathbb{R}$ and let $x_0 \in \mathbb{R}$. Assume that $\lim_{x \to x_0} f(x)$ exists.

- If $f \geq 0$ on a neighborhood of x_0 then $\lim_{x \to x_0} f(x) \geq 0$ or $\lim_{x \to x_0} f(x) = +\infty$.
- If there exists $M > 0$ s.t. $f \geq 0$ on $\{x > M\}$ then $\lim_{x \to +\infty} f(x) \geq 0$ or $\lim_{x \to +\infty} f(x) = +\infty$
- If there exists $M < 0$ s.t. $f \geq 0$ on $\{x < M\}$ then $\lim_{x \to -\infty} f(x) \geq 0$ or $\lim_{x \to -\infty} f(x) = +\infty$.

Analogous statements hold if these limits are negative.

Proof. We prove the first claim (the others follow a similar strategy). By contradiction, assume that $f \geq 0$ on a neighborhood of x_0 and that $\lim_{x \to x_0} f(x) < 0$ or $\lim_{x \to x_0} f(x) = -\infty$. We immediately obtain a contradiction to Theorem 5.2. $\square$

Theorem 5.4 (Local boundedness): Let $f : \mathbb{R} \to \mathbb{R}$ and let $x_0 \in \mathbb{R}$.

- If $\lim_{x \to x_0} f(x)$ exists and is finite, then f is bounded on a neighborhood of x_0: there exist $\delta > 0$ and $A > 0$ such that for all $0 < |x - x_0| < \delta$, $|f(x)| < A$.
- If $\lim_{x \to +\infty} f(x)$ exists and is finite, then f is bounded for all large x: there exist $A > 0$ and $M > 0$ such that for all $x > M$, $|f(x)| < A$.
- If $\lim_{x \to -\infty} f(x)$ exists and is finite, then f is bounded for all large negative x: there exist $M < 0$ and $A > 0$ such that for all $x < M$, $|f(x)| < A$.

Proof. We prove the first claim. Denote $\ell = \lim_{x \to x_0} f(x) \in \mathbb{R}$. By definition of the limit, for any $\varepsilon > 0$ there exists $\delta > 0$ such that for any $0 < |x - x_0| < \delta$, we have $f(x) \in (\ell - \varepsilon, \ell + \varepsilon)$. Choosing $A = |\ell| + \varepsilon$ will do the job: $|f(x)| < A$ for all $x \in (x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$. $\square$

5.2 Algebra of limits

Given some functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$, our goal is to understand how limits of $f \pm g$ might behave, as well as limits of $f \cdot g$ and $\frac{f}{g}$. In general we should exercise caution!

Example 5.1: Let $C \in \mathbb{R}$. Consider the sequences $a_n = n + C$ and $b_n = -n$. Then:

$$\lim_{n\to\infty} a_n = +\infty \qquad \lim_{n\to\infty} b_n = -\infty$$

and

$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty}(n + C - n) = \lim_{n\to\infty} C = C.$$

This suggests that $+\infty - \infty = C$ for any C. The analogous example for functions would involve $f(x) = x + C$ and $g(x) = -x$.

Example 5.2: Consider the sequences $a_n = 2n$ and $b_n = -n$. Then:

$$\lim_{n\to\infty} a_n = +\infty \qquad \lim_{n\to\infty} b_n = -\infty$$

and

$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty}(2n - n) = \lim_{n\to\infty} n = +\infty.$$

Example 5.3: Consider the sequences $a_n = n$ and $b_n = -2n$. Then:

$$\lim_{n\to\infty} a_n = +\infty \qquad \lim_{n\to\infty} b_n = -\infty$$

and

$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty}(n - 2n) = \lim_{n\to\infty}(-n) = -\infty.$$

These examples demonstrate that we should be cautious when taking limits of sums or differences of sequences and functions. In particular, these examples show that $+\infty - \infty$ is a meaningless expression. Similar absurdities can be achieved by multiplication or division of sequences or functions. Let's list the meaningless expressions:

Meaningless expressions

The following expressions do not make sense:

$$+\infty - \infty \qquad -\infty + \infty \qquad \pm\infty \cdot 0 \qquad \frac{\pm\infty}{\pm\infty} \qquad \frac{0}{0}$$

In contrast, here are the meaningful expressions:

Meaningful expressions

The following expressions *do* make sense:

$$\begin{array}{rcll}
+\infty + C & = & +\infty & (\text{if } C \in \mathbb{R} \cup \{+\infty\}) \\
-\infty + C & = & -\infty & (\text{if } C \in \{-\infty\} \cup \mathbb{R}) \\
\pm\infty \cdot C & = & \pm\infty & (\text{if } C \in \mathbb{R}_+ \cup \{+\infty\}) \\
\pm\infty \cdot C & = & \mp\infty & (\text{if } C \in \{-\infty\} \cup \mathbb{R}_-) \\
\dfrac{\pm\infty}{C} & = & \pm\infty & (\text{if } C \in \mathbb{R}_+) \\
\dfrac{\pm\infty}{C} & = & \mp\infty & (\text{if } C \in \mathbb{R}_-) \\
\dfrac{C}{\pm\infty} & = & 0 & (\text{if } C \in \mathbb{R})
\end{array}$$

A more delicate meaningful expression is

$$\frac{C}{0} = \infty \quad (\text{if } C \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\} \setminus \{0\})$$

However to determine whether it is $+\infty$ or $-\infty$ we need to look at the numerator and the denominator in a neighborhood of the point in question. This could lead to different left and right limits (in the case that the denominator changes sign at the point in question).

Example 5.4: We start with a simple example:

$$f(x) = \frac{1}{x}.$$

At $x = 0$, we have an expression of the form $\frac{C}{0}$. The constant $C = 1 > 0$ in this case, and the denominator changes sign from negative (left of 0) to positive (right of 0). Hence

$$\lim_{x \to 0^-} \frac{1}{x} = -\infty \qquad \text{and} \qquad \lim_{x \to 0^+} \frac{1}{x} = +\infty.$$

Example 5.5: We analyze the rational function

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

Numerator: $x - 2$ is < 0 when $x < 2$, $= 0$ when $x = 2$ and > 0 when $x > 2$.
Denominator: $x^2 - 2x + 1 = (x-1)^2$ is always ≥ 0, and $= 0$ only at $x = 1$.

At $x = 1$ we encounter the situation $\frac{C}{0}$ with $C < 0$. Near $x = 1$, the denominator is > 0, so that both left and right limits must be $-\infty$ and we have:

$$\lim_{x \to 1} \frac{x-2}{x^2 - 2x + 1} = -\infty.$$

The rules we stated allow us to state the following important theorem:

Theorem 5.5 (Algebra of Limits)**:** Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$, and let x_0 either be a real number, or $+\infty$ or $-\infty$. Suppose that as $x \to x_0$, both f and g have limits

$$\lim_{x\to x_0} f(x) = \ell_f \qquad \text{and} \qquad \lim_{x\to x_0} f(x) = \ell_g$$

(these limits can be finite of infinite). Then the following equalities hold ***only*** *when the right hand side is a meaningful expression*:

$$\begin{aligned}\lim_{x\to x_0} (f(x) \pm g(x)) &= \ell_f \pm \ell_g \\ \lim_{x\to x_0} (f(x) \cdot g(x)) &= \ell_f \cdot \ell_g \\ \lim_{x\to x_0} \frac{f(x)}{g(x)} &= \frac{\ell_f}{\ell_g},\end{aligned}$$

where in the last case, $g(x)$ must be nonzero for all x tending to x_0, and it is possible that there are different left and right limits (for instance, if g changes sign at x_0).

Proof. We prove just one case to demonstrate the idea of the proof. Let us prove that

$$\lim_{x\to x_0} (f(x) + g(x)) = \ell_f + \ell_g$$

in the case that both ℓ_f and ℓ_g are real numbers (and not $\pm\infty$). That is, we need to show that for any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for all $0 < |x - x_0| < \delta$, we have that $|(f(x) + g(x)) - (\ell_f + \ell_g)| < \varepsilon$.

Fix $\tilde{\varepsilon} > 0$. We start with f: since $\lim_{x\to x_0} f(x) = \ell_f$,

$$\exists \delta_f = \delta_f(\tilde{\varepsilon}) > 0 \text{ such that } \forall 0 < |x - x_0| < \delta_f,\ |f(x) - \ell_f| < \tilde{\varepsilon}.$$

Similarly, for g, using the same $\tilde{\varepsilon}$:

$$\exists \delta_g = \delta_g(\tilde{\varepsilon}) > 0 \text{ such that } \forall 0 < |x - x_0| < \delta_g,\ |g(x) - \ell_g| < \tilde{\varepsilon}.$$

Define $\delta = \min\{\delta_f, \delta_g\}$. Then for all $0 < |x - x_0| < \delta$,

$$\begin{aligned}|(f(x) + g(x)) - (\ell_f + \ell_g)| &= |(f(x) - \ell_f) + (g(x) - \ell_g)| \\ \textit{(triangle inequality)} \qquad &\le |f(x) - \ell_f| + |g(x) - \ell_g| \\ &< \tilde{\varepsilon} + \tilde{\varepsilon} = 2\tilde{\varepsilon}.\end{aligned}$$

By defining $\varepsilon = 2\tilde{\varepsilon}$ we are done: we have shown that for every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $0 < |x - x_0| < \delta$, we have $|\underbrace{(f(x) + g(x))}_{\text{this is the function}} - \underbrace{(\ell_f + \ell_g)}_{\text{this is the limit}}| < \varepsilon$. ☐

An immediate corollary is this:

Corollary 5.6: If f and g are continuous at $x_0 \in \mathbb{R}$, then so are $f \pm g$, $f \cdot g$ and $\frac{f}{g}$ continuous at x_0 (the last one only if $g(x_0) \neq 0$).

Another corollary is this:

Corollary 5.7: Rational functions are continuous on their domains and polynomials are continuous on $\mathbb{R}$.

5.3 Comparison theorems

Theorem 5.8: Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ and let $x_0 \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$. Suppose that

$$\lim_{x \to x_0} f(x) = \ell_f \qquad \text{and} \qquad \lim_{x \to x_0} g(x) = \ell_g.$$

If in a neighborhood of x_0 (excluding x_0 itself) we have $f(x) \le g(x)$, then $\ell_f \le \ell_g$.

Proof. We skip the proof here. It can be found in the book. $\square$

Theorem 5.9 (Squeeze Theorem)**:** Let $f, g, h : \mathbb{R} \to \mathbb{R}$ and let $x_0 \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$. If

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} h(x) = \ell \in \mathbb{R} \qquad \text{and} \qquad f \le g \le h \text{ in a neighborhood of } x_0$$

(excluding x_0 itself)

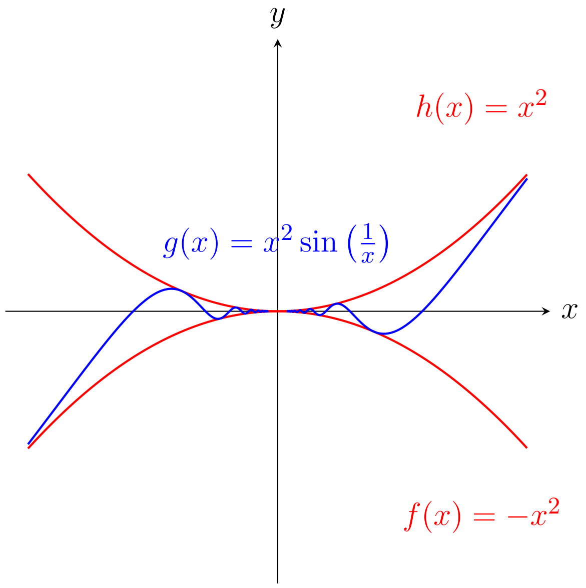
then

$$\lim_{x \to x_0} g(x) = \ell.$$

Proof. We skip the proof here. It can be found in the book. $\square$

Example 5.6: Using this theorem we can immediately conclude that

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



Example 5.7: The Squeeze Theorem justifies our computation in Example 4.12 (the removable singularity of $\frac{\sin x}{x}$ at $x = 0$). In that problem we found that

$$\underbrace{\cos x}_{f(x)} < \underbrace{\frac{\sin x}{x}}_{g(x)} < \underbrace{1}_{h(x)}$$

and used the fact that $\lim_{x \to 0} \cos x = 1$ to conclude that

$$\lim_{x \to 0} \frac{\sin x}{x} = 1.$$

Corollary 5.10 (Squeeze to 0 Theorem): Let $x_0 \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$. Let $f : \mathbb{R} \to \mathbb{R}$ be bounded in a neighborhood of x_0. Let $g : \mathbb{R} \to \mathbb{R}$ satisfy $\lim_{x \to x_0} g(x) = 0$. Then

$$\lim_{x \to x_0} f(x) \cdot g(x) = 0.$$

Proof. Observe that $\lim_{x \to x_0} g(x) = 0$ is satisfied if and only if $\lim_{x \to x_0} |g(x)| = 0$. By assumption, f is bounded on a neighborhood of x_0. This means that there exists $M > 0$ such that $|f(x)| < M$ for all x in this neighborhood. Hence, on this neighborhood of x_0 we have

$$0 \leq |f(x) \cdot g(x)| \leq M g(x)$$

and the claim follows from Theorem 5.9. □

Theorem 5.11 (Squeeze to $\pm\infty$ Theorem): Let $f, g : \mathbb{R} \to \mathbb{R}$ and let $x_0 \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$. If

$$\lim_{x \to x_0} f(x) = +\infty \qquad \text{and} \qquad f \leq g \text{ in a neighborhood of } x_0$$

(excluding x_0 itself)

then

$$\lim_{x \to x_0} g(x) = +\infty.$$

An analogous statement (with obvious modifications) can be made for the case when the limit is $-\infty$.

Proof. The proof is a simple adaptation of previous proofs, we skip it here. □

Example 5.8: Show that

$$\lim_{x \to -\infty} \left(x^2 - e^x - 3 \sin x - 8\right) = +\infty.$$

Observe that for $x < 0$ we have:

$$x^2 - e^x - 3 \sin x - 8 \geq x^2 - 1 - 3 - 8 = x^2 - 12.$$

Letting $f(x) = x^2 - 12$, we see that $\lim_{x \to -\infty} f(x) = +\infty$, and the theorem implies the required result.

5.4 Indeterminate forms of algebraic type

We go back to the meaningless expressions:

Meaningless expressions

$$+\infty - \infty \qquad -\infty + \infty \qquad \pm\infty \cdot 0 \qquad \frac{\pm\infty}{\pm\infty} \qquad \frac{0}{0}$$

Here we want to show that for algebraic functions (i.e. polynomials, rational functions or functions involving roots of polynomials) we can sometimes make sense of such expressions, by careful inspection and simple manipulation.

Simple examples

Example 5.9: The indeterminate form $+\infty - \infty$ can yield any result:

- $\lim_{x\to+\infty} ((x+1) - x) = 1$
- $\lim_{x\to+\infty} (x - (x+1)) = -1$
- $\lim_{x\to+\infty} (2x - x) = +\infty$
- $\lim_{x\to+\infty} (x - 2x) = -\infty$
- $\lim_{x\to+\infty} ((x + \sin x) - x)$ does not exist (oscillates)

Example 5.10: The indeterminate form $+\infty \cdot 0$ can yield any result:

- $\lim_{x\to+\infty} \left(x \cdot \frac{1}{x}\right) = 1$
- $\lim_{x\to+\infty} \left(x \cdot \frac{1}{x^2}\right) = 0$
- $\lim_{x\to+\infty} \left(x^2 \cdot \frac{1}{x}\right) = +\infty$
- $\lim_{x\to+\infty} \left(x \cdot \frac{\sin x}{x}\right)$ does not exist (oscillates)

Example 5.11: The indeterminate form $\frac{+\infty}{+\infty}$ can yield any result:

- $\lim_{x\to+\infty} \left(\frac{x}{x}\right) = 1$
- $\lim_{x\to+\infty} \left(\frac{x}{x^2}\right) = 0$
- $\lim_{x\to+\infty} \left(\frac{x^2}{x}\right) = +\infty$
- $\lim_{x\to+\infty} \left(\frac{x + x\sin x}{x}\right)$ does not exist (oscillates between 0 and 2)

Example 5.12: The indeterminate form $\frac{0}{0}$ can yield any result:

- $\lim_{x\to 0} \left(\frac{x}{x}\right) = 1$
- $\lim_{x\to 0} \left(\frac{x^2}{x}\right) = 0$
- $\lim_{x\to 0} \left(\frac{x}{x^2}\right) = +\infty$
- $\lim_{x\to 0} \left(\frac{\sin x}{x}\right) = 1$
- $\lim_{x\to 0} \left(\frac{x\sin(1/x)}{x}\right)$ does not exist (oscillates)

Polynomials

Consider the following polynomial (assume $a_n \neq 0$):

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

It can be rewritten as

$$p(x) = x^n \underbrace{\left(a_n + \frac{a_{n-1}}{x} + \cdots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}\right)}_{\text{converges to } a_n \text{ as } x \to \pm\infty}.$$

Since we know how the part in the brackets behaves, we can deduce that:

$$\lim_{x\to+\infty} p(x) = \begin{cases} +\infty & \text{if } a_n > 0 \\ -\infty & \text{if } a_n < 0 \end{cases}$$

and

$$\lim_{x\to-\infty} p(x) = \begin{cases} +\infty & \text{if } a_n > 0 \text{ and } n \text{ is even} \\ -\infty & \text{if } a_n > 0 \text{ and } n \text{ is odd} \\ +\infty & \text{if } a_n < 0 \text{ and } n \text{ is odd} \\ -\infty & \text{if } a_n < 0 \text{ and } n \text{ is even} \end{cases}$$

Rational functions

Consider the following rational function (assume $a_n \neq 0$, $b_m \neq 0$):

$$\begin{aligned} r(x) = \frac{p(x)}{q(x)} &= \frac{a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0}{b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0} \\ &= x^{n-m} \underbrace{\frac{a_n + a_{n-1}x^{-1} + \cdots + a_1 x^{-n+1} + a_0 x^{-n}}{b_m + b_{m-1}x^{-1} + \cdots + b_1 x^{-m+1} + b_0 x^{-m}}}_{\text{converges to } \frac{a_n}{b_m} \text{ as } x\to\pm\infty} \end{aligned}$$

We therefore have

$$\lim_{x\to\pm\infty} r(x) = \begin{cases} \infty & \text{if } n > m \\ \frac{a_n}{b_m} & \text{if } n = m \\ 0 & \text{if } n < m \end{cases}$$

The first case ($n > m$) requires further analysis (as in the case of a polynomial) to determine the type of infinity (i.e. whether the limit is $+\infty$ or $-\infty$).

Other algebraic functions

If we encounter a problem with roots, our first goal is to get rid of these roots, at least in the numerator. This can often be achieved by using the fact that $(a+b)(a-b) = a^2 - b^2$:

$$\frac{\sqrt{f(x)} + \sqrt{g(x)}}{h(x)} = \frac{\sqrt{f(x)} + \sqrt{g(x)}}{h(x)} \cdot \frac{\sqrt{f(x)} - \sqrt{g(x)}}{\sqrt{f(x)} - \sqrt{g(x)}} = \frac{f(x) - g(x)}{h(x)(\sqrt{f(x)} - \sqrt{g(x)})}.$$

We have thereby gotten rid of the roots in the numerator, and, hopefully, the resulting expression is easier to deal with.

Examples

Example 5.13: Compute

$$\lim_{x\to 0}\frac{\sqrt{1+5x}-\sqrt{1-2x}}{3x}.$$

Observe that, in the limit, we get an expression of the form $\frac{0}{0}$, so we cannot determine the limit. To proceed, we multiply numerator and denominator by $\sqrt{1+5x}+\sqrt{1-2x}$ to get:

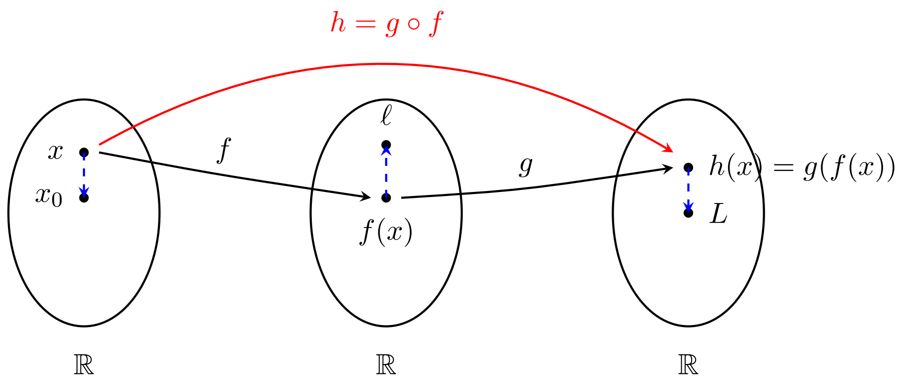
$$\begin{aligned}\frac{\sqrt{1+5x}-\sqrt{1-2x}}{3x} &= \frac{\sqrt{1+5x}-\sqrt{1-2x}}{3x}\cdot\frac{\sqrt{1+5x}+\sqrt{1-2x}}{\sqrt{1+5x}+\sqrt{1-2x}}\\ &= \frac{1+5x-(1-2x)}{3x\left(\sqrt{1+5x}+\sqrt{1-2x}\right)}\\ &= \frac{7x}{3x\left(\sqrt{1+5x}+\sqrt{1-2x}\right)}\\ &= \frac{7}{3}\cdot\underbrace{\frac{1}{\sqrt{1+5x}+\sqrt{1-2x}}}_{\text{this tends to } \frac{1}{2} \text{ as } x\to 0}=\frac{7}{6}.\end{aligned}$$

We have therefore found that:

$$\lim_{x\to 0}\frac{\sqrt{1+5x}-\sqrt{1-2x}}{3x}=\frac{7}{6}.$$

5.5 Substitution Theorem

We now want to understand how limits behave under composition of functions: if we have $f,g:\mathbb{R}\to\mathbb{R}$, with $\lim_{x\to x_0} f(x)=\ell$ and $\lim_{y\to\ell} g(y)=L$, then we want to conclude that $\lim_{x\to x_0} g(f(x))=L$.



This is indeed true:

Theorem 5.12 (Substitution Theorem)**:** Let $f,g:\mathbb{R}\to\mathbb{R}$ and let $x_0,\ell\in\{-\infty\}\cup\mathbb{R}\cup\{+\infty\}$. Suppose that

$$\lim_{x\to x_0} f(x)=\ell$$

and that g is defined on a neighborhood of ℓ (possibly excluding ℓ itself), satisfying:

- if $\ell \in \mathbb{R}$ then g is continuous at ℓ, and
- if $\ell = \pm\infty$ then $\lim_{y\to\ell} g(y)$ exists (possibly infinite).

Then $h = g \circ f$ has a limit as $x \to x_0$ and:

$$\lim_{x\to x_0} g(f(x)) = \lim_{y\to\ell} g(y).$$

Proof. We skip the proof. It is straightforward, and can be found in the book. □

Composition of continuous functions

Observe that if $\ell \in \mathbb{R}$ and g is continuous at ℓ, then $\lim_{y\to\ell} g(y) = g(\ell)$ so that the conclusion of the theorem simplifies to:

$$\lim_{x\to x_0} g(f(x)) = g(\ell) = g(\lim_{x\to x_0} f(x)).$$

That is, in this case, the operation of applying the function g and the operation of taking the limit $x \to x_0$ commute (the order at which we take them can be replaced).

Corollary 5.13: If f is continuous at x_0 and g is continuous at ℓ then $g \circ f$ is continuous at x_0.

Proof. This is immediate from the last comment. Denote $h = g \circ f$. Then:

$$\lim_{x\to x_0} h(x) = \lim_{x\to x_0} g(f(x)) = g(\lim_{x\to x_0} f(x)) = g(f(x_0)) = h(x_0).$$

□

Example 5.14: Compute

$$\lim_{x\to 0} \frac{\sin x^2}{x^2}.$$

We see that this function is the composition of $f(x) = x^2$ with $g(y) = \frac{\sin y}{y}$, for $y \neq 0$. We know that $\lim_{y\to 0} g(y) = 1$, so we complete g by defining $g(0) = 1$. Now f and g are continuous on $\mathbb{R}$. Using the fact that $\lim_{x\to 0} f(x) = 0$ we have

$$\lim_{x\to 0} \frac{\sin x^2}{x^2} = \lim_{y\to 0} \frac{\sin y}{y} = 1.$$

Example 5.15: Compute

$$\lim_{x\to +\infty} \ln\left(\sin\left(\frac{1}{x}\right)\right).$$

Here we have

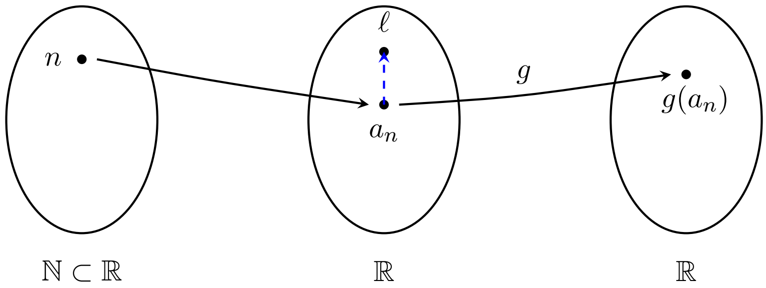
$$f(x) = \sin\left(\frac{1}{x}\right) \qquad \text{and} \qquad g(y) = \ln y.$$

As $x \to +\infty$, $\frac{1}{x} \to 0$. We know that $\sin 0 = 0$. The logarithm $\ln y$ isn't defined for $y \leq 0$, however we know that $\lim_{y\to 0^+} \ln y = -\infty$. So we have:

$$\lim_{x\to +\infty} \ln\left(\sin\left(\frac{1}{x}\right)\right) = \lim_{y\to 0^+} \ln y = -\infty.$$

Nonexistence of a limit

To show that a limit $\lim_{x\to x_0} g(y)$ *doesn't* exist we can rely on the previous results. A common method is as follows: as in the figure below, compose g with a sequence a_n such that $\lim_{n\to\infty} a_n = \ell$. Then try to find another sequence, $\{b_n\}_{n\in\mathbb{N}}$, also satisfying $\lim_{n\to\infty} b_n = \ell$, but for which $\lim_{n\to\infty} g(a_n) \neq \lim_{n\to\infty} g(b_n)$. Then g doesn't have a limit at ℓ.



Theorem 5.14 (Criterion for nonexistence of a limit)**:** Let $g : \mathbb{R} \to \mathbb{R}$ be defined in a neighborhood of $\ell \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$ (possibly excluding ℓ itself). Suppose that there exist sequences $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} a_n = \ell = \lim_{n\to\infty} b_n$ and such that

$$\lim_{n\to\infty} g(a_n) \neq \lim_{n\to\infty} g(b_n).$$

Then $g(y)$ does not have a limit as $y \to \ell$.

Proof. By contradiction. If the limit existed, then the Substitution Theorem would imply that

$$\lim_{n\to\infty} g(a_n) = \lim_{y\to\ell} g(y) = \lim_{n\to\infty} g(b_n),$$

in contradiction to the assumption. □

5.6 Theorems on limits of sequences

We can now continue the analysis of sequences, which we begun in Section 4.2. To simplify the presentation, let us agree that we say that a sequence $\{a_n\}_{n\in\mathbb{N}}$ satisfies a property **for all large** n if there exists $N \in \mathbb{N}$ such that for all $n > N$ the sequence satisfies this property. The results we obtained for functions all carry over to sequences, so we can state the following 'big' theorem:

Theorem 5.15:

1. The limit of a sequence (if exists) is unique.
2. A convergent sequence is bounded.
3. A sequence that is monotone for all large n cannot be indeterminate.
4. For sequences $\{a_n\}_{n\in\mathbb{N}}, \{b_n\}_{n\in\mathbb{N}}$, if $a_n \leq b_n$ for all large n, then $\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n$.
5. For sequences $\{a_n\}_{n\in\mathbb{N}}, \{b_n\}_{n\in\mathbb{N}}, \{c_n\}_{n\in\mathbb{N}}$, if for all large n, $a_n \leq b_n \leq c_n$, and $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n$, then b_n has a limit and it is the same limit.

6. If two sequences $\{a_n\}_{n\in\mathbb{N}}$, $\{b_n\}_{n\in\mathbb{N}}$ have limits $\lim_{n\to\infty} a_n = \ell_a$ and $\lim_{n\to\infty} b_n = \ell_b$, then

   (a) $\lim_{n\to\infty}(a_n \pm b_n) = \ell_a \pm \ell_b$

   (b) $\lim_{n\to\infty}(a_n \cdot b_n) = \ell_a \cdot \ell_b$

   (c) $\lim_{n\to\infty}\left(\frac{a_n}{b_n}\right) = \frac{\ell_a}{\ell_b}$, if for all large n, $b_n \neq 0$,

   whenever the expressions on the right hand side are meaningful.

7. If $\{a_n\}_{n\in\mathbb{N}}$ has the limit ℓ and $g : \mathbb{R} \to \mathbb{R}$ is defined in a neighborhood of ℓ, then

   (a) if $\ell \in \mathbb{R}$ and g is continuous at ℓ, then $\lim_{n\to\infty} g(a_n) = g(\ell)$,

   (b) if $\ell = \pm\infty$ and $\lim_{y\to\ell} g(y)$ exists, then $\lim_{n\to\infty} g(a_n) = \lim_{y\to\ell} g(y)$.

8. $\lim_{n\to\infty} a_n = 0$ if and only if $\lim_{n\to\infty} |a_n| = 0$.

9. If a sequence $\{a_n\}_{n\in\mathbb{N}}$ is bounded, and another sequence $\{b_n\}_{n\in\mathbb{N}}$ satisfies $\lim_{n\to\infty} b_n = 0$, then $\lim_{n\to\infty}(a_n b_n) = 0$.

Proof. The proof is completely analogous to the various proofs we've seen for functions. We skip it here. □

For sequences there is an additional useful tool, which relies in the fact that sequences are discrete (as opposed to functions on $\mathbb{R}$):

Theorem 5.16 (Ratio Test)**:** Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence that for is positive for all large n (i.e. there exists $N \in \mathbb{N}$ such that for all $n > N$, $a_n > 0$). Assume that the limit $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = q$ exists (it may be finite or infinite). Then

- if $q < 1$, then $\lim_{n\to\infty} a_n = 0$,
- if $q > 1$, then $\lim_{n\to\infty} a_n = +\infty$,
- if $q = 1$, it is impossible to determine whether or not the sequence has a limit.

Note that the theorem applies also for sequences that are negative for all large n.

Proof. The proof is simple and we skip it here. □

Example 5.16: Consider a sequence that we have previously seen:

$$a_n = \frac{n!}{n^{100}}.$$

We have

$$\frac{a_{n+1}}{a_n} = \frac{\frac{(n+1)!}{(n+1)^{100}}}{\frac{n!}{n^{100}}} = \frac{n^{100}(n+1)!}{(n+1)^{100}n!} = \underbrace{\left(\frac{n}{n+1}\right)^{100}}_{\to 1} \underbrace{(n+1)}_{\to+\infty},$$

hence

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

so that the sequence a_n diverges.

5.7 Fundamental limits and indeterminate forms of exponential type

We have seen before that
$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e.$$
Now we show the same result for the function $\left(1+\frac{1}{x}\right)^x$:

Claim: The function $\left(1+\frac{1}{x}\right)^x$ has limits as $x\to\pm\infty$, and
$$\lim_{x\to\pm\infty}\left(1+\frac{1}{x}\right)^x = e.$$

Proof. Observe that the function $\left(1+\frac{1}{x}\right)^x$ is defined when $1+\frac{1}{x}>0$ and $x\neq 0$. Hence it is defined when either $x>0$ or $x<-1$. We prove for the case $x\to+\infty$.

Let $n=\lceil x\rceil$. Then
$$n\leq x<n+1.$$
Hence
$$\frac{1}{n+1}<\frac{1}{x}\leq\frac{1}{n}$$
$$\Downarrow$$
$$1+\frac{1}{n+1}<1+\frac{1}{x}\leq 1+\frac{1}{n}$$
$$\Downarrow$$
$$\left(1+\frac{1}{n+1}\right)^n\leq\left(1+\frac{1}{n+1}\right)^x<\left(1+\frac{1}{x}\right)^x\leq\left(1+\frac{1}{n}\right)^x<\left(1+\frac{1}{n}\right)^{n+1}$$

So we have:
$$\underbrace{\left(1+\frac{1}{n+1}\right)^{n+1}}_{\to e}\underbrace{\left(1+\frac{1}{n+1}\right)^{-1}}_{\to 1}<\left(1+\frac{1}{x}\right)^x<\underbrace{\left(1+\frac{1}{n}\right)^n}_{\to e}\underbrace{\left(1+\frac{1}{n}\right)}_{\to 1}$$
So by the Squeeze Theorem,
$$\lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^x = e.$$
The proof in the case $x\to-\infty$ follows similarly, taking caution with signs. □

Observe that by substituting $y=\frac{1}{x}$ we have:
$$\lim_{x\to\pm\infty}\left(1+\frac{1}{x}\right)^x=\lim_{y\to 0}(1+y)^{\frac{1}{y}}=e.$$

Claim:
$$\lim_{x\to 0}\frac{\ln(1+x)}{x}=1.$$

Proof.

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

Hence, using the continuity of the logarithm, we have:

$$\lim_{x\to 0}\frac{\ln(1+x)}{x} = \ln\underbrace{\left(\lim_{x\to 0}(1+x)^{\frac{1}{x}}\right)}_{=e} = 1$$

where in the last equality we have used the previous remark (above) about $(1+x)^{\frac{1}{x}}$. ☐

Claim:

$$\lim_{x\to 0}\frac{e^x-1}{x} = 1.$$

Proof. Follows from the previous claim with an appropriate substitutions, we skip this here. ☐

Useful identities

$$\lim_{x\to 0}\frac{\sin x}{x} = 1$$

$$\lim_{x\to 0}\frac{1-\cos x}{x^2} = \frac{1}{2}$$

$$\lim_{x\to\pm\infty}\left(1+\frac{a}{x}\right)^x = e^a, \qquad (a\in\mathbb{R})$$

$$\lim_{x\to 0}\frac{\log_a(1+x)}{x} = \frac{1}{\ln a}, \qquad (a>0,\ a\neq 1)$$

$$\lim_{x\to 0}\frac{a^x-1}{x} = \ln a, \qquad (a>0)$$

$$\lim_{x\to 0}\frac{(1+x)^\alpha-1}{x} = \alpha, \qquad (\alpha\in\mathbb{R})$$

Power functions

Limits of powers of functions

Let $h(x) = [f(x)]^{g(x)}$, let $x_0 \in \{-\infty\}\cup\mathbb{R}\cup\{+\infty\}$ and assume that f, g have limits as $x\to x_0$ and that $f>0$ near x_0. Observe that $h = e^{g\ln f}$. Hence, by the continuity of the exponential and the fact that for continuous functions we can commute the operations of taking the limit and applying the function:

$$\lim_{x\to x_0}\left([f(x)]^{g(x)}\right) = e^{\lim_{x\to x_0}(g(x)\ln f(x))}.$$

So we need to study the exponential of $\lim_{x\to x_0}(g(x)\ln f(x))$. This is the limit of the product of two functions. We know that it is problematic if we get $0\cdot\infty$. Hence we need to investigate thoroughly in these cases:

1. $\lim_{x\to x_0} g(x) = \pm\infty$ and $\lim_{x\to x_0} f(x) = 1$, so that we get 1^∞.
2. $\lim_{x\to x_0} g(x) = 0$ and $\lim_{x\to x_0} f(x) = 0$, so that we get 0^0.
3. $\lim_{x\to x_0} g(x) = 0$ and $\lim_{x\to x_0} f(x) = +\infty$, so that we get ∞^0.

Example 5.17: Determine $\lim_{x\to+\infty} x^{\frac{1}{x}}$.

We see that this has the form ∞^0. Let $y = \frac{1}{x}$, so that the problem becomes $\lim_{y\to 0^+}(1/y)^y$. We see that

$$x^{\frac{1}{x}} = \left(\frac{1}{y}\right)^y = e^{y\ln\frac{1}{y}} = e^{-y\ln y}.$$

We will later prove that $\lim_{y\to 0^+}(y\ln y) = 0$, so that

$$\lim_{x\to+\infty} x^{\frac{1}{x}} = \lim_{y\to 0^+} e^{-y\ln y} = e^{\lim_{y\to 0^+}(-y\ln y)} = e^0 = 1.$$

Chapter 6

Local comparison of functions

In this chapter we shall gather some tools that will allow us to study the *asymptotic* behavior of functions. The asymptotic behavior of a function $f(x)$ can refer either to its behavior as $x \to \pm\infty$, or as $x \to x_0 \in \mathbb{R}$, where $f(x)$ might tend to 0 or to $\pm\infty$.

6.1 Landau symbols

Let $f, g : \mathbb{R} \to \mathbb{R}$ and let $x_0 \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$. Assume that the limit

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \ell$$

exists (can be finite or infinite). We introduce the **Landau symbols**:

Big $\mathcal{O}$

If ℓ is finite, then we say that f **is controlled by** g as $x \to x_0$, and we write

$$f = \mathcal{O}(g), \quad x \to x_0.$$

We often say that f **is big $\mathcal{O}$ of** g as $x \to x_0$.

Same order

If ℓ is finite and $\ell \neq 0$, then we say that f **has the same order of magnitude as** g as $x \to x_0$, and we write

$$f \asymp g, \quad x \to x_0.$$

Asymptotically equivalent

If $\ell = 1$, then we say that f **is equivalent to** g as $x \to x_0$, and we write

$$f \sim g, \quad x \to x_0.$$

Little o

If $\ell = 0$, then we say that f **is negligible with respect to** g as $x \to x_0$, and we write

$$f = o(g), \quad x \to x_0.$$

We often say that f **is little o of** g as $x \to x_0$.

$\ell = \pm\infty$

If $\ell = \pm\infty$, we need to look at $\frac{g}{f}$ instead!

Example 6.1: 1. Since $\lim_{x\to 0} \frac{\sin x}{x} = 1$ we have

$$\sin x \sim x, \quad x \to 0.$$

2. Since $\lim_{x\to +\infty} \frac{\sin x}{x} = 0$ we have

$$\sin x = o(x), \quad x \to +\infty.$$

3. Since $\lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}$ we have

$$1 - \cos x \asymp x^2, \quad x \to 0.$$

Properties of Landau symbols

1. Observe that f and g having the same order, or being asymptotically equivalent, or being little o, are all subcases of being Big $\mathcal{O}$, that is:

$$\begin{aligned} f \asymp g \quad &\Rightarrow \quad f = \mathcal{O}(g) \\ f \sim g \quad &\Rightarrow \quad f = \mathcal{O}(g) \\ f = o(g) \quad &\Rightarrow \quad f = \mathcal{O}(g) \end{aligned}$$

all as $x \to x_0$. Another implication is:

$$f \sim g \quad \Rightarrow \quad f \asymp g$$

as $x \to x_0$. Conversely:

$$f \asymp g \quad \Rightarrow \quad f \sim \ell g$$

as $x \to x_0$.

2.
$$f \sim g \quad \Leftrightarrow \quad f = g + o(g)$$

as $x \to x_0$.

3.

$$\begin{aligned} f = \mathcal{O}(g) \qquad &\Leftrightarrow \qquad f = \mathcal{O}(\lambda g),\ \forall \lambda \neq 0 \\ f = o(g) \qquad &\Leftrightarrow \qquad f = o(\lambda g),\ \forall \lambda \neq 0 \end{aligned}$$

all as $x \to x_0$.

4.

$$\begin{aligned} f = \mathcal{O}(1),\ x \to x_0 \qquad &\Leftrightarrow \qquad f(x) \xrightarrow{x \to x_0} \ell \in \mathbb{R} \\ f = o(1),\ x \to x_0 \qquad &\Leftrightarrow \qquad f(x) \xrightarrow{x \to x_0} 0 \end{aligned}$$

In particular, in both cases, f is bounded in a neighborhood of x_0.

5.

$$f \text{ is continuous at } x_0 \qquad \Leftrightarrow \qquad f(x) = f(x_0) + o(1),\ \text{as } x \to x_0.$$

Monomials

One of the simplest functions is the monomial power of x: x^n. We therefore want to be able to compare them near where they vanish ($x_0 = 0$) and near where they tend to infinity ($x_0 = \pm\infty$).

Near 0

Near $x_0 = 0$ we have:

$$x^n = o(x^m),\ x \to 0, \qquad \Leftrightarrow \qquad n > m,$$

since $x^{n-m} \to 0$ as $x \to 0$ when $n > m$. This implies that **near 0, bigger powers are negligible**.

Near $\pm\infty$

Near $x_0 = \pm\infty$ we have:

$$x^n = o(x^m),\ x \to \pm\infty, \qquad \Leftrightarrow \qquad n < m,$$

since $x^{n-m} = \frac{1}{x^{m-n}} \to 0$ as $x \to \pm\infty$ when $n < m$. This implies that **near $\pm\infty$, smaller powers are negligible**.

Further properties of Landau symbols

Proposition 6.1: Consider functions $f, \tilde{f}, g, \tilde{g} : \mathbb{R} \to \mathbb{R}$ such that $f \sim \tilde{f}$ and $g \sim \tilde{g}$ as $x \to x_0$. Then

$$\begin{aligned} \lim_{x \to x_0} f(x)g(x) &= \lim_{x \to x_0} \tilde{f}(x)\tilde{g}(x) \\ \lim_{x \to x_0} \frac{f(x)}{g(x)} &= \lim_{x \to x_0} \frac{\tilde{f}(x)}{\tilde{g}(x)} \end{aligned}$$

Proof. Let us prove the first claim. We multiply and divide by the *tilde* functions as follows

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

and the proof is complete. *Note that we have used the fact that we obtained the product of functions that in the limit give meaningful expressions (otherwise we would not have been allowed to take limits of the individual parts of the product).* □

Corollary 6.2: Consider functions $f, f_1, g, g_1 : \mathbb{R} \to \mathbb{R}$ and assume that $f_1 = o(f)$ and $g_1 = o(g)$ both as $x \to x_0$. Then

$$\begin{aligned}\lim_{x\to x_0} (f(x) \pm f_1(x)(g(x) \pm g_1(x)) &= \lim_{x\to x_0} f(x)g(x) \\ \lim_{x\to x_0} \frac{f(x) \pm f_1(x)}{g(x) \pm g_1(x)} &= \lim_{x\to x_0} \frac{f(x)}{g(x)}\end{aligned}$$

Corollary 6.3: Consider functions $f, \tilde{f}, g, \tilde{g} : \mathbb{R} \to \mathbb{R}$ such that $f \sim \tilde{f}$ and $g \sim \tilde{g}$ as $x \to x_0$. Then

$$\begin{aligned} f = \mathcal{O}(g) \quad &\Leftrightarrow \quad f = \mathcal{O}(\tilde{g}) \quad \Leftrightarrow \quad \tilde{f} = \mathcal{O}(g) \quad \Leftrightarrow \quad \tilde{f} = \mathcal{O}(\tilde{g}) \\ f = o(g) \quad &\Leftrightarrow \quad f = o(\tilde{g}) \quad \Leftrightarrow \quad \tilde{f} = o(g) \quad \Leftrightarrow \quad \tilde{f} = o(\tilde{g})\end{aligned}$$

all as $x \to x_0$.

Warning!

These rules do not apply to *sums* and *differences*. For instance, consider

$$\begin{aligned} f(x) &= x & \tilde{f}(x) &= x + 1 \\ g(x) &= x & \tilde{g}(x) &= x\end{aligned}$$

Then $f \sim \tilde{f}$ and $g \sim \tilde{g}$ as $x \to +\infty$. However,

$$\lim_{x\to+\infty} (f(x) - g(x)) = \lim_{x\to+\infty} 0 = 0 \neq 1 = \lim_{x\to+\infty} 1 = \lim_{x\to+\infty} (\tilde{f}(x) - \tilde{g}(x)).$$

Example 6.2: Compute

$$\lim_{x\to 0} \frac{\sin 2x + x^3}{4x + 5\ln(1 + x^2)}.$$

1. First we simplify the numerator. We know that $\lim_{x\to 0} \frac{\sin x}{x} = 1$. This implies that $\sin 2x \sim 2x$ as $x \to 0$. Hence (using Corollary 6.3):

$$x^3 = o(2x) = o(\sin 2x), \quad x \to 0.$$

2. Now we turn to the denominator. We use the fact (which we have not proved yet) that $\ln(1+x) \sim x$ as $x \to 0$. Hence, $\ln(1+x^2) \sim x^2$ as $x \to 0$. Hence:

$$5\ln(1+x^2) \sim 5x^2 = o(4x), \quad x \to 0.$$

3. So we have:

$$\lim_{x\to 0} \frac{\sin 2x + x^3}{4x + 5\ln(1+x^2)} = \lim_{x\to 0} \frac{\sin 2x}{4x} = \frac{1}{2} \underbrace{\lim_{x\to 0} \frac{\sin 2x}{2x}}_{=1} = \frac{1}{2}.$$

Important takeaway

When we want to study the limit of a complicated expression, we need to understand the asymptotic behavior of all the terms that it includes, and try to convert them to monomials.

Fundamental limits

We have the following, all as $x \to 0$: *(some of these will be proven later)*

$$\sin x \sim x$$
$$1 - \cos x \sim \frac{1}{2}x^2$$
$$\ln(1+x) \sim x$$
$$e^x - 1 \sim x$$
$$(1+x)^\alpha - 1 \sim \alpha x$$

6.2 Infinitesimal and infinite functions

As we have seen in the previous section, we are interested in the asymptotic study of the behavior of functions as we either approach their zeros (i.e. points where the function vanishes) or where they 'blow up' (i.e. points where they tend to $\pm\infty$). If a function $f : \mathbb{R} \to \mathbb{R}$ tends to 0 as $x \to x_0$, f is said to be **infinitesimal** at x_0. If it tends to $\pm\infty$ as $x \to x_0$, it is said to be **infinite** at x_0. *Here, as always, x_0 could be any finite value, or $\pm\infty$.*

1) If f and g are two *infinitesimal* functions at x_0, then:

- if $f = o(g)$ at x_0 then f is said to be *infinitesimal of a higher order*. Sometimes we write:

  $$|f| \ll |g| \ll 1$$

  to signify that $f = o(g)$ and $g = o(1)$.

- if $f \sim g$ at x_0 then f and g are said to be *infinitesimal of the same order*.

2) If f and g are two *infinite* functions at x_0, then:

- if $f = o(g)$ at x_0 then f is said to be *infinite of a lower order*. Sometimes we write:
$$1 \ll f \ll g$$
to signify that $f = o(g)$ and $\lim_{x \to x_0} f = +\infty$.
- if $f \sim g$ at x_0 then f and g are said to be *infinite of the same order*.

Ordering of important infinite functions

The following functions are ordered in terms of their infinite order as $x \to +\infty$: *(we will prove this later)*
$$\log_a x \qquad x^s \qquad b^x$$
for any $a > 1,\ s > 0,\ b > 1$.

This means that:
$$\left.\begin{aligned} \log_a x &= o(x^s) \\ x^s &= o(b^x) \end{aligned}\right\} \qquad \text{as } x \to +\infty$$

Ordering of important infinite sequences

The following sequences are ordered in terms of their infinite order as $n \to \infty$: *(we will prove this later)*
$$\log_a n \qquad n^s \qquad b^n \qquad n! \qquad n^n$$
for any $a > 1,\ s > 0,\ b > 1$.

This means that:
$$\left.\begin{aligned} \log_a n &= o(n^s) \\ n^s &= o(b^n) \\ b^n &= o(n!) \\ n! &= o(n^n) \end{aligned}\right\} \qquad \text{as } n \to \infty$$

Stirling formula

There is a precise formula for the relationship between $n!$ and n^n, known as the **Stirling formula**, it provides a way to approximate $n!$:
$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n, \qquad n \to \infty.$$
This is useful in many applications, including in statistics, where factorials appear in the binomial formula.

We don't prove it here. There is something surprising about the fact that $\sqrt{2}$, π and e all appear in this formula....

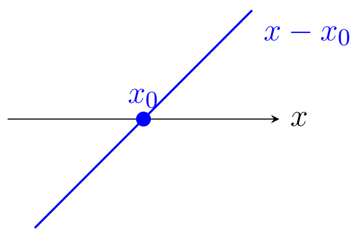
6.3 Order and principal part of infinitesimals and infinites

The most naive thing we can do, when trying to understand the behavior of an infinite or infinitesimal function, is to compare it to powers of x. Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is *infinite* at x_0 if $\lim_{x \to x_0} f(x) = \pm\infty$, and it is *infinitesimal* at x_0 if $\lim_{x \to x_0} f(x) = 0$.

Infinitesimal functions at a finite point x_0

Suppose that $x_0 \in \mathbb{R}$, and that $\lim_{x \to x_0} f(x) = 0$. Then

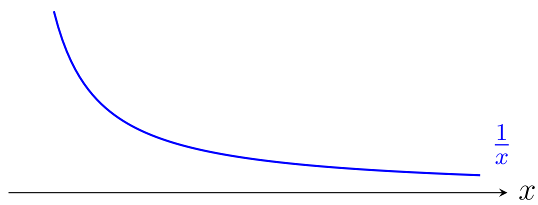
$$\textit{we want to compare } f \textit{ to powers of } \varphi(x) = x - x_0$$



Infinitesimal functions at an infinite point x_0

Suppose that $x_0 \in \{\pm\infty\}$, and that $\lim_{x \to x_0} f(x) = 0$. Then

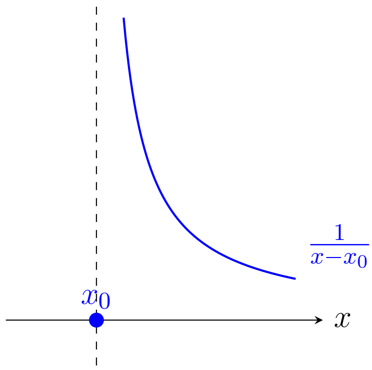
$$\textit{we want to compare } f \textit{ to powers of } \varphi(x) = \frac{1}{x}$$



Infinite functions at a finite point x_0

Suppose that $x_0 \in \mathbb{R}$, and that $\lim_{x \to x_0} f(x) = \pm\infty$. Then

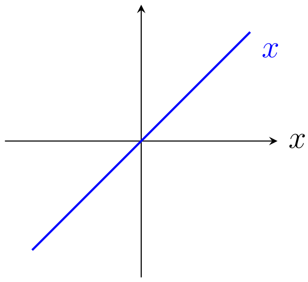
$$\textit{we want to compare } f \textit{ to powers of } \varphi(x) = \frac{1}{x - x_0}$$



Infinite functions at an infinite point x_0

Suppose that $x_0 \in \{\pm\infty\}$, and that $\lim_{x\to x_0} f(x) = \pm\infty$. Then

we want to compare f to powers of $\varphi(x) = x$



In all the above cases, sometimes we may need φ to be non-negative. So we may need to introduce an absolute value in the definitions of the various φ.

Order of an infinitesimal/infinite function

Let f be infinitesimal or infinite at x_0. If there exists $\alpha > 0$ such that

$$f \asymp \varphi^\alpha, \quad x \to x_0$$

then we say that α is the **order** of f at x_0 with respect to φ.

Observe that if f has an order it is *unique*. Furthermore, by definition of the symbol $\asymp$, $f \asymp \varphi^\alpha$, $x \to x_0$, means that

$$\lim_{x\to x_0} \frac{f(x)}{\varphi^\alpha(x)} = \ell \in \mathbb{R} \setminus \{0\}.$$

This can also be written as:

$$f \sim \ell\varphi^\alpha, \quad x \to x_0$$

which can be rewritten as

$$f = \ell\varphi^\alpha + o(\varphi^\alpha), \quad x \to x_0.$$

Principal part of an infinitesimal/infinite function

If $f = \ell\varphi^\alpha + o(\varphi^\alpha)$ as $x \to x_0$ then

$$p(x) = \ell\varphi^\alpha(x)$$

is called the **principal part** of f at x_0 with respect to φ.

Examples

Example 6.3: Consider the function

$$f(x) = \sin x - \tan x, \qquad \text{near } x_0 = 0.$$

Since $\sin 0 = \tan 0 = 0$, f is infinitesimal at $x = 0$. We can write:

$$\begin{aligned}
\sin x - \tan x &= \sin x - \frac{\sin x}{\cos x} \\
&= \frac{\sin x \cdot (\cos x - 1)}{\cos x} \\
&\sim \frac{x \cdot (-\frac{1}{2}x^2)}{1} \\
&= -\frac{1}{2}x^3, \qquad x \to 0.
\end{aligned}$$

Hence f is infinitesimal of order 3 at $x = 0$ with respect to $\varphi(x) = x$. The principal part is $p(x) = -\frac{1}{2}x^3$ and we can write:

$$\sin x - \tan x = -\frac{1}{2}x^3 + o(x^3), \quad x \to 0.$$

Example 6.4: Consider the function

$$f(x) = \sqrt{x^2 + 3} - \sqrt{x^2 - 1}.$$

As $x \to +\infty$, $f(x)$ tends to 0 (verify that you can see why!). So f is infinitesimal at $x_0 = +\infty$. We'll want to compare it to $\varphi(x) = \frac{1}{x}$. First we want the roots in the denominator, so we multiply and divide by $\sqrt{x^2 + 3} + \sqrt{x^2 - 1}$ to get:

$$\begin{aligned}
f(x) &= \sqrt{x^2 + 3} - \sqrt{x^2 - 1} \\
&= \frac{(x^2 + 3) - (x^2 - 1)}{\sqrt{x^2 + 3} + \sqrt{x^2 - 1}} \\
&= \frac{1}{x} \cdot \frac{4}{\sqrt{1 + \frac{3}{x^2}} + \sqrt{1 - \frac{1}{x^2}}}
\end{aligned}$$

To compare to $\varphi(x) = \frac{1}{x}$ we need to look at

$$\lim_{x \to +\infty} \frac{f(x)}{\varphi^\alpha(x)}$$

and try to identify the correct α. We see that we get:

$$\lim_{x\to+\infty} \frac{f(x)}{\varphi^\alpha(x)} = \lim_{x\to+\infty} \frac{x^\alpha}{x} \cdot \frac{4}{\sqrt{1+\frac{3}{x^2}}+\sqrt{1-\frac{1}{x^2}}}$$

so that with $\alpha = 1$ we have

$$\lim_{x\to+\infty} \frac{f(x)}{\varphi^\alpha(x)} = \lim_{x\to+\infty} \frac{4}{\sqrt{1+\frac{3}{x^2}}+\sqrt{1-\frac{1}{x^2}}} = 2.$$

Hence f is infinitesimal of order 1 at $+\infty$ with respect to $\varphi(x) = x^{-1}$. The principal part is $p(x) = 2x^{-1}$ and we can write

$$\sqrt{x^2+3} - \sqrt{x^2-1} = 2x^{-1} + o(x^{-1}), \qquad x \to +\infty.$$

Example 6.5: Consider the function

$$f(x) = \sqrt{9x^5 + 7x^3 - 1}$$

which is infinite as $x \to +\infty$. To determine the order we compare it with $\varphi(x) = x$:

$$\begin{aligned}\lim_{x\to+\infty} \frac{f(x)}{x^\alpha} &= \lim_{x\to+\infty} \frac{\sqrt{9x^5+7x^3-1}}{x^\alpha} \\ &= \lim_{x\to+\infty} \frac{x^{\frac{5}{2}}\sqrt{9+7x^{-2}-x^{-5}}}{x^\alpha}.\end{aligned}$$

This suggests choosing $\alpha = \frac{5}{2}$. Then we have:

$$\lim_{x\to+\infty} \frac{f(x)}{x^{\frac{5}{2}}} = \lim_{x\to+\infty} \sqrt{9+7x^{-2}-x^{-5}} = 3.$$

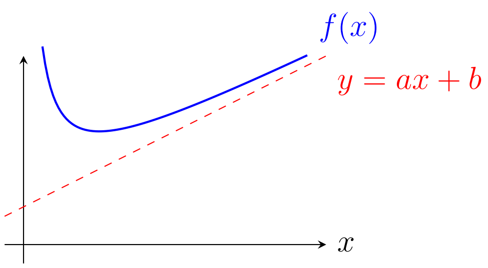
Hence f is infinite of order $\frac{5}{2}$ at $+\infty$ with respect to $\varphi(x) = x$. The principal part is $p(x) = 3x^{\frac{5}{2}}$ and we can write

$$\sqrt{9x^5+7x^3-1} = 3x^{\frac{5}{2}} + o(x^{\frac{5}{2}}), \qquad x \to +\infty.$$

6.4 Asymptotes

We have already seen *horizontal* and *vertical* asymptotes. However it is possible to have slanted asymptotes. We say that a function $f(x)$ **behaves asymptotically as** $x \to +\infty$ like the affine function $y = ax + b$ if

$$\lim_{x\to+\infty} (f(x) - (ax+b)) = 0.$$



A similar definition can be made for $x \to -\infty$.

In principle we can compare a function $f : \mathbb{R} \to \mathbb{R}$ to any function $g : \mathbb{R} \to \mathbb{R}$ as $x \to \pm\infty$. In the case of $+\infty$, we say that f behaves asymptotically as $x \to +\infty$ like g if

$$\lim_{x \to +\infty} (f(x) - g(x)) = 0.$$

Chapter 7

Global properties of continuous functions

As opposed to the previous chapters, where we studied *local* properties of functions (that is, properties of functions as $x \to x_0$, where x_0 can be a finite point or $\pm\infty$), in the current chapter we study *global* properties of functions. That is, we consider functions defined on intervals $[a, b] \subset \mathbb{R}$ (where $a < b$) and try to draw some conclusions on their behavior on this entire interval.

7.1 Theorem of Existence of Zeroes (Bolzano's Theorem)

The first main global result is one for finding zeroes of a function f – i.e. points x_0 where $f(x_0) = 0$:

Theorem 7.1 (Bolzano's Theorem)**:** Let $f : \mathrm{dom}(f) \subseteq \mathbb{R} \to \mathbb{R}$ and suppose it is continuous on an interval $[a, b] \subseteq \mathrm{dom}(f)$ where $a < b$. If f changes sign between a and b then there exists $x_0 \in (a, b)$ such that $f(x_0) = 0$. Moreover, if f is *strictly* monotone on $[a, b]$ then x_0 is the *unique* point in $[a, b]$ where f equals zero.

Such an x_0 is called a **zero of** f.

Proof. Assume that $f(a) < 0 < f(b)$. The proof for the case $f(a) > 0 > f(b)$ will follow the exact same ideas. We split the proof into several steps.

Step A. Constructing sequences $\{a_n\}_{n\in\mathbb{N}}$, $\{b_n\}_{n\in\mathbb{N}}$ in search of x_0:

Step A0. Define $a_0 = a$ and $b_0 = b$. Define the center point between a_0 and b_0 to be $c_0 = \frac{a_0+b_0}{2}$. If $f(c_0) = 0$ then $x_0 = c_0$ and the proof is done. Otherwise, if $f(c_0) < 0$ we define $a_1 = c_0$ and $b_1 = b_0$ and if $f(c_0) > 0$ we define $a_1 = a_0$ and $b_1 = c_0$.

Step A1. Define the center point between a_1 and b_1 to be $c_1 = \frac{a_1+b_1}{2}$. If $f(c_1) = 0$ then $x_0 = c_1$ and the proof is done. Otherwise, if $f(c_1) < 0$ we define $a_2 = c_1$ and $b_2 = b_1$ and if $f(c_1) > 0$ we define $a_2 = a_1$ and $b_2 = c_1$.

Step A2. Define the center point between a_2 and b_2 to be $c_2 = \frac{a_2+b_2}{2}$. If $f(c_2) = 0$ then $x_0 = c_2$ and the proof is done. Otherwise, if $f(c_2) < 0$ we define $a_3 = c_2$ and $b_3 = b_2$ and if $f(c_2) > 0$ we define $a_3 = a_2$ and $b_3 = c_2$.

...and so on...

Step An. The points a_n and b_n have been defined in the previous step. Define the center point between a_n and b_n to be $c_n = \frac{a_n+b_n}{2}$. If $f(c_n) = 0$ then $x_0 = c_n$ and the proof is done. Otherwise, if $f(c_n) < 0$ we define $a_{n+1} = c_n$ and $b_{n+1} = b_n$ and if $f(c_n) > 0$ we define $a_{n+1} = a_n$ and $b_{n+1} = c_n$.

...and so on...

Step B. Properties of the sequences $\{a_n\}_{n\in\mathbb{N}}$, $\{b_n\}_{n\in\mathbb{N}}$.

If a zero x_0 hasn't been located in any finite step, then by induction we have constructed sequences $\{a_n\}_{n\in\mathbb{N}}$, $\{b_n\}_{n\in\mathbb{N}}$ satisfying:

$$\begin{aligned}
&a_0 \le a_1 \le \cdots \le a_n \le \cdots \le b_n \le \cdots b_1 \le b_0 \\
&[a_0, b_0] \supset [a_1, b_1] \supset \cdots \supset [a_n, b_n] \supset \cdots \\
&f(a_n) < 0 < f(b_n) \quad \text{and} \quad b_n - a_n = \frac{a_0 - b_0}{2^n}, \quad \forall n \in \mathbb{N}.
\end{aligned}$$

By the first property, $\{a_n\}_{n\in\mathbb{N}}$ is a bounded, monotonically increasing sequence (it is bounded by b_0, for instance). Hence its limit exists, denote it by $a_\infty \in \mathbb{R}$:

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

Similarly, $\{b_n\}_{n\in\mathbb{N}}$ is bounded and monotonically decreasing, hence its limit exists. We denote it by $b_\infty \in \mathbb{R}$:

$$\lim_{n\to\infty} b_n = b_\infty.$$

Step C. Conclusion.

Observe that

$$b_\infty - a_\infty = \lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = \lim_{n\to\infty} (b_n - a_n) = \lim_{n\to\infty} \frac{b_0 - a_0}{2^n} = 0$$

so that $b_\infty = a_\infty$. Denote this number by x_0:

$$x_0 = b_\infty = a_\infty.$$

Since f is continuous, we may use Theorem 5.15(7a) (the substitution theorem for a continuous function of a convergent sequence) to get:

$$\lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(b_n) = f(x_0).$$

However we know that

$$f(a_n) < 0 < f(b_n), \quad \forall n \in \mathbb{N}.$$

By Theorem 5.15(4),

$$f(x_0) = \lim_{n\to\infty} f(a_n) \le 0 \quad \text{and} \quad 0 \le \lim_{n\to\infty} f(b_n) = f(x_0)$$

Hence it necessarily follows that

$$f(x_0) = 0.$$

Finally, if f is strictly monotone on $[a, b]$, then by Proposition 2.1, f is injective on $[a, b]$. This means that for every $y \in f([a, b])$ there exists a unique $x \in [a, b]$ such that $f(x) = y$. Hence x_0 is the unique zero of f in $[a, b]$. □

Example 7.1: The function

$$f(x) = \begin{cases} 1 & x \geq 10 \\ -1 & x < 10 \end{cases}$$

changes sign on the interval $[9, 11]$. However, there is no zero in $[9, 11]$ (i.e. there is no $x \in [9, 11]$ such that $f(x) = 0$. Why? The problem with this f is that it is not continuous on $[9, 11]$ (there's a jump discontinuity at $x = 10$).

Example 7.2: The function $f(x) = x^2$ is a continuous function on $\mathbb{R}$ that has a zero at $x = 0$, however it is always non-negative (i.e. it does not change sign). This shows that a continuous function can have a zero without changing sign. That is, changing sign is a *sufficient* condition for a continuous function to have a zero, but not a *necessary* condition.

Example 7.3: The function $f(x) = e^x + \sin x$ is continuous on $\mathbb{R}$. Let's look at the interval $[-\frac{\pi}{2}, 0]$. For $x = -\frac{\pi}{2}$ the function is negative: $f(-\frac{\pi}{2}) = e^{-\pi/2} + \sin(-\frac{\pi}{2}) = e^{-\pi/2} - 1 < e^0 - 1 = 1 - 1 = 0$ and for $x = 0$ the function is positive: $f(0) = e^0 - \sin 0 = 1 - 0 = 1 > 0$. Hence there exists $x_0 \in (-\frac{\pi}{2}, 0)$ such that $f(x_0) = 0$. Moreover, e^x is strictly increasing on $\mathbb{R}$, and $\sin x$ is strictly increasing on $[-\frac{\pi}{2}, 0]$, so that x_0 is the unique zero within this interval.

Corollary 7.2: Let $(\alpha, \beta) \subseteq \mathbb{R}$ be an interval (α may be $-\infty$ and β may be $+\infty$), and let $f : \mathbb{R} \to \mathbb{R}$ be defined on (a, b), satisfy

$$\lim_{x \to \alpha^+} f(x) = \ell_\alpha \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$$
$$\lim_{x \to \beta^-} f(x) = \ell_\beta \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$$

with ℓ_α and ℓ_β having opposite signs. Then f has a zero x_0 in (α, β): $f(x_0) = 0$. Moreover, if f is strictly monotone in (α, β) then x_0 is unique.

Proof. This proof is an exercise. □

Corollary 7.3: Let $f, g : \mathbb{R} \to \mathbb{R}$ be continuous on an interval $[a, b]$. If

$$f(a) < g(a) \quad \text{and} \quad f(b) > g(b)$$

or

$$f(a) > g(a) \quad \text{and} \quad f(b) < g(b)$$

then there exists a point $x_0 \in (a, b)$ satisfying

$$f(x_0) = g(x_0).$$

Moreover, if f and g are strictly monotone then x_0 is unique.

Proof. This proof is very simple: we consider the function

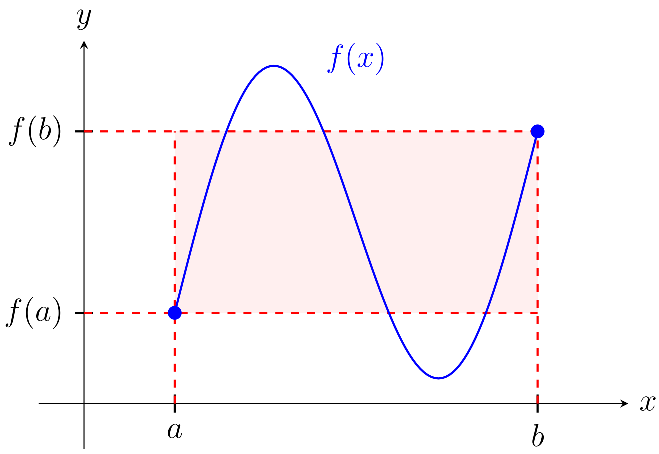
$$h(x) = f(x) - g(x).$$

Then h is continuous on $[a, b]$ (since it is the difference of two continuous functions on $[a, b]$). Furthermore, $h(a)$ and $h(b)$ have different sign (due to the assumptions on f and g). So h satisfies the conditions of Theorem 7.1, and there exists $x_0 \in (a, b)$ such that $h(x_0) = 0$. But this means (by definition of h) that $f(x_0) = g(x_0)$.

The strictly monotone case is also a consequence of Theorem 7.1 (can you think why if f and g are both strictly monotone, then h is strictly monotone? it is not immediately evident. Look at Lemma 2.2). □

7.2 Range of a continuous function defined on an interval

Theorem 7.4 (Intermediate Value Theorem)**:** Let $f : \mathbb{R} \to \mathbb{R}$ be continuous on the closed interval $[a, b]$, where $a < b$. Then f attains all values in the closed interval $[f(a), f(b)]$.



Proof. If $f(a) = f(b)$ the statement of the theorem is trivial: $[f(a), f(b)]$ is a single point, and it has both a and b in its pre-image. So we assume that $f(a) < f(b)$. The case $f(a) > f(b)$ follows similarly, and is left as an exercise.

We need to prove that for any $f(a) < y_0 < f(b)$ there exists $a < x_0 < b$ such that $f(x_0) = y_0$. Consider the constant function $g(x) = y_0$ for all $x \in [a, b]$. Then

$$f(a) < y_0 = g(a) \quad \text{and} \quad g(b) = y_0 < f(b).$$

Thus $f(a) < g(a)$ and $f(b) > g(b)$. By Corollary 7.3 there is a point $x_0 \in (a, b)$ where $f(x_0) = g(x_0)$. By the definition of g, we have that $f(x_0) = y_0$ and the proof is complete. □

Corollary 7.5: Let $f : \mathbb{R} \to \mathbb{R}$ be continuous on an interval $I \subseteq \mathbb{R}$. Then $f(I)$ is also an interval, with endpoints give by $\inf_I f$ and $\sup_I f$.

Remark: The interval I can be open, closed, or half open half closed. Moreover, it may be infinite on one or both ends.

Proof. If $f(I)$ is a single point there's nothing to prove. So assume that there exist $y_1 \neq y_2$ both belonging to $f(I)$, and assume, without loss of generality, that $y_1 < y_2$. Then there exist $x_1 \neq x_2$, both elements of I, such that $f(x_1) = y_1$ and $f(x_2) = y_2$. Assume that $x_1 < x_2$ (the other case will follow a similar proof). Then by the Intermediate Value Theorem (Theorem 7.4), f attains all values $y \in [y_1, y_2]$:

$$[y_1, y_2] \subseteq f([x_1, x_2]) \subseteq f(I).$$

Therefore we have shown that for any two points in $f(I)$, the closed interval that has these points as end-points is a subset of $f(I)$. This implies that $f(I)$ is itself an interval (possibly infinite). This fact is discussed in Lemma 7.6 below. By the definition of *infimum* and *supremum* of a set, it follows that the endpoints of $f(I)$ are $\inf_I f$ and $\sup_I f$ (they can be infinite). □

Lemma 7.6: Let $I \subseteq \mathbb{R}$ be a subset of $\mathbb{R}$ that satisfies the following condition: for any $x_1, x_2 \in I$ with $x_1 < x_2$, the entire closed interval $[x_1, x_2]$ is a subset of I. Then I is an interval, i.e. I could be an open, closed, or half-open-half-closed interval, and one or both of its endpoints can be infinite.

Proof. This is a simple proof (by contradiction) which we skip here. $\square$

Example 7.4: Let $f(x) = \tan x$. Then $f((-\frac{\pi}{2}, \frac{\pi}{2})) = (-\infty, +\infty)$.

Example 7.5: Let $f(x) = \cos x$. Then $f((-\infty, +\infty)) = [-1, 1]$.

Example 7.6: Let $f(x) = \arctan x$. Then $f((-\infty, +\infty)) = (-\frac{\pi}{2}, \frac{\pi}{2})$.

Example 7.7: Let $f(x) = e^x$. Then $f([0, +\infty)) = [1, +\infty)$.

The following theorem tells us that the image of a closed interval under a continuous function is always a closed interval:

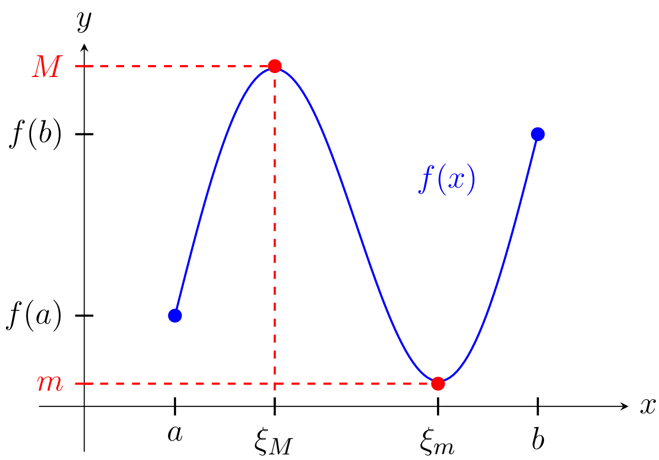
Weierstrass' Theorem

Let f be continuous on the interval $[a, b]$, where $a < b$ are real numbers. Then f is bounded on $[a, b]$ and it attains its minimum and maximum on $[a, b]$:

$$m = \min_{x \in [a,b]} f(x) \qquad \text{and} \qquad M = \max_{x \in [a,b]} f(x).$$

It follows that

$$f([a, b]) = [m, M].$$



Proof. **Step A.** The supremum. Define

$$M = \sup_{x \in [a,b]} f(x)$$

which can be a finite number or $+\infty$.

Case 1: M is finite. In this case, we know by definition of the supremum that for any $\varepsilon > 0$ there exists $x_\varepsilon \in [a, b]$ satisfying $M - \varepsilon < f(x_\varepsilon) \leq M$. Our goal is to construct a sequence $\{a_n\}_{n\in N_+}$, as follows: instead of ε take, for each $n \geq 1$, $\frac{1}{n}$, so that there exists $x_n \in [a,b]$ such that

$$M - \frac{1}{n} < f(x_n) \leq M.$$

By the Squeeze Theorem for sequences (Theorem 5.15(5)) if follows that

$$\lim_{n\to\infty} f(x_n) = M.$$

Case 2: M is infinite. In this case, again by the definition of the supremum, for each $n \in \mathbb{N}_+$ there exists $x_n \in [a,b]$ such that

$$f(x_n) > n.$$

By the Squeeze to $\pm\infty$ Theorem for sequences (Theorem 5.15(4)) we must have

$$\lim_{n\to\infty} f(x_n) = +\infty = M.$$

Step B. The sequence $\{x_n\}_{n\in\mathbb{N}_+}$. In both cases, we obtained a sequence $\{x_n\}_{n\in\mathbb{N}_+} \subset [a,b]$. This is a bounded sequence (it is contained in a bounded interval). By the Bolzano-Weierstrass Theorem (see below) it has a convergent subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$. Call its limit $\xi_M \in [a,b]$. Then we have:

$$\lim_{k\to\infty} x_{n_k} = \xi_M.$$

Moreover, since $f(x_n)$ converges, so does its subsequence $f(x_{n_k})$ and they share the same limit:

$$M = \lim_{n\to\infty} f(x_n) = \lim_{k\to\infty} f(x_{n_k}).$$

Using the continuity of f at the point ξ_M, we have

$$M = \lim_{k\to\infty} f(x_{n_k}) = f(\lim_{k\to\infty} x_{n_k}) = f(\xi_M).$$

But if $M = f(\xi_M)$ it must be a real number, so M cannot be $+\infty$. Furthermore, since M is attained at ξ_M, it belongs to the range f on $[a,b]$, hence

$$M = \max_{x\in[a,b]} f(x).$$

The proof for m follows the same ideas.

Finally, the fact that $f([a,b]) = [m,M]$ is an immediate consequence of Corollary 7.5. $\square$

The Bolzano-Weierstrass Theorem

Let $\{x_n\}_{n\in\mathbb{N}}$ be a bounded sequence (i.e., there exist real numbers $a < b$ such that $a < x_n < b$ for all $n \in \mathbb{N}$). Then $\{x_n\}_{n\in\mathbb{N}}$ has a convergent subsequence: there exists a sequence of indices $\{n_k\}_{k\in\mathbb{N}} \subseteq \mathbb{N}$ such that the subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ converges.

Proof. We do not give the full proof here, just a sketch. The proof follows the same bisection method that we've seen in the proof of Theorem 7.1:

There exist infinitely many points of the sequence in $[a, b]$. Split $[a, b]$ into two halves, left and right. At least one of them will contain infinitely many points of the sequence. Pick n_1 such that x_{n_1} belongs to that half. Divide that half into two. At least one of those two halves contains infinitely many points of the sequence. Pick n_2 such that x_{n_2} belongs to that half. If we keep dividing the interval and picking points, we obtain a subsequence

$$x_{n_1}, x_{n_2}, x_{n_3}, x_{n_4}, \dots, x_{n_k}, \dots$$

where the indices $\{n_k\}_{n\in\mathbb{N}_+} \subseteq \mathbb{N}_+$ form an infinite subset of the integers. By construction, this subsequence will have a limit. □

Example 7.8: The sequence $a_n = (-1)^n$, $n \in \mathbb{N}$, has values

$$1, -1, 1, -1, 1, -1, \dots, (-1)^n, (-1)^{n+1}, \dots$$

This sequence is indeterminate. However, if we choose only the even indices we find that a_{n_k} (where $n_k = 2k$) is the subsequence $a_0, a_2, a_4, \dots = 1, 1, 1, \dots$ does have the trivial limit of 1. Similarly, the odd indices ($n_k = 2k + 1$) will give us a subsequence with the limit -1.

Remark: The above example demonstrates that a bounded sequence can have more than one convergent subsequence. The Bolzano-Weierstrass theorem tells us that there's *at least* one.

7.3 Invertibility of continuous functions

Lemma 7.7: Let f be continuous and invertible on an interval I. Let $x_1 < x_2 < x_3$ be points in I. Then exactly one of the following holds:

$$(i) \quad f(x_1) < f(x_2) < f(x_3)$$

or

$$(ii) \quad f(x_1) > f(x_2) > f(x_3).$$

Proof. Since f is invertible, it is 1-1. Therefore, since $x_1 \neq x_3$, also $f(x_1) \neq f(x_3)$.

Consider the case $f(x_1) < f(x_3)$ and assume *by contradiction* that (i) isn't satisfied. So $f(x_2) \notin (f(x_1), f(x_3))$.

Suppose that $f(x_2)$ lies to the right of the interval $(f(x_1), f(x_3))$, so that $f(x_1) < f(x_3) < f(x_2)$. Consider f on $[x_1, x_2]$. Since it is continuous there, by Theorem 7.4 (The Intermediate Value Theorem), $f|_{[x_1,x_2]}$ must assume all values in $[f(x_1), f(x_2)]$. Since $f(x_1) < f(x_3) < f(x_2)$, there will be $z \in [x_1, x_2]$ such that $f(z) = f(x_3)$. But this contradicts the fact that f is 1-1. Hence (i) is satisfied and $f(x_1) < f(x_2) < f(x_3)$.

The case that $f(x_2)$ lies to the left of the interval $(f(x_1), f(x_3))$ is handled in similar way.

The proof for the case $f(x_1) > f(x_3)$ follows the same ideas, and the lemma follows. □

Theorem 7.8: Let f be continuous on an interval $I \subseteq \mathbb{R}$. Then

$$f \text{ is 1-1 on } I \qquad \Leftrightarrow \qquad f \text{ is strictly monotone on } I.$$

Proof. The direction $\Leftarrow$ was already proven in Proposition 2.1.

Hence we only need to prove the implication $\Rightarrow$. Recall that a 1-1 function is invertible, so that it is enough to prove

$$f \text{ is invertible on } I \qquad \Rightarrow \qquad f \text{ is strictly monotone on } I.$$

Let $x_1 < x_2$ be points in I. We'll show that if $f(x_1) < f(x_2)$ then f is strictly increasing on I. Let z_1, z_2 be two points such that

$$x_1 < z_1 < z_2 < x_2.$$

We want to show that $f(z_1) < f(z_2)$.

Using Lemma 7.7 for the trio $x_1 < z_1 < x_2$, we find that

$$f(x_1) < f(z_1) < f(x_2).$$

Using Lemma 7.7 for the trio $z_1 < z_2 < x_2$, we find that

$$f(z_1) < f(z_2) < f(x_2).$$

In particular, we have found that $z_1 < z_2$ implies $f(z_1) < f(z_2)$, so f is strictly increasing.

The strictly decreasing case $(f(x_1) > f(x_2))$ follows the same idea of proof, which completes the proof. □

Theorem 7.9: Let f be continuous and invertible on an interval I. Then f^{-1} is continuous on the interval $J = f(I)$.

Proof. We skip the proof. □

7.4 Lipschitz and uniformly continuous functions

Lipschitz functions

A function $f : \mathbb{R} \to \mathbb{R}$ is **Lipschitz on an interval** $I \subseteq \mathbb{R}$ if there exists a constant $L \geq 0$ such that

$$|f(x_1) - f(x_2)| \leq L|x_1 - x_2| \qquad \forall x_1, x_2 \in I.$$

For a Lipschitz function on I, the smallest L that satisfies this inequality is called the **Lipschitz constant** of f on I.

Example 7.9: The function $f(x) = x$ is Lipschitz on $\mathbb{R}$ with Lipschitz constant 1: for any $x_1, x_2 \in \mathbb{R}$,

$$|f(x_1) - f(x_2)| = |x_1 - x_2|.$$

Example 7.10: The function $f(x) = x^2$ is Lipschitz on the interval $[0, 2]$ with Lipschitz constant 4. Indeed, for any $x_1, x_2 \in [0, 2]$,

$$|f(x_1) - f(x_2)| = |x_1^2 - x_2^2| = |x_1 - x_2| \cdot \underbrace{|x_1 + x_2|}_{\leq 2+2=4} \leq 4|x_1 - x_2|.$$

Example 7.11: The function $f(x) = \sqrt{x}$ is not Lipschitz on the interval $[0, 1]$. This can be shown by taking $x_1 = 0$ and any other $x_2 \in (0, 1]$:

$$|f(0) - f(x_2)| = |\sqrt{0} - \sqrt{x_2}| = \sqrt{x_2} = \frac{1}{\sqrt{x_2}} x_2 = \frac{1}{\sqrt{x_2}} |0 - x_2|$$

As $x_2 \to 0^+$, the coefficient $\frac{1}{\sqrt{x_2}}$ tends to $+\infty$, so there's no finite L satisfying the conditions of a Lipschitz function.

However, $f(x) = \sqrt{x}$ *is* Lipschitz on any interval whose lower bound is greater than 0. For instance, consider $I = (\frac{1}{2}, 3]$, then for any $x_1, x_2 \in I$:

$$|f(x_1) - f(x_2)| = |\sqrt{x_1} - \sqrt{x_2}| = \frac{1}{\sqrt{x_1} + \sqrt{x_2}} |x_1 - x_2| \leq \frac{1}{\sqrt{\frac{1}{2}} + \sqrt{\frac{1}{2}}} |x_1 - x_2| = \frac{\sqrt{2}}{2} |x_1 - x_2|.$$

Example 7.12: The function $f(x) = \sin x$ is Lipschitz on $\mathbb{R}$ with constant $L = 1$ (which means that the Lipschitz constant is *at most* 1). Indeed, we have

$$\begin{aligned} |\sin x_1 - \sin x_2| &= 2 \left| \sin \frac{x_1 - x_2}{2} \right| \left| \cos \frac{x_1 + x_2}{2} \right| \\ &\leq 2 \left| \sin \frac{x_1 - x_2}{2} \right| \\ &\leq |x_1 - x_2|. \end{aligned}$$

Proposition 7.10: If f is Lipschitz on I, then it is also continuous on I.

Proof. Fix some point $x_0 \in I$. Since f is Lipschitz, there exists some $L \geq 0$ such that

$$|f(x_0) - f(x)| \leq L|x_0 - x|, \qquad \forall x \in I.$$

Let $\varepsilon > 0$ and define $\delta = \frac{\varepsilon}{L}$. Then whenever $|x_0 - x| < \delta$ we have

$$|f(x_0) - f(x)| \leq L|x_0 - x| < L\delta = \varepsilon$$

which proves that f is continuous at x_0. Since x_0 was arbitrary, f is continuous on I. □

Uniform continuity

A function $f : \mathbb{R} \to \mathbb{R}$ is **uniformly continuous on an interval** $I \subseteq \mathbb{R}$ if for any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\forall x_1, x_2 \in I, \qquad |x_1 - x_2| < \delta \quad \Rightarrow \quad |f(x_1) - f(x_2)| < \varepsilon.$$

The crucial aspect in the above definition is that the same $\delta(\varepsilon)$ works for the *entire interval* I, not just at a single point. Recall that the definition of continuity held only at a given point x_0!

Example 7.13: The function $f(x) = \frac{1}{x}$ is *not* uniformly continuous on $(0,1)$: as $x \to 0^+$, compare, for instance x and $\sqrt{x}$:

$$\lim_{x\to 0^+} |x - \sqrt{x}| = 0$$

while

$$\lim_{x\to 0^+} |f(x) - f(\sqrt{x})| = \lim_{x\to 0^+} \left|\frac{1}{x} - \frac{1}{\sqrt{x}}\right| = \lim_{x\to 0^+} \left|\frac{1-\sqrt{x}}{x}\right| = +\infty.$$

So, points that become arbitrarily close on the x-axis, have their images become arbitrarily distant on the y-axis.

However, $f(x) = \frac{1}{x}$ is uniformly continuous on any interval further away from 0. For instance, consider $I = [a, +\infty)$ where $a > 0$. Then

$$|f(x_1) - f(x_2)| = \left|\frac{1}{x_1} - \frac{1}{x_2}\right| = \frac{|x_1 - x_2|}{x_1 x_2} \leq \frac{1}{a^2}|x_1 - x_2|.$$

Hence for any $\varepsilon > 0$ we can take $\delta = a^2\varepsilon$ to get that if $|x_1 - x_2| < \delta$ then $|f(x_1) - f(x_2)| \leq \frac{1}{a^2}|x_1 - x_2| < \frac{1}{a^2}\delta = \frac{1}{a^2}a^2\varepsilon = \varepsilon$.

Example 7.14: The function $f(x) = \sin(\frac{1}{x})$ is also *not* uniformly continuous on $(0,1)$. Indeed, as $x \to 0^+$ there are points that are arbitrarily close on the x-axis, while their distance on the y-axis is 2 (we're thinking here about points where the sine function achieves the values ± 1).

Theorem 7.11 (Heine-Cantor Theorem)**:** If f is continuous on a closed interval $I = [a, b]$ then is it uniformly continuous there.

Proof. We prove *by contradiction*. If the claim is not true, then f is not uniformly continuous on I. That is, $\exists \varepsilon > 0$ such that $\forall \delta > 0$, $\exists x', x'' \in I$, such that $|x' - x''| < \delta$ and $|f(x') - f(x'') \geq \varepsilon$.

So ε is fixed, but we can choose δ. Let $\delta = \frac{1}{n}$, $n \in \mathbb{N}_+$. Then for each n we have

$$|x'_n - x''_n| < \frac{1}{n} \qquad \text{and} \qquad |f(x'_n) - f(x''_n)| \geq \varepsilon.$$

We immediately observe that

$$\lim_{n\to\infty} |x'_n - x''_n| = 0,$$

while the corresponding function values will always remain at least ε apart, thus cannot converge to 0. Both sequences $\{x'_n\}_{n\in\mathbb{N}_+}$ and $\{x''_n\}_{n\in\mathbb{N}_+}$ are bounded, so the Bolzano-Weierstrass theorem applies.

Let's start with the sequence $\{x'_n\}_{n\in\mathbb{N}_+}$: it is bounded, and therefore there is a convergent subsequence $\{x'_{n_k}\}_{k\in\mathbb{N}_+}$. Call its limit $\overline{x}$:

$$\lim_{k\to\infty} x'_{n_k} = \overline{x}.$$

We note that since the interval $[a, b]$ is closed, the limit point $\overline{x}$ must also belong to $[a, b]$.

Now we consider the sequence $\{x''_n\}_{n\in\mathbb{N}_+}$: it is also bounded, but we need to be cautious about working with it. We know that $|x'_n - x''_n| < \frac{1}{n}$, where it is important that both elements have the same index n. For the first sequence we already took a subsequence with index $\{n_k\}_{k\in\mathbb{N}_+}$. This forces us to only consider the elements of $\{x''_n\}_{n\in\mathbb{N}_+}$ that also have the indices $\{n_k\}_{k\in\mathbb{N}_+}$. Hence we should start with the subsequence $\{x''_{n_k}\}_{k\in\mathbb{N}_+}$. This is again a bounded sequence, so it has a convergent subsequence $\{x''_{n_{k_j}}\}_{j\in\mathbb{N}_+}$. Let us show that its limit must also be $\overline{x}$. First, we summarize all that we have established so far:

$$\begin{aligned}
&1)\ \lim_{n\to\infty} |x'_n - x''_n| = 0,\\
&2)\ \lim_{k\to\infty} x'_{n_k} = \overline{x},\\
&3)\ \lim_{j\to\infty} x''_{n_{k_j}} \ \text{exists.}
\end{aligned}$$

So we have:

$$\begin{aligned}
\lim_{j\to\infty}\left(x''_{n_{k_j}} - \overline{x}\right) &= \lim_{j\to\infty}\left(x''_{n_{k_j}} - x'_{n_{k_j}} + x'_{n_{k_j}} - \overline{x}\right)\\
&= \lim_{j\to\infty}\left(x''_{n_{k_j}} - x'_{n_{k_j}}\right) + \lim_{j\to\infty}\left(x'_{n_{k_j}} - \overline{x}\right) = 0
\end{aligned}$$

So, indeed,

$$\lim_{j\to\infty} x''_{n_{k_j}} = \overline{x}.$$

The function f is continuous on I, so in particular it is continuous at $\overline{x} \in I$. This means that:

$$\lim_{j\to\infty} f(x'_{n_{k_j}}) = f(\lim_{j\to\infty} x'_{n_{k_j}}) = f(\overline{x}) = f(\lim_{j\to\infty} x''_{n_{k_j}}) = \lim_{j\to\infty} f(x''_{n_{k_j}}).$$

So we get that

$$\lim_{j\to\infty}\left(f(x'_{n_{k_j}}) - f(x''_{n_{k_j}})\right) = 0,$$

in contradiction to the assumption that

$$|f(x'_n) - f(x''_n)| \geq \varepsilon > 0, \qquad \forall n \in \mathbb{N}_+.$$

□

Example 7.15: The function $f(x) = \sqrt{x}$ is uniformly continuous on $[0, 1]$ due to the Heine-Cantor theorem. However, as we saw in Example 7.11 it is not Lipschitz on $[0, 1]$.

Chapter 8

Differential calculus

In a bid to study the *rate of change* of a function, in this chapter we introduce the notion of the *derivative*. For instance, if $x : \mathbb{R} \to \mathbb{R}$ given by $x(t)$ describes the location of a car along a straight road at time t, then the rate of change of x gives the speed of the car. The rate of change of the speed is the acceleration.

8.1 The derivative

To define the derivative, we start by defining a small increment in x and in y. For a function $f : \mathrm{dom}(f) \subseteq \mathbb{R} \to \mathbb{R}$, we fix $x_0 \in \mathrm{dom}(f)$ and assume that f is defined in a neighborhood $I_r(x_0)$ of x_0. Then for $x \in I_r(x_0)$ we denote

$$\Delta x = x - x_0$$

which is called the **increment in** x. Correspondingly we denote

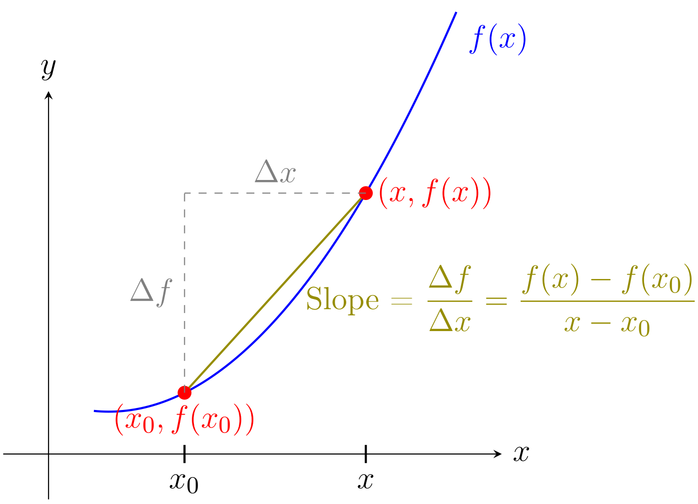
$$\Delta f = f(x) - f(x_0)$$

which is called the **increment in** f. We therefore have

$$x = x_0 + \Delta x \qquad \text{and} \qquad f(x) = f(x_0) + \Delta f.$$

The first step towards defining the derivative, is defining the **difference quotient of f between x_0 and x**:

$$\frac{\Delta f}{\Delta x} = \frac{f(x) - f(x_0)}{x - x_0} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$



The line connecting $(x_0, f(x_0))$ and another fixed point $(x_1, f(x_1))$ is called a **secant** and its equation is given by

$$s(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0)$$

By taking the limit of the secant as $x_1 \to x_0$ we obtain the *tangent line* to f at x_0, whose slope is the derivative of f at x_0:

Differentiability and the derivative

A function $f : \mathrm{dom}(f) \subseteq \mathbb{R} \to \mathbb{R}$ defined in a neighborhood of $x_0 \in \mathrm{dom}(f)$ is called **differentiable at** x_0 if $\lim_{x \to x_0} \frac{\Delta f}{\Delta x}$ exists and is finite. This number is called the **derivative of** f **at** x_0 and is denoted by $f'(x_0)$:

$$f'(x_0) = \lim_{x \to x_0} \frac{\Delta f}{\Delta x} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{\Delta x \to x_0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Other symbols used to denote the derivative at x_0 include

$$y'(x_0) \qquad \frac{\mathrm{d}f}{\mathrm{d}x}(x_0) \qquad \frac{\mathrm{d}}{\mathrm{d}x}f(x_0) \qquad \frac{\mathrm{d}y}{\mathrm{d}x}(x_0) \qquad \frac{\mathrm{d}}{\mathrm{d}x}y(x_0) \qquad \mathrm{D}f(x_0)$$

Given $I \subseteq \mathrm{dom}(f)$, we say that f **is differentiable on** I if f is differentiable at each $x_0 \in I$.

We observe that this defines a new function:

The derivative function

The definition of the derivative at a point induces the definition of the **derivative function** f' whose value at x_0 is $f'(x_0)$. Its domain is

$$\mathrm{dom}(f') = \{x \in \mathrm{dom}(f) \;\mid\; f \text{ is differentiable at } x\}.$$

Proposition 8.1: If f is differentiable at x_0 it is also continuous at x_0.

Proof. We have

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

□

Remark: The other implication is not true. For instance, the function $f(x) = |x|$ is continuous on $\mathbb{R}$, and in particular at $x_0 = 0$, however it is *not differentiable* at $x_0 = 0$. This can easily be seen by taking limits from the left and from the right and seeing that they are different:

$$\lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to 0^+} \frac{x}{x} = 1,$$
$$\lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to 0^-} \frac{-x}{x} = -1.$$

Example 8.1: Consider the affine function $f(x) = ax + b$. Then:

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{ax + b - ax_0 - b}{x - x_0} = a \lim_{x \to x_0} \frac{x - x_0}{x - x_0} = a.$$

Since this is true for all $x_0 \in \mathbb{R}$, we find that the derivative function of $f(x)$ is $f'(x) = a$.

Example 8.2: Consider the quadratic function $f(x) = ax^2$:

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{ax^2 - ax_0^2}{x - x_0} = a \lim_{x \to x_0} \frac{(x - x_0)(x + x_0)}{x - x_0} = 2ax_0.$$

Since this is true for all $x_0 \in \mathbb{R}$, we find that the derivative function of $f(x)$ is $f'(x) = 2ax$.

Example 8.3: Consider a monomial $f(x) = ax^n$. Using the formula

$$a^n - b^n = (a - b)\left(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \cdots + ab^{n-2} + b^{n-1}\right) = (a - b)\sum_{k=0}^{n-1} a^k b^{n-1-k}$$

we have

$$\begin{aligned} f'(x_0) &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \\ &= \lim_{x \to x_0} \frac{ax^n - ax_0^n}{x - x_0} \\ &= a \lim_{x \to x_0} \frac{(x - x_0)\sum_{k=0}^{n-1} x^k x_0^{n-1-k}}{x - x_0} \\ &= anx_0^{n-1}. \end{aligned}$$

Since this is true for all $x_0 \in \mathbb{R}$, we find that the derivative function of $f(x)$ is $f'(x) = anx^{n-1}$.

Example 8.4: We can also consider the more general power function $f(x) = x^\alpha$, $\alpha \in \mathbb{R}$. Here we will use the identity (which was given in Section 5.7)

$$\lim_{y \to 0} \frac{(1 + y)^\alpha - 1}{y} = \alpha, \qquad (\alpha \in \mathbb{R})$$

which we did not prove (but we *nearly* proved, only a few minor manipulations were missing). Suppose the $x_0 \in \mathrm{dom}(f)$, then

$$
\begin{aligned}
f'(x_0) &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \\
&= \lim_{\Delta x \to 0} \frac{(x_0 + \Delta x)^\alpha - x_0^\alpha}{\Delta x} \\
&= x_0^{\alpha - 1} \lim_{\Delta x \to 0} \frac{(1 + \frac{\Delta x}{x_0})^\alpha - 1}{\frac{\Delta x}{x_0}} \\
\left(\textit{substitute } y = \frac{\Delta x}{x_0}\right) \qquad &= x_0^{\alpha-1} \lim_{y \to 0} \frac{(1+y)^\alpha - 1}{y} \\
&= \alpha x_0^{\alpha - 1}.
\end{aligned}
$$

This is true for all x_0 for which the expression $x_0^{\alpha-1}$ is well-defined. This means that the derivative function is $f'(x) = \alpha x^{\alpha-1}$ and its domain consists of all $x \in \mathrm{dom}(f)$ for which $x^{\alpha-1}$ is well-defined.

Example 8.5: The function

$$f(x) = \sqrt{x} = x^{\frac{1}{2}} \quad \text{has domain } \mathrm{dom}(f) = [0, +\infty)$$

while its derivative

$$f'(x) = \frac{1}{2} x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}} \quad \text{has domain } \mathrm{dom}(f') = (0, +\infty).$$

Example 8.6: For the function $f(x) = \sin x$ we use the formula

$$\sin(\alpha) - \sin(\beta) = 2 \cos\left(\frac{\alpha + \beta}{2}\right) \sin\left(\frac{\alpha - \beta}{2}\right)$$

as follows:

$$
\begin{aligned}
f'(x_0) &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \\
&= \lim_{x \to x_0} \frac{\sin x - \sin x_0}{x - x_0} \\
&= \lim_{x \to x_0} \frac{2\cos\left(\frac{x + x_0}{2}\right) \sin\left(\frac{x - x_0}{2}\right)}{x - x_0} \\
&= \underbrace{\lim_{x \to x_0} \cos\left(\frac{x + x_0}{2}\right)}_{= \cos x_0} \underbrace{\lim_{x \to x_0} \frac{\sin\left(\frac{x - x_0}{2}\right)}{\frac{x - x_0}{2}}}_{=1} \\
&= \cos x_0.
\end{aligned}
$$

Since this is true for all $x_0 \in \mathbb{R}$, we find that the derivative function of $f(x)$ is $f'(x) = \cos x$.

Example 8.7: Similarly it can be shown that for $f(x) = \cos x$, $f'(x) = -\sin x$.

Example 8.8: Let $f(x) = a^x$. Then

$$\begin{aligned} f'(x_0) &= \lim_{\Delta x \to 0} \frac{a^{x_0+\Delta x} - a^{x_0}}{\Delta x} \\ &= a^{x_0} \underbrace{\lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}}_{=\ln a} \\ &= a^{x_0} \ln a \end{aligned}$$

Since this is true for all $x_0 \in \mathbb{R}$, we find that the derivative function of $f(x)$ is $f'(x) = a^x \ln a$. In particular,

$$(e^x)' = e^x.$$

This is the only function that is equal to its own derivative at every point, and that is one reason why the number e is so special.

Example 8.9: The area A of a disc of radius r is given by πr^2. Thinking of A as being a function of r, we can write

$$A(r) = \pi r^2.$$

Its derivative is

$$A'(r) = 2\pi r$$

which is exactly the circumference of the circle of radius r. That is, the rate of change of area of a disc of radius r is equal to the circumference of its boundary.

Example 8.10: If the location $x(t)$ at time t of a car driving along the x-axis is given by

$$x(t) = z + vt, \qquad z, v \in \mathbb{R},$$

(here the independent variable is *time* t, and x is the *dependent* variable) then the rate of change of the location is give by

$$x'(t) = v$$

which is the *speed* of the car.

8.2 Differentiation rules

Theorem 8.2 (Algebraic operations and linearity of the derivative)**:** Let $f, g : \mathbb{R} \to \mathbb{R}$ be differentiable at $x_0 \in \mathbb{R}$. Then their sum, difference, product and ratio are all also differentiable at x_0 and satisfy:

$$\begin{aligned} (f \pm g)'(x_0) &= f'(x_0) \pm g'(x_0), \\ (fg)'(x_0) &= f'(x_0)g(x_0) + f(x_0)g'(x_0), \\ \left(\frac{f}{g}\right)'(x_0) &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2}. \end{aligned}$$

Moreover, the formula for the sum and product implies that *differentiation is a linear operation*:

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0), \qquad \forall \alpha, \beta \in \mathbb{R}.$$

Proof. Let us prove the formula for the product of the two functions. The other proofs follow similar ideas.

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

□

The linearity of the derivative allows us to take a derivative of a **polynomial** since we already know the rule for a monomial $\frac{\mathrm{d}}{\mathrm{d}x}(x^n) = nx^{n-1}$. Let $p(x) = \sum_{k=0}^{n} a_k x^k$, then

$$p'(x) = \sum_{k=0}^{n} k a_k x^{k-1} = n a_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + 2a_2 x + a_1.$$

Example 8.11: Let $p(x) = 10x^6 - 3x^4 - 2x^3 + x^2 - 50$. Then

$$p'(x) = 60x^5 - 12x^3 - 6x^2 + 2x.$$

Example 8.12: We can also differentiate **rational functions**. Let

$$r(x) = \frac{5x^3 - 4x + 2}{3x - 10}.$$

Then

$$\begin{aligned}r'(x) &= \frac{(15x^2-4)(3x-10)-(5x^3-4x+2)\cdot 3}{(3x-10)^2}\\ &= \frac{45x^3-150x^2-12x+40-15x^3+12x-6}{9x^2-60x+100}\\ &= \frac{30x^3-150x^2+34}{9x^2-60x+100}.\end{aligned}$$

Example 8.13: Let $f(x) = \tan x$. Then

$$\begin{aligned}f'(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\sin x}{\cos x}\right)\\ &= \frac{\sin' x\cos x - \sin x\cos' x}{\cos^2 x}\\ &= \frac{\cos^2 x + \sin^2 x}{\cos^2 x}\\ &= \frac{1}{\cos^2 x} = 1 + \tan^2 x.\end{aligned}$$

Theorem 8.3 (Chain Rule): Let $f, g : \mathbb{R} \to \mathbb{R}$. Assume that f is differentiable at x_0 and that g is differentiable at $y_0 = f(x_0)$. Then $h = g \circ f$ is differentiable at x_0 and the derivative is given by:

$$h'(x_0) = g'(y_0)f'(x_0) = g'(f(x_0))f'(x_0).$$

Proof. Recall the definition of the derivative of g at y_0:

$$g'(y_0) = \lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0}$$

which can be rewritten as follows:

$$\begin{aligned} 0 &= \lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} - g'(y_0) \\ &= \lim_{y \to y_0} \frac{g(y) - g(y_0) - g'(y_0)(y - y_0)}{y - y_0}. \end{aligned}$$

Call the numerator in the above limit $\psi(y)$:

$$\psi(y) = g(y) - g(y_0) - g'(y_0)(y - y_0).$$

The function $\psi(y)$ is defined in a small neighborhood of y_0. Then we know that not only $\lim_{y \to y_0} \psi(y) = 0$, but also

$$\lim_{y \to y_0} \frac{\psi(y)}{y - y_0} = 0$$

(this means that $\psi(y) = o(y - y_0)$ as $y \to y_0$). That is, as $y \to y_0$, $\psi(y)$ tends to 0 faster than $y - y_0$. Denote:

$$\varphi(y) = \frac{\psi(y)}{y - y_0} \qquad \text{so that} \qquad \psi(y) = \varphi(y)(y - y_0).$$

Then

$$\lim_{y \to y_0} \varphi(y) = 0.$$

Returning to the definition of ψ, we have:

$$g(y) - g(y_0) - g'(y_0)(y - y_0) = \psi(y) = \varphi(y)(y - y_0).$$

Hence, for all y in a neighborhood of y_0:

$$g(y) - g(y_0) = g'(y_0)(y - y_0) + \varphi(y)(y - y_0).$$

Plugging in $y = f(x)$ and $y_0 = f(x_0)$, and dividing by $x - x_0$ this becomes

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = g'(f(x_0))\frac{f(x) - f(x_0)}{x - x_0} + \varphi(f(x))\frac{f(x) - f(x_0)}{x - x_0}.$$

Recalling that $h = g \circ f$ we have

$$\frac{h(x) - h(x_0)}{x - x_0} = g'(f(x_0))\frac{f(x) - f(x_0)}{x - x_0} + \varphi(f(x))\frac{f(x) - f(x_0)}{x - x_0}.$$

Taking the limit $x \to x_0$, we get

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

Hence we find that the derivative of $h = g \circ f$ at x_0 exists and is equal to

$$h'(x_0) = (g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

□

Chain rule: avoiding confusion!

It is easy to get confused with notation when dealing with the chain rule, so it is best to be cautious. If we write $h = g \circ f$ then the chain rule can be expressed as

$$\frac{\mathrm{d}h}{\mathrm{d}x}(x_0) = \frac{\mathrm{d}g}{\mathrm{d}y}(y_0)\frac{\mathrm{d}f}{\mathrm{d}x}(x_0)$$

where $y_0 = f(x_0)$. Alternatively, if we write $y = f(x)$ and $z = g(y)$ then we can write the chain rule also as:

$$\frac{\mathrm{d}z}{\mathrm{d}x}(x_0) = \frac{\mathrm{d}z}{\mathrm{d}y}(y_0)\frac{\mathrm{d}y}{\mathrm{d}x}(x_0)$$

where $y_0 = f(x_0)$. This latter expression is easier to remember, because one could imagine that the terms $\mathrm{d}y$ on the right hand side cancel out (though we are not allowed to actually do that!).

Example 8.14: Consider the function $z = h(x) = \sqrt{1-x^2}$ which is the composition of $y = f(x) = 1 - x^2$ with $z = g(y) = \sqrt{y}$. Recalling that

$$f'(x) = -2x \qquad \text{and} \qquad g'(y) = \frac{1}{2\sqrt{y}}$$

we have:

$$\frac{\mathrm{d}h}{\mathrm{d}x}(x) = \frac{\mathrm{d}g}{\mathrm{d}y}(y)\frac{\mathrm{d}f}{\mathrm{d}x}(x) = \frac{1}{2\sqrt{1-x^2}}(-2x) = -\frac{x}{\sqrt{1-x^2}}.$$

Theorem 8.4 (Derivative of the inverse function)**:** Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and invertible in a neighborhood of $x_0 \in \mathbb{R}$. Moreover, suppose that f is differentiable at x_0 and that $f'(x_0) \neq 0$. Then the inverse function $f^{-1}(y)$ is differentiable at $y_0 = f(x_0)$ and the derivative there is given by

$$\frac{\mathrm{d}}{\mathrm{d}y}f^{-1}(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$

Proof. By Theorem 7.9, f^{-1} is continuous on a neighborhood of $y_0 = f(x_0)$. For y in the neighborhood we have

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{\frac{f(x)-f(x_0)}{x-x_0}}.$$

The Substitution Theorem (Theorem 5.12) implies that taking $y \to y_0$ on the left hand side, is the same as taking $x \to x_0$ on the right hand side:

$$\lim_{y\to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \lim_{x\to x_0} \frac{1}{\frac{f(x)-f(x_0)}{x-x_0}} = \frac{1}{f'(x_0)}.$$

□

Example 8.15: Let $y = f(x) = \tan x$. We saw that $y'(x) = 1 + \tan^2 x = 1 + y^2(x)$. Denote $x = f^{-1}(y) = \arctan(y)$. Then:

$$\frac{\mathrm{d}}{\mathrm{d}y} \arctan(y) = \frac{1}{1 + y^2}.$$

Example 8.16: Let $y = f(x) = \sin x$. We saw that $y'(x) = \cos x = \sqrt{1 - \sin^2 x} = \sqrt{1 - y^2(x)}$. Then:

$$\frac{\mathrm{d}}{\mathrm{d}y} \arcsin(y) = \frac{1}{\sqrt{1 - y^2}}.$$

Similarly, it can be shown that

$$\frac{\mathrm{d}}{\mathrm{d}y} \arccos(y) = -\frac{1}{\sqrt{1 - y^2}}.$$

Example 8.17: Let $y = f(x) = a^x$. Then $y'(x) = (\ln a)a^x = (\ln a)y$. The inverse function is $x = f^{-1}(y) = \log_a y$, and we have:

$$\frac{\mathrm{d}}{\mathrm{d}y} \log_a y = \frac{1}{(\ln a)y}$$

One can check that the same result is obtained for the derivative of $\log_a(-y)$, so that we have

$$\frac{\mathrm{d}}{\mathrm{d}y} \log_a |y| = \frac{1}{(\ln a)y} \qquad y \neq 0.$$

In particular:

$$\frac{\mathrm{d}}{\mathrm{d}y} \ln |y| = \frac{1}{y} \qquad y \neq 0.$$

Logarithmic derivative

If $f : \mathbb{R} \to \mathbb{R}_+$ attains positive values, then $h(x) = \ln(f(x))$ is well-defined. Moreover, if f is differentiable, then so is h and we have $h'(x) = \ln'(f(x))f'(x)$. Using the formula for the derivative of the logarithm, we have:

$$\frac{\mathrm{d}}{\mathrm{d}x} \ln(f(x)) = \frac{f'(x)}{f(x)}.$$

Proposition 8.5: Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable. Then:

$$\begin{aligned} f \text{ odd} \quad &\Rightarrow \quad f' \text{ even} \\ f \text{ even} \quad &\Rightarrow \quad f' \text{ odd} \end{aligned}$$

Proof. Let us consider the first case. The second one follows a similar line of reasoning. If f is odd, then $f(-x) = -f(x)$. Taking derivatives of both sides we have $-f'(-x) = -f'(x)$ so that $f'(-x) = f'(x)$ and the derivative function is even. □

Important derivatives

$$\frac{\mathrm{d}}{\mathrm{d}x} x^\alpha = \alpha x^{\alpha-1} \qquad (\forall \alpha \in \mathbb{R})$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \sin x = \cos x$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \cos x = -\sin x$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \tan x = 1 + \tan^2 x = \frac{1}{\cos^2 x}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \arcsin x = \frac{1}{\sqrt{1-x^2}}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \arccos x = -\frac{1}{\sqrt{1-x^2}}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \arctan x = \frac{1}{1+x^2}$$

$$\frac{\mathrm{d}}{\mathrm{d}x} a^x = (\ln a)\, a^x \qquad (\forall a > 0) \qquad \text{in particular, } \frac{\mathrm{d}}{\mathrm{d}x} e^x = e^x$$

$$\frac{\mathrm{d}}{\mathrm{d}x} \log_a |x| = \frac{1}{(\ln a)\, x} \qquad (\forall a > 0,\ a \neq 1) \quad \text{in particular, } \frac{\mathrm{d}}{\mathrm{d}x} \ln |x| = \frac{1}{x}$$

8.3 Where differentiability fails

We have already seen that the function $f(x) = |x|$ is not differentiable at $x_0 = 0$, since the right- and left-limits of the expression $\frac{\Delta f}{\Delta x}$ are different (they are ± 1). This suggests that we can define differentiability on both sides:

Definition 8.6: Suppose that $f : \mathbb{R} \to \mathbb{R}$ is defined on an interval $[x_0, x_0 + r)$ for some $r > 0$. Then f is **right-differentiable** at x_0 if

$$\lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \quad \text{exists and is finite. It is denoted } f'_+(x_0)$$

Similarly, if f is defined on $(x_0 - r, x_0]$ then we say that f is **left-differentiable** at x_0 if

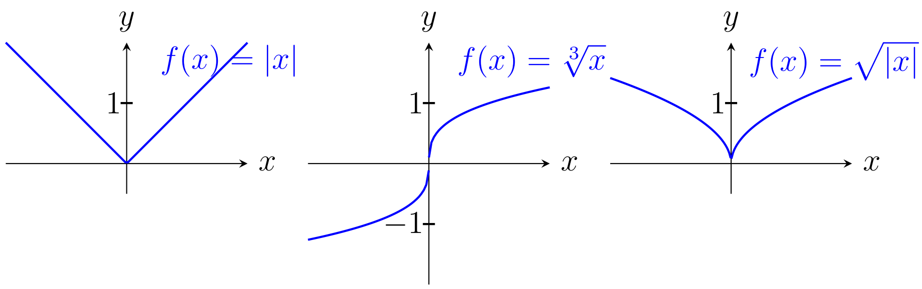
$$\lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \quad \text{exists and is finite. It is denoted } f'_-(x_0).$$

Proposition 8.7: A function f is differentiable at x_0 if and only if it is both left- and right-differentiable at x_0, and $f'_+(x_0) = f'_-(x_0)$. In this case $f'(x_0) = f'_+(x_0) = f'_-(x_0)$.

Proof. The proof is almost immediate, and left as an exercise (see Proposition 4.4). □

Remark:

1. If $f'_+(x_0) \neq f'_-(x_0)$ are both finite, then x_0 is a **corner point** of f (e.g. 0 is a corner point of $|x|$).
2. If one or both of $f'_+(x_0)$ and $f'_-(x_0)$ are $+\infty$ or $-\infty$, then x_0 is a point with **vertical tangent** of f. For example for $f(x) = \sqrt{x}$
$$f'_+(0) = +\infty.$$
Another example is $f(x) = \sqrt[3]{x}$, for which $f'(x) = \frac{1}{3}x^{-\frac{2}{3}}$ is defined for $x \neq 0$. However,
$$f'_\pm(0) = \lim_{x \to 0^\pm} \frac{1}{3}x^{-\frac{2}{3}} = +\infty.$$
3. If $f'_+(x_0)$ and $f'_-(x_0)$ are infinite of opposite signs, then x_0 is a **cusp** point of f. For example, consider $f(x) = \sqrt{|x|}$. Then by definition
$$f'_\pm(0) = \lim_{x \to 0^\pm} \frac{\sqrt{|x|}}{x} = \lim_{x \to 0^\pm} \frac{\sqrt{|x|}}{\operatorname{sign}(x)|x|} = \lim_{x \to 0^\pm} \frac{1}{\operatorname{sign}(x)\sqrt{|x|}} = \pm\infty.$$



Theorem 8.8: If a function f is continuous at x_0 and differentiable in a neighborhood of x_0 (excluding x_0 itself) then f is differentiable at x_0 if $\lim_{x \to x_0} f'(x)$ exists and is finite. If so, then $f'(x_0) = \lim_{x \to x_0} f'(x)$.

Proof. We don't have (yet) the tools to prove this theorem. We'll come back to it later. □

Example 8.18: Are there numbers $a, b \in \mathbb{R}$ such that the function

$$f(x) = \begin{cases} a\sin(2x) - 4 & x < 0 \\ b(x-1) + e^x & x \geq 0 \end{cases}$$

is differentiable at $x_0 = 0$?

First, f needs to be continuous at $x_0 = 0$. So we need:

$$\lim_{x \to 0^-} a \sin(2x) - 4 = -4 \qquad \text{to be equal to} \qquad -b + 1$$

Hence $b = 5$. Now we need the left- and right-derivatives to agree.

$$\lim_{x \to 0^+} f'(x) = \lim_{x \to 0^+} (b + e^x) = 6,$$
$$\lim_{x \to 0^-} f'(x) = \lim_{x \to 0^-} 2a \cos(2x) = 2a.$$

For these to be equal, we impose $a = 3$.

8.4 Extrema and critical points

We can now dig deeper into our previous definitions of the supremum, infimum, maximum and minimum of sets.

Local maximum

A point x_0 is a **local maximum point** for f if there exists a neighborhood $I_r(x_0)$ such that

$$f(x) \leq f(x_0), \qquad \forall x \in \operatorname{dom}(f) \cap I_r(x_0).$$

Then $f(x_0)$ is a **local maximum** of f.

Global maximum

A point x_0 is a **global maximum point** for f if

$$f(x) \leq f(x_0), \qquad \forall x \in \operatorname{dom}(f).$$

Then $f(x_0)$ is the **global maximum** of f. The maximum is **strict** if $f(x) < f(x_0)$ for all $x \neq x_0$.

Local minimum

A point x_0 is a **local minimum point** for f if there exists a neighborhood $I_r(x_0)$ such that

$$f(x) \geq f(x_0), \qquad \forall x \in \operatorname{dom}(f) \cap I_r(x_0).$$

Then $f(x_0)$ is a **local minimum** of f.

Global minimum

A point x_0 is a **global minimum point** for f if

$$f(x) \geq f(x_0), \qquad \forall x \in \mathrm{dom}(f).$$

Then $f(x_0)$ is the **global minimum** of f. The minimum is **strict** if $f(x) > f(x_0)$ for all $x \neq x_0$.

Any one of the points described above will be called an ***extremum point*** of the function.

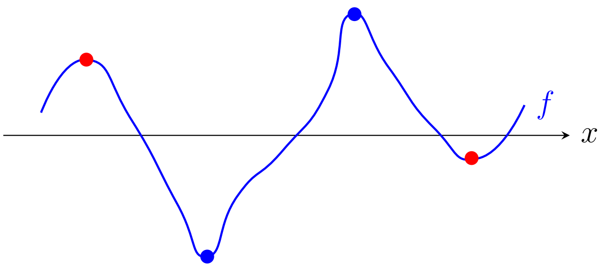


Figure 8.1: Examples of global min/max in blue, and local min/max in red

For a differentiable function, minima and maxima are points where the derivative vanishes (i.e. the tangent is parallel to the x-axis). Points where the derivative vanishes are called *critical points*:

Critical point

A **critical point** of a function $f : \mathbb{R} \to \mathbb{R}$ is a point $x_0 \in \mathbb{R}$ at which f is differentiable and $f'(x_0) = 0$.

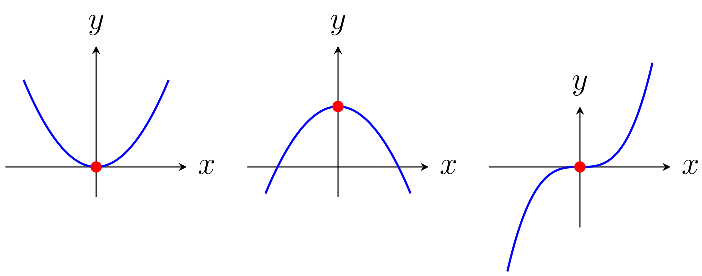


Figure 8.2: Three types of critical points: local minimum (left), local maximum (center), and inflection point (right).

Theorem 8.9 (Fermat's Theorem)**:** If f is differentiable at an extremum point x_0 then $f'(x_0) = 0$.

Proof. Suppose that x_0 is a local maximum. Let $I_r(x_0)$ be a neighborhood on which $f(x) \leq f(x_0)$ for all $x \in I_r(x_0)$. Then within this neighborhood, $\Delta f = f(x) - f(x_0) \leq 0$. We therefore have:

$$\text{The fraction} \quad \frac{f(x) - f(x_0)}{x - x_0} \quad \text{is} \quad \begin{cases} \leq 0 & \text{if } x > x_0 \\ \geq 0 & \text{if } x < x_0 \end{cases}$$

By Proposition 8.7 the left- and right-derivatives at x_0 must equal $f'(x_0)$. The only way this is possible is if they are 0. □

So we see that at an extremum the derivative (if exists) is 0, i.e. it is a critical point. An extremum might also be found at a point where f is not differentiable (think about $|x|$) or at boundary points of the domain (think about $\arcsin x$). So we summarize:

Finding extrema

To find extreme points of a function we must look at the following points:

- Critical points,
- Points where f is not differentiable,
- Points on the boundary of the domain.

Among these we will find the extrema (but we have to check case-by-case).

8.5 The Theorems of Rolle, Lagrange and Cauchy

Theorem 8.10 (Rolle's Theorem)**:** Let $a, b \in \mathbb{R}$, $a < b$, and let $f : \mathbb{R} \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$ then there exists $x_0 \in (a, b)$ with $f'(x_0) = 0$. That is, f has at least one critical point in (a, b).

Proof. From Weierstrass' Theorem, we know that $f([a, b]) = [m, M]$ where

$$m = \min_{x \in [a,b]} f(x) = f(x_m) \qquad \text{and} \qquad M = \max_{x \in [a,b]} f(x) = f(x_M)$$

where x_m is a global minimum for f on $[a, b]$ and x_M is a global maximum for f on $[a, b]$.

If $m = M$ then f is constant on $[a, b]$ so that $f'(x) = 0$ for every $x \in [a, b]$ and the proof is done. Otherwise, $m < M$. Hence we have

$$m \leq f(a) = f(b) \leq M.$$

Since $m < M$, at least one of the $\leq$ above must be a strict inequality.

If $f(a) = f(b) < M$, then x_M cannot be a or b, so $x_M \in (a, b)$. By Fermat's Theorem (Theorem 8.9), since x_M is a differentiable extremum point, $f'(x_M) = 0$ and the proof is done. The case $m < f(a) = f(b)$ follows in a similar way. □

Remark: We have just proven that there is a critical point between a and b. It is important to note that there could be more than one critical point. The proof only shows that there exists *at least* one.

Theorem 8.11 (Mean Value Theorem, a.k.a. Lagrange's Theorem)**:** Let $a, b \in \mathbb{R}$, $a < b$, and let $f : \mathbb{R} \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists a point $x_0 \in (a, b)$ such that

$$\frac{f(b) - f(a)}{b - a} = f'(x_0).$$

Every such point x_0 is called a **Lagrange point** for f in (a, b).

Proof. The idea of the proof is to 'tilt' the function so that the values at the endpoints become equal, and we can apply Rolle's Theorem. To this end, define:

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a), \qquad \forall x \in [a, b].$$

Then g is continuous on $[a, b]$, differentiable on (a, b), with

$$g(a) = f(a) \qquad \text{and} \qquad g(b) = f(b) - (f(b) - f(a)) = f(a).$$

Hence Rolle's Theorem can be applied to g. Noting that

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a},$$

we see that $g'(x_0) = 0$ if and only if

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}$$

and the proof is complete. □

Theorem 8.12 (Cauchy's Theorem)**:** Let $a, b \in \mathbb{R}$, $a < b$, and let $f, g : \mathbb{R} \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Suppose that $g'(x) \neq 0$ for all $x \in (a, b)$. Then there exists a point $x_0 \in (a, b)$ such that

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x_0)}{g'(x_0)}.$$

Remark: Note that this theorem generalizes the Mean Value Theorem. Indeed, by taking $g(x) = x$ we recover the Mean Value Theorem.

Proof. First we claim the $g(a) \neq g(b)$. Indeed, by contradiction, if those values were equal, then Rolle's Theorem would imply that there exists some $x_0 \in (a, b)$ such that $g'(x_0) = 0$, contrary to our assumption. Define

$$h(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)), \qquad \forall x \in [a, b].$$

Then h is continuous on $[a, b]$, differentiable on (a, b), with

$$h(a) = f(a) \qquad \text{and} \qquad h(b) = f(b) - (f(b) - f(a)) = f(a).$$

Hence Rolle's Theorem can be applied to h. Noting that

$$h'(x) = f'(x) - \frac{f(b) - f(a)}{g(b) - g(a)}g'(x),$$

we see that $h'(x_0) = 0$ if and only if

$$f'(x_0) = \frac{f(b) - f(a)}{g(b) - g(a)}g'(x_0)$$

which completes the proof. □

8.6 First and second finite increment formulas

For a function $f : \mathbb{R} \to \mathbb{R}$ that is differentiable at $x_0 \in \mathbb{R}$, from the definition of the derivative

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

we can subtract $f'(x_0)$, multiply and divide it by $x - x_0$, and insert it into the limit, to obtain

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

By definition of the little o symbol, this means that the numerators in the above three expressions are little o's of their denominators, as $x \to x_0$. This leads us to

First increment formula

The first increment formula for a differentiable function f at x_0 states that

$$f(x) - f(x_0) = f'(x_0)(x - x_0) + o(x - x_0), \quad x \to x_0,$$

or, equivalently,

$$\Delta f = f'(x_0)\Delta x + o(\Delta x), \quad \Delta x \to 0.$$

For f that is differentiable at x_0, this formula gives us an approximation of $f(x)$ at nearby points x. Figure 8.3 demonstrates this formula graphically.

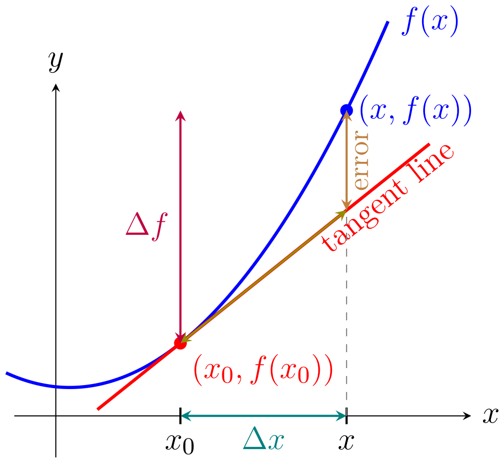


Figure 8.3: The 'error' is of order $o(\Delta x)$ as $\Delta x \to 0$

Another approximation method uses the Mean Value Theorem (Theorem 8.11). Suppose now that f is differentiable on an entire interval $I \subseteq \mathbb{R}$. Let x_1, x_2 be two points in I. Then by the Mean Value Theorem, there exists $\overline{x}$ between x_1 and x_2 such that

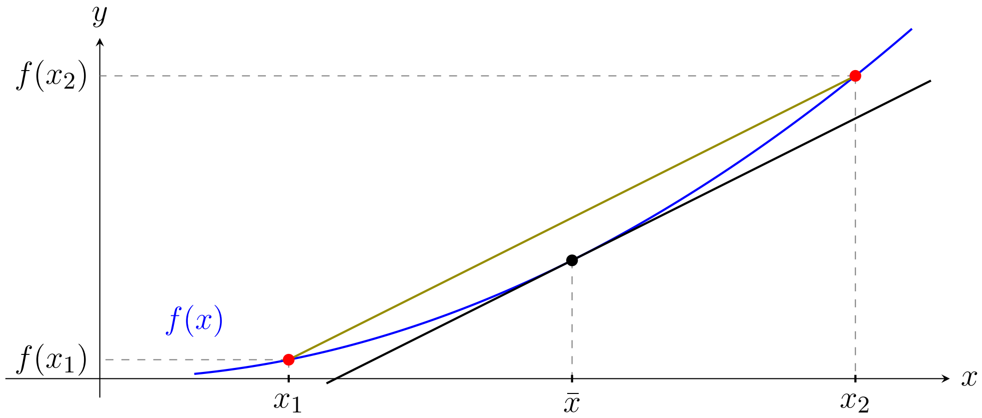
$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\overline{x}).$$

Second increment formula

The second increment formula for a differentiable function f on an interval I with x_1, x_2 in I, states that

$$f(x_2) - f(x_1) = f'(\overline{x})(x_2 - x_1)$$

for some $\overline{x}$ between x_1 and x_2.



The following proposition follows as a consequence:

Proposition 8.13: Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable on an interval $I \subseteq \mathbb{R}$. Then

$$f \text{ is constant on } I \qquad \Leftrightarrow \qquad f'(x) = 0, \quad \forall x \in I.$$

Proof. **Direction $\Rightarrow$.** If f is constant on I, then for any $x_0 \in I$, we have $\frac{f(x)-f(x_0)}{x-x_0}$ for any other $x \in I$, so the $f'(x_0) = 0$ by the definition of the derivative.

Direction $\Leftarrow$. Take any two $x_1, x_2 \in I$. By the second increment formula

$$f(x_2) - f(x_1) = f'(\overline{x})(x_2 - x_1)$$

for some $\overline{x}$ between x_1 and x_2. But $f'(x) = 0$ for any $x \in I$ by assumption, so that $f(x_2) = f(x_1)$ and the proof is complete. □

Proposition 8.14: Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable on an interval $I \subseteq \mathbb{R}$ with bounded derivative on I. Define

$$L = \sup_{x \in I} |f'(x)|$$

which cannot be $+\infty$ since the derivative is bounded on I. Then f is Lipschitz on I with Lipschitz constant L.

Proof. Our goal is simple: verify that

$$|f(x_1) - f(x_2)| \leq L|x_1 - x_2| \qquad \forall x_1, x_2 \in I$$

with L being the *optimal* (i.e. smallest) constant satisfying this inequality. By the second increment formula, for any $x_1, x_2 \in I$ there exists some $\overline{x}$ between x_1 and x_2 such that $f(x_2) - f(x_1) = f'(\overline{x})(x_2 - x_1)$. Taking absolute values and estimating, we have

$$|f(x_2) - f(x_1)| = |f'(\overline{x})| \cdot |x_2 - x_1| \leq L|x_2 - x_1|$$

where the inequality follows from our assumption. This is enough to prove that f is Lipschitz on I, but it doesn't prove that L is the *optimal (or Lipschitz)* constant in the inequality. Denote by L_{opt} the optimal constant. Then we know that

$$L_{\text{opt}} \leq L.$$

Now we'll show that $L_{\text{opt}} \geq L$, thus concluding that $L_{\text{opt}} = L$. Fix $x_0 \in I$. We know that

$$\left|\frac{f(x) - f(x_0)}{x - x_0}\right| = \frac{|f(x) - f(x_0)|}{|x - x_0|} \leq L_{\text{opt}} \qquad \forall x, x_0 \in I,\ x \neq x_0.$$

Taking the limit $x \to x_0$, we have

$$|f'(x_0)| = \left|\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}\right| = \lim_{x \to x_0} \left|\frac{f(x) - f(x_0)}{x - x_0}\right| \leq L_{\text{opt}}$$

[Note that we are allowed to exchange the order of the limit and the absolute value due to the Substitution Theorem (Theorem 5.12) using the fact that the limit of $\frac{f(x)-f(x_0)}{x-x_0}$ *exists.]* Taking the supremum over all $x_0 \in I$ of the above and using the definition of L, we arrive at

$$L \leq L_{\text{opt}}$$

and the proof is complete. □

8.7 Monotonicity intervals

Theorem 8.15: Let $f : \mathbb{R} \to \mathbb{R}$ and suppose that f is differentiable on an interval I. Then (a)

$$f'(x) \text{ has the same sign (or 0) throughout } I \qquad \Leftrightarrow \qquad f \text{ is monotone on } I$$

and (b)

$$f'(x) \text{ has a strict sign throughout } I \qquad \Rightarrow \qquad f \text{ is strictly monotone on } I.$$

Proof. **We start by proving (a)(⇐).** Suppose that f is monotone increasing on I (the monotone decreasing case will be similar). Let $x_0 \in I$ and assume that it is not at the boundary of I (i.e. there are other points to its left and to its right that are in I).

So for any $x \in I$ with $x \leq x_0$ we have $f(x) - f(x_0) \leq 0$. Therefore $\frac{f(x)-f(x_0)}{x-x_0} \geq 0$.

For any $x \in I$ with $x \geq x_0$ we have $f(x) - f(x_0) \geq 0$. Again, $\frac{f(x)-f(x_0)}{x-x_0} \geq 0$.

Therefore, for any $x_0, x \in I$

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0.$$

By Theorem 5.3 (local sign of limits) the limit as $x \to x_0$ is also non-negative:

$$f'(x_0) \geq 0.$$

This proves the assertion for all x_0 that are not on the boundary of I. If x_0 is on the boundary of I, the same argument can be repeated with one-sided limits.

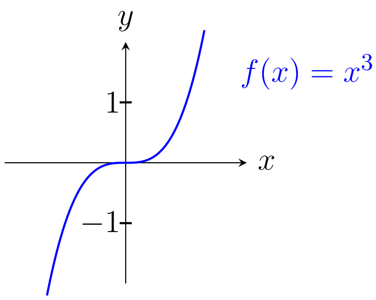
Now we prove (a)($\Rightarrow$). Let $x_1, x_2 \in I$ with $x_1 < x_2$. By the second increment formula, there exists $\overline{x} \in (x_1, x_2)$ such that

$$f(x_2) - f(x_1) = f'(\overline{x})(x_2 - x_1).$$

By assumption, $f'(\overline{x}) \geq 0$ and $x_2 - x_1 > 0$, so that $f(x_2) \geq f(x_1)$ which completes the proof.

The proof of (b) follows immediately, since in the above argument $f'(\overline{x}) > 0$, hence $f(x_2) > f(x_1)$. $\square$

Observe that part (b) has a one-sided implication; the other implication is not true. For example, the function $f(x) = x^3$ is strictly increasing on $\mathbb{R}$, however its derivative function is not strictly positive (it vanishes at $x = 0$).



Corollary 8.16: Let $f : \mathbb{R} \to \mathbb{R}$ and suppose that f is differentiable on an interval I. Let $x_0 \in I$ be in the interior of I (not on the boundary). Then:

- If $f'(x) \geq 0$ to the left of x_0 and $f'(x) \leq 0$ to the right of x_0, then x_0 is a local maximum.
- If $f'(x) \leq 0$ to the left of x_0 and $f'(x) \geq 0$ to the right of x_0, then x_0 is a local minimum.

Proof. This simple proof is left as an exercise. $\square$

Finding extrema and monotonicity intervals of a function

Using Theorem 8.15 and Corollary 8.16, we see that to find extrema and monotonicity intervals of a function, all we need to do is to know the sign and zeroes of its derivative.

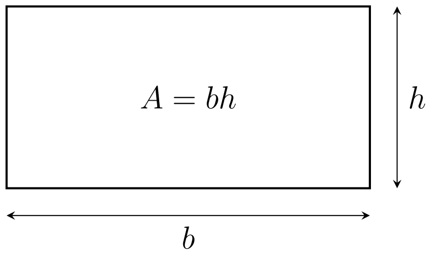
Example 8.19: Let $f(x) = xe^{2x}$. Then

$$f'(x) = e^{2x}(1 + 2x)$$

vanishes only when $1 + 2x = 0$, i.e. for $x = -\frac{1}{2}$. We see that $f'(x) > 0$ for $x > \frac{-}{1}2$ (f is strictly increasing on $[-\frac{1}{2}, +\infty)$) and $f'(x) < 0$ for $x < -\frac{1}{2}$ (f is strictly decreasing on $(-\infty, -\frac{1}{2}]$), so that $f(-\frac{1}{2})$ is a global minimum.

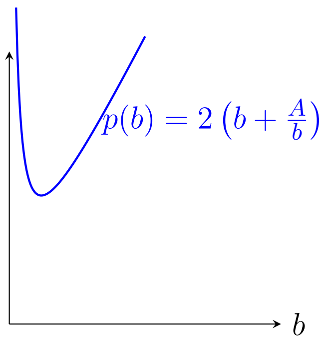
Example 8.20: Consider a rectangle with sides of length b and h, area $A = bh$ and perimeter

$$p = 2(b + h).$$



1. **Q: for fixed area A, what is the minimal perimeter?**

   We fix A, retain b as a variable, and express $h = \frac{A}{b}$. Then $p = 2(b + \frac{A}{b})$.

   We compute the derivative function

   $$p'(b) = 2\left(1 - \frac{A}{b^2}\right)$$

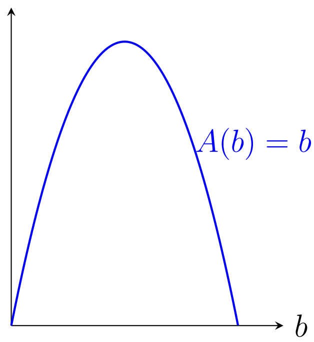
   which vanishes when $b = \sqrt{A}$. It is easy to check that $p'(b) < 0$ for $b < \sqrt{A}$ and $p'(b) > 0$ for $b > \sqrt{A}$. Hence $b = \sqrt{A}$ is the global minimum of this function, and

   $$p(\sqrt{A}) = 4\sqrt{A}.$$

   In this case $b = h$ and the rectangle is a square. This shows that **among all rectangles with a fixed area, the square minimizes the perimeter**.

2. **Q: for fixed perimeter p, what is the maximal area?**

   We fix p, retain b as a variable, and express $h = \frac{p}{2} - b$. Then $A = b(\frac{p}{2} - b)$.



We see that A vanishes for $b = 0$ and $b = \frac{p}{2}$. We compute

$$A'(b) = \frac{p}{2} - 2b$$

which vanishes when $b = \frac{p}{4}$. It is easy to check that $A'(b) > 0$ for $b < \frac{p}{4}$ and $A'(b) < 0$ for $b > \frac{p}{4}$. Hence $b = \frac{p}{4}$ is the global maximum of this function, and

$$A(\frac{p}{4}) = \frac{p^2}{16}.$$

In this case $b = h = \frac{p}{4}$ and the rectangle is a square. This shows that **among all rectangles with a fixed perimeter, the square maximizes the area**.

8.8 Higher-order derivatives

Second-order derivatives

Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable. If the derivative function f' is differentiable at $x_0 \in \mathbb{R}$, we say that f **is twice differentiable at** x_0. We write

$$f''(x_0) = (f')'(x_0)$$

to express the **second derivative of** f **at** x_0. The derivative function of f' is denoted f'' and is called the **second derivative of** f. Some common symbols include

$$y''(x_0) \qquad \frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x_0) \qquad \frac{\mathrm{d}^2}{\mathrm{d}x^2} f(x_0) \qquad \frac{\mathrm{d}^2 y}{\mathrm{d}x^2}(x_0) \qquad \frac{\mathrm{d}^2}{\mathrm{d}x^2} y(x_0) \qquad \mathrm{D}^2 f(x_0)$$

Higher-order derivatives

More generally, we can define derivatives of any integer order $k \in \mathbb{N}$ inductively.

Since it is not convenient to write $f\overbrace{''''''}^{k\text{ times}}(x_0)$ we write $f^{(k)}$, defined inductively as

$$f^{(k)}(x_0) = (f^{(k-1)})'(x_0)$$

to express the k^{th} **derivative of** f **at** x_0. The k^{th} derivative function of f is denoted $f^{(k)}$ and is called the k^{th} **derivative of** f. Some common symbols include

$$y^{(k)}(x_0) \qquad \frac{\mathrm{d}^k f}{\mathrm{d}x^k}(x_0) \qquad \frac{\mathrm{d}^k}{\mathrm{d}x^k}f(x_0) \qquad \frac{\mathrm{d}^k y}{\mathrm{d}x^k}(x_0) \qquad \frac{\mathrm{d}^k}{\mathrm{d}x^k}y(x_0) \qquad \mathrm{D}^k f(x_0)$$

Example 8.21: The function $f(x) = x^n$, $n \in \mathbb{N}_+$ has the following derivatives, defined on $\mathbb{R}$:

$$\begin{aligned}
f'(x) &= nx^{n-1} \\
f''(x) &= n(n-1)x^{n-2} \\
&\vdots \\
f^{(k)}(x) &= n(n-1)\cdots(n-k+2)(n-k+1)x^{n-k} \\
&\vdots \\
f^{(n)}(x) &= n(n-1)\cdots 3 \cdot 2 \cdot 1x^{n-n} = n! \\
f^{(n+1)}(x) &= 0.
\end{aligned}$$

Example 8.22: Since $\frac{\mathrm{d}}{\mathrm{d}x}e^x = e^x$, we have $\frac{\mathrm{d}^n}{\mathrm{d}x^n}e^x = e^x$ for all $n \in \mathbb{N}$.

Example 8.23: For sin and cos we have:

$$\begin{aligned}
\frac{\mathrm{d}^n}{\mathrm{d}x^n}\sin x &= \sin\left(x + n\frac{\pi}{2}\right) \\
\frac{\mathrm{d}^n}{\mathrm{d}x^n}\cos x &= \cos\left(x + n\frac{\pi}{2}\right)
\end{aligned}$$

Verify these formulas!!

The functional spaces $\mathcal{C}^k(I)$ and $\mathcal{C}^\infty(I)$

Given an interval $I \subseteq \mathbb{R}$, we can define the set $\mathcal{C}^0(I)$ of all functions that are continuous on I:

$$\mathcal{C}^0(I) = \{f : \mathbb{R} \to \mathbb{R} \ \mid \ f \text{ is continuous on } I\}$$

and the set $C^k(I)$ of all functions that are k times differentiable ($k \in \mathbb{N}$) on I, with the k^{th} derivative being continuous in I:

$$\mathcal{C}^k(I) = \left\{ f : \mathbb{R} \to \mathbb{R} \ \mid \ f^{(j)}(x) \text{ exist } \forall j = 1, \ldots, k, \ \forall x \in I, \text{ and } f^{(k)} \in \mathcal{C}^0(I) \right\}.$$

A function belonging to $\mathcal{C}^k(I)$ is said to be a $\mathcal{C}^k$ **function on** I. We can further define

$$\mathcal{C}^\infty(I) = \left\{ f : \mathbb{R} \to \mathbb{R} \ \mid \ f^{(j)}(x) \text{ exist } \forall j \in \mathbb{N}, \ \forall x \in I \right\}.$$

the set of all infinitely-differentiable functions. A function $f \in \mathcal{C}^\infty(I)$ is said to be a $\mathcal{C}^\infty$ **function on** I.

The sets $\mathcal{C}^k(I)$ and $\mathcal{C}^\infty(I)$ are in fact *functional spaces*, which means that they possess certain useful properties. This is not within the scope of this course, though.

8.9 Convexity and inflection points

Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable at a point $x_0 \in \mathbb{R}$. Then

$$t_{x_0}(x) = f(x_0) + f'(x_0)(x - x_0)$$

is the equation of the line tangent to the graph of f at x_0.

Convexity

The function f is called **convex at** x_0 if there exists a neighohbrood $I_r(x_0)$ of x_0 on which it lies above the tangent at x_0:

$$f(x) \geq t_{x_0}(x), \qquad \forall x \in I_r(x_0).$$

The function f is called **strictly convex at** x_0 if the inequality is strict:

$$f(x) > t_{x_0}(x), \qquad \forall x \in I_r(x_0) \setminus \{x_0\}.$$

Concavity

The function f is called **concave at** x_0 if there exists a neighobrhood $I_r(x_0)$ of x_0 on which it lies below the tangent at x_0:

$$f(x) \leq t_{x_0}(x), \qquad \forall x \in I_r(x_0).$$

The function f is called **strictly concave at** x_0 if the inequality is strict:

$$f(x) < t_{x_0}(x), \qquad \forall x \in I_r(x_0) \setminus \{x_0\}.$$

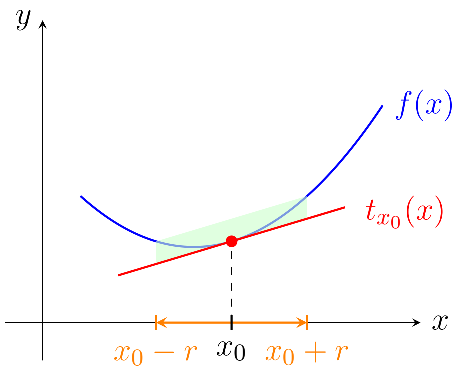


Figure 8.4: Illustration of a convex function at x_0: $f(x) \geq t_{x_0}(x)$ for all x in some neighborhood $I_r(x_0)$ of x_0 ($r > 0$ is small). The green shaded region shows where the function lies above its tangent.

Example 8.24: Let us show that $f(x) = x^2$ is strictly convex at $x_0 = 1$. The tangent line to the graph is

$$t_1(x) = 1 + 2(x - 1) = 2x - 1.$$

Comparing to f we find

$$f(x) - t_1(x) = x^2 - (2x - 1) = x^2 - 2x + 1 = (x - 1)^2 \geq 0$$

with an equality if and only if $x = 1$. Hence f is *strictly* convex at $x_0 = 1$.

Convexity and Concavity on an Interval

A differentiable function f is called **convex on** I if it is convex at all points in I. Similarly, f is called **concave on** I if it is concave at all points in I.

Theorem 8.17: Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable on an interval $I \subseteq \mathbb{R}$. Then f is convex on I if and only if for every $x_0 \in I$,

$$f(x) \geq t_{x_0}(x) = f(x_0) + f'(x_0)(x - x_0) \qquad \forall x \in I.$$

Similarly, f is concave on I if and only if for every $x_0 \in I$,

$$f(x) \leq t_{x_0}(x) = f(x_0) + f'(x_0)(x - x_0) \qquad \forall x \in I.$$

Remark: Observe that the definition of convexity/concavity is a definition at a single point x_0 (see Figure 8.4). Then we said that f is convex/concave on I it is it convex/-concave at all $x_0 \in I$. However the property remained a property of the point x_0. The novelty of this theorem is that the property is now a property of the entire interval I, not just a single point x_0. To understand this visually, let's focus on the convex case. A function is convex at x_0 if *near* x_0 its graph lies above the tangent t_{x_0} (see Figure 8.5).

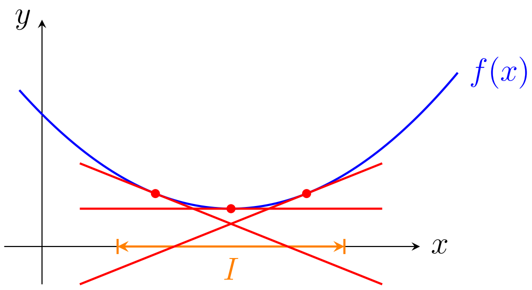


Figure 8.5: A convex function on the interval $I = [a, b]$: $f(x)$ lies above all its tangent lines within the interval, not just at a single point.

Proof of Theorem 8.17. We prove for the convex case. The concave case follows a similar argument.

(Direction $\Leftarrow$) This direction is immediate by definition of convexity.

(Direction $\Rightarrow$) We assume that f is convex on I, and want to prove that for every $x_0 \in I$, $f(x) \geq t_{x_0}(x)$ for all $x \in I$. Equivalently, define for each x_0 the function

$$g_{x_0}(x) = f(x) - t_{x_0}(x).$$

Then we want to show that $g_{x_0}(x) \geq 0$ for all $x \in I$. Observe that

$$g_{x_0}(x_0) = f(x_0) - t_{x_0}(x_0) = 0$$

and

$$g'_{x_0}(x_0) = f'(x_0) - t'_{x_0}(x_0) = 0.$$

For simplicity, assume that x_0 is not a boundary point of I.

Since f is convex at x_0, we know that there exists $r > 0$ such that $g_{x_0}(x) \geq 0$ for all $x \in I_r(x_0)$, where $I_r \subseteq I$. Let us show that $g_{x_0}(x) \geq 0$ to the *right* of x_0. The proof for the *left* of x_0 is similar.

To the right of x_0, we know that for all $x \in [x_0, x_0 + r)$, $g_{x_0}(x) \geq 0$. If this inequality can be extended to any $x \in I \cap \{x > x_0\}$ then we're done. Otherwise, define

$$P = \{x > x_0 \ \mid \ g_{x_0}(y) \geq 0, \ \forall y \in [x_0, x]\}$$

to be the set of points x to the right of x_0 such that g_{x_0} is non-negative on $[x_0, x]$. The maximal such interval $[x_0, x_1]$ has a right endpoint given by

$$x_1 = \sup P.$$

By contradiction, assume that x_1 is an internal point of I.

By definition of the supremum, immediately to the right of x_1 there are points where g_{x_0} is negative. By the continuity of g_{x_0}, it must hold that

$$g_{x_0}(x_1) = 0.$$

Claim: $g_{x_0}(x) = 0$ for all $x \in [x_0, x_1]$.

Proof of claim. **By contradiction**. If the claim is not true, then since g_{x_0} is non-negative, it must hold that $M = \max_{x\in[x_0,x_1]} g_{x_0}(x)$ is strictly positive: $M > 0$. By Weierstrass' Theorem, the maximum of a continuous function on a closed interval is attained, so that there exists $\overline{x} \in (x_0, x_1)$ such that

$$g_{x_0}(\overline{x}) = M.$$

By Fermat's Theorem (Theorem 8.9),

$$g'_{x_0}(\overline{x}) = 0$$

since $\overline{x}$ is an extremum.

Now recall that g_{x_0} is a convex function. Since $g_{x_0}(\overline{x}) = M$ and $g'_{x_0}(\overline{x}) = 0$, convexity implies that $g_{x_0}(x) \geq M$ on a neighborhood of $\overline{x}$. Since M is the maximum of g_{x_0}, it must hold that $g_{x_0}(x) = M$ on that neighborhood. How far to the right can this neighborhood extend? Define

$$Q = \{x > \overline{x} \ \mid \ g_{x_0}(y) = M, \ \forall y \in [\overline{x}, x]\}$$

and

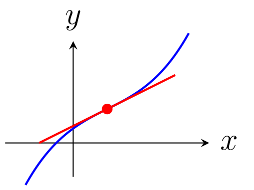
$$x_2 = \max Q$$

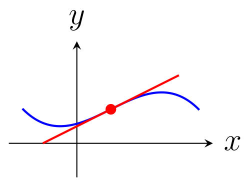
(the maximum is attained since g_{x_0} is a continuous function). Observe that $x_2 < x_1$ since $g_{x_0}(x_1) = 0$. So we have found a point $x_2 \in (x_0, x_1)$ where the function g_{x_0} attains the value M, and, moreover, g_{x_0} attains the value M in a left-neighborhood of x_2. Hence the left-derivative of g_{x_0} at x_2 must be 0. Since g_{x_0} is differentiable at x_2, it must hold that $g'_{x_0}(x_2) = 0$. But then due to convexity at x_2 it must hold that $g_{x_0} \geq 0$ in a neighborhood of x_2, and in particular in a right-neighborhood. But this contradicts the definition of x_2 as $\max Q$. Therefore the claim is proved. □

We can now conclude the proof of the theorem. The claim implies that $g_{x_0}(x) = 0$ for all $x \in [x_0, x_1]$. But then, as in the preceding argument, the left-derivative of g_{x_0} at x_1 is 0. Consequently, (by convexity at x_1) it must hold that $g_{x_0} \geq 0$ to the right of x_1, in contradiction to the definition of x_1 as $\sup P$ and the assumption that x_1 is an internal point of I. Therefore it must hold that $g_{x_0} \geq 0$ on I. □

Inflection point

An inflection point is a point x_0 where the graph of the function lies above (or at) the tangent line t_{x_0} on one side, and below (or at) on the other side.





Theorem 8.18: Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable on $I \subseteq \mathbb{R}$. Then

$$f \text{ is convex on } I \qquad \Leftrightarrow \qquad f' \text{ is increasing on } I,$$

and

$$f \text{ is strictly convex on } I \qquad \Leftarrow \qquad f' \text{ is strictly increasing on } I.$$

Proof. We skip the proof. □

Corollary 8.19: Let $f : \mathbb{R} \to \mathbb{R}$ be twice-differentiable on $I \subseteq \mathbb{R}$. Then

$$f \text{ is convex on } I \qquad \Leftrightarrow \qquad f''(x) \geq 0,\ \forall x \in I,$$

and

$$f \text{ is strictly convex on } I \qquad \Leftarrow \qquad f''(x) > 0,\ \forall x \in I.$$

Proof. The corollary follows immediately from Theorem 8.18 by applying Theorem 8.15 to f'. □

Corollary 8.20: Let f be a twice-differentiable function in a neighborhood of x_0. Then:

$$x_0 \text{ is an inflection point} \qquad \Rightarrow \qquad f''(x_0) = 0,$$

and

$$f''(x_0) = 0 \text{ and } f'' \text{ changes sign at } x_0 \qquad \Rightarrow \qquad x_0 \text{ is an inflection point.}$$

Moreover,

$$f''(x_0) = 0 \text{ and } f'' \text{ doesn't change sign at } x_0 \qquad \Rightarrow \qquad x_0 \text{ isn't an inflection point.}$$

Proof. We still do not have all the tools to prove this. □

Remark: The preceding results can be stated for concave functions as well.

8.10 Qualitative study of a function

Qualitative study of a function f

STEP 1. Understand the domain of f and possible symmetries (such as the function being even or odd).

STEP 2. What happens at the end-points of the domain? Any asymptotes? Discontinuities?

STEP 3. Are there points where f is not differentiable?

STEP 4. Compute f' and understand its domain, where it vanishes and where it is positive/negative. This determines monotonicity intervals and (some) extrema of f.

STEP 5. Find inflection points and determine intervals of convexity/concavity using the tools we learned for f', or by differentiating again (if possible).

8.11 De l'Hôpital's Theorem

De l'Hôpital's Theorem, which we state below, is a very important theorem enabling us to evaluate some indeterminate forms of the form $\frac{0}{0}$ or $\frac{\infty}{\infty}$. It also has an interesting history attached to it. It is named after Guillaume de l'Hôpital, a 17th century French mathematician, who appears to have been a decent mathematician, researching analysis and geometry. However, he was a contemporary of the Bernoulli brothers (Johann and Jacob) who were (and still are) considered to be exceptional mathematicians. De l'Hôpital, in addition to being a mathematician, also came from nobility. This meant that he had money. Enough money to offer Johann Bernoulli annual payments in return for his mathematical results. In short, the following theorem may be called "De l'Hôpital's Theorem", but it is, in fact, Johann Bernoulli's Theorem.

De l'Hôpital's Theorem

Let $f, g : \mathbb{R} \to \mathbb{R}$ and let $x_0 \in \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$. Suppose that f, g are differentiable in a neighborhood of x_0 (possibly excluding x_0 itself) and their limits as $x \to x_0$ exist and are equal:

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = L$$

where

$$L \text{ is either } 0 \text{ or } +\infty \text{ or } -\infty.$$

Then

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}$$

if the limit on the right hand side exists.

Proof. We prove only for the case $L = 0$, $x_0 \in \mathbb{R}$, and for the right-limit at x_0. In this case, even if f, g are not defined at x_0, we can define $f(x_0) = g(x_0) = L = 0$ to make the two functions continuous at x_0. Let $x > x_0$ be a nearby point to the right of x_0. Then since $f(x_0) = g(x_0) = 0$ we have

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)}.$$

By Cauchy's Theorem (Theorem 8.12) there exists a point $t \in (x_0, x)$, depending on x (so we write $t(x)$) such that

$$\frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f'(t(x))}{g'(t(x))}.$$

Now, letting $x \to x_0$, the point $t(x)$ will satisfy $\lim_{x \to x_0} t(x) = x_0$ by the Squeeze Theorem. Hence, using the Substitution Theorem (Theorem 5.12) we can replace the limit

$x \to x_0$ with $t \to x_0$ and obtain:

$$\begin{aligned}\lim_{x \to x_0} \frac{f(x)}{g(x)} &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{g(x) - g(x_0)} \\ &= \lim_{x \to x_0} \frac{f'(t(x))}{g'(t(x))} \\ &= \lim_{t \to x_0} \frac{f'(t)}{g'(t)}\end{aligned}$$

and the proof (in this case) is complete. We skip the proofs of the other cases. □

Example 8.25: Compute

$$\lim_{x \to 0} \frac{e^{2x} - e^{-2x}}{\sin 5x}.$$

We see that we obtain a limit of the form $\frac{0}{0}$. Using De l'Hôpital:

$$\lim_{x \to 0} \frac{e^{2x} - e^{-2x}}{\sin 5x} = \lim_{x \to 0} \frac{\frac{\mathrm{d}}{\mathrm{d}x}(e^{2x} - e^{-2x})}{\frac{\mathrm{d}}{\mathrm{d}x}(\sin 5x)} = \frac{2}{5} \lim_{x \to 0} \frac{e^{2x} + e^{-2x}}{\cos 5x} = \frac{4}{5}.$$

Example 8.26: Compute

$$\lim_{x \to 0} \frac{1 + 3x - (1 + 2x)^{3/2}}{x \sin x}.$$

This has the indeterminate form $\frac{0}{0}$. Differentiating numerator and denominator, we consider the limit:

$$\lim_{x \to 0} \frac{3 - \frac{3}{2} \cdot 2 \cdot (1 + 2x)^{1/2}}{\sin x + x \cos x}.$$

This is still indeterminate of the form $\frac{0}{0}$. So we differentiate *again*:

$$\lim_{x \to 0} \frac{-\frac{3}{2} \cdot 2 \cdot \frac{1}{2} \cdot 2 \cdot (1 + 2x)^{-1/2}}{\cos x + \cos x - x \sin x} = -\frac{3}{2}.$$

Hence:

$$\lim_{x \to 0} \frac{1 + 3x - (1 + 2x)^{3/2}}{x \sin x} = -\frac{3}{2}.$$

Example 8.27: We can now prove that $x^n = o(e^x)$ as $x \to +\infty$. First we write

$$\lim_{x \to +\infty} \frac{x^n}{e^x}$$

which is of the indeterminate form $\frac{+\infty}{+\infty}$. Differentiating over and over again, we have:

$$\begin{aligned}\lim_{x \to +\infty} \frac{x^n}{e^x} &= \lim_{x \to +\infty} \frac{n x^{n-1}}{e^x} \\ &= \lim_{x \to +\infty} \frac{n(n-1) x^{n-2}}{e^x} \\ &= \lim_{x \to +\infty} \frac{n(n-1)(n-2) x^{n-3}}{e^x} \\ &= \cdots \\ \textit{after } n \textit{ derivatives} &= \lim_{x \to +\infty} \frac{n!}{e^x} = 0.\end{aligned}$$

Example 8.28: We can also show that $\ln x = o(x^\alpha)$ for any $\alpha > 0$ as $x \to +\infty$:

$$\lim_{x\to+\infty} \frac{\ln x}{x^\alpha} = \lim_{x\to+\infty} \frac{\frac{1}{x}}{\alpha x^{\alpha-1}} = \frac{1}{\alpha} \lim_{x\to+\infty} \frac{1}{x^\alpha} = 0$$

Example 8.29: The function $f(x) = e^x - 1 - \sin x$ is infinitesimal as $x \to 0$. What is its order with respect to $\varphi(x) = x$ and what is its principal part?

We consider the limit, for $\alpha > 0$ to be determined:

$$\lim_{x\to 0} \frac{f(x)}{x^\alpha} = \lim_{x\to 0} \frac{e^x - 1 - \sin x}{x^\alpha}$$

which is indeterminate of the form $\frac{0}{0}$. Differentiating we have

$$\lim_{x\to 0} \frac{e^x - \cos x}{\alpha x^{\alpha-1}}$$

which is still of the form $\frac{0}{0}$. So we differentiate again:

$$\lim_{x\to 0} \frac{e^x + \sin x}{\alpha(\alpha-1)x^{\alpha-2}}.$$

Now the numerator tends to 1 as $x \to 0$, and if we choose $\alpha = 2$ then the denominator is 2. So we have:

$$\lim_{x\to 0} \frac{f(x)}{x^2} = \frac{1}{2}$$

which means that $f(x)$ is infinitesimal of order 2 with respect to $\varphi(x) = x$ at $x = 0$, and the principal part is $p(x) = \frac{1}{2}x^2$.

Chapter 9

Taylor expansions and applications

Our goal in this chapter is to approximate complicated functions using polynomials, which are easier to study. This is a powerful tool, with many applications, not least in engineering. Moreover, it is the first example of an approximation of a function using a well-known set of functions (polynomials in this case). Later on (not in this course), you will see examples of approximation using sines and cosines, as well as other sets of approximating functions.

9.1 Taylor formulas

Consider a function $f : \mathbb{R} \to \mathbb{R}$, and suppose that we can differentiate it at a given point x_0 as many times as we'd like. Then consider the following sequence of polynomials with increasing order, starting from 0th order:

0^{TH} **order polynomial.** Define

$$(Tf)_{0,x_0}(x) = f(x_0)$$

to simply be the constant function with value x_0 for all $x \in \mathbb{R}$. Then, obviously,

$$f(x_0) = (Tf)_{0,x_0}(x_0).$$

1^{ST} **order polynomial.** Define

$$(Tf)_{1,x_0}(x) = f(x_0) + f'(x_0)(x - x_0)$$

to be the affine function that passes through the point $(x_0, f(x_0))$ and has slope $f'(x_0)$. Then, since

$$(Tf)'_{1,x_0}(x) = f'(x_0), \quad \forall x \in \mathbb{R},$$

we have

$$\begin{aligned} f(x_0) &= (Tf)_{1,x_0}(x_0), \\ f'(x_0) &= (Tf)'_{1,x_0}(x_0). \end{aligned}$$

2^{ND} **order polynomial.** Define

$$(Tf)_{2,x_0}(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2$$

to be a parabola. Let's check its properties. Differentiating twice we have:

$$
\begin{aligned}
(Tf)'_{2,x_0}(x) &= f'(x_0) + f''(x_0)(x - x_0) \\
(Tf)''_{2,x_0}(x) &= f''(x_0)
\end{aligned}
$$

Thus, plugging $x = x_0$ into $(Tf)_{2,x_0}$ and its first two derivatives, we have

$$
\begin{aligned}
(Tf)_{2,x_0}(x_0) &= f(x_0) \\
(Tf)'_{2,x_0}(x_0) &= f'(x_0) \\
(Tf)''_{2,x_0}(x_0) &= f''(x_0).
\end{aligned}
$$

Hence $(Tf)_{2,x_0}(x)$ is a second-order polynomial (parabola) satisfying that: (a) it coincides with f at x_0, (b) its derivative coincides with the derivative of f at x_0, and (c) its *second* derivative coincides with the second derivative of f at x_0.

3^RD order polynomial. Defining

$$
(Tf)_{3,x_0}(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \frac{1}{3 \cdot 2}f'''(x_0)(x - x_0)^3
$$

we have

$$
\begin{aligned}
(Tf)'_{3,x_0}(x) &= f'(x_0) + f''(x_0)(x - x_0) + \frac{1}{2}f'''(x_0)(x - x_0)^2 \\
(Tf)''_{3,x_0}(x) &= f''(x_0) + f'''(x_0)(x - x_0) \\
(Tf)'''_{3,x_0}(x) &= f'''(x_0)
\end{aligned}
$$

so that

$$
\begin{aligned}
(Tf)_{3,x_0}(x_0) &= f(x_0) \\
(Tf)'_{3,x_0}(x_0) &= f'(x_0) \\
(Tf)''_{3,x_0}(x_0) &= f''(x_0) \\
(Tf)'''_{3,x_0}(x_0) &= f'''(x_0).
\end{aligned}
$$

$n^{\mathbf{TH}}$ order polynomial. Following the previous ideas, we can define

$$
\begin{aligned}
(Tf)_{n,x_0}(x) = &f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \frac{1}{3 \cdot 2}f'''(x_0)(x - x_0)^3 + \cdots \\
&+ \frac{1}{n(n-1)(n-2)\cdots 3 \cdot 2}f^{(n)}(x_0)(x - x_0)^n \\
= &\sum_{k=0}^{n} \frac{1}{k!}f^{(k)}(x_0)(x - x_0)^k
\end{aligned}
$$

to find that

$$
\begin{aligned}
(Tf)_{n,x_0}(x_0) &= f(x_0) \\
(Tf)'_{n,x_0}(x_0) &= f'(x_0) \\
(Tf)''_{n,x_0}(x_0) &= f''(x_0) \\
&\vdots \\
(Tf)^{(n)}_{n,x_0}(x_0) &= f^{(n)}(x_0).
\end{aligned}
$$

Taylor and Maclaurin polynomials

Given a function $f : \mathbb{R} \to \mathbb{R}$ that is n-times differentiable at a point x_0, its **Taylor polynomial (expansion) of order (degree)** n **at** x_0 is the polynomial

$$
(Tf)_{n,x_0}(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k.
$$

When the point $x_0 = 0$ is the origin, this polynomial is sometimes called the **Maclaurin polynomial (expansion)** of f.

Lemma 9.1: The derivative of the Taylor polynomial of f is the Taylor polynomial of the derivative f' of one lesser order:

$$
(Tf)'_{n,x_0}(x) = (Tf')_{n-1,x_0}(x).
$$

Proof. We compute

$$
\begin{aligned}
(Tf)'_{n,x_0}(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k\right) \\
&= \sum_{k=1}^{n} \frac{k}{k!} f^{(k)}(x_0)(x - x_0)^{k-1} \\
&= \sum_{k=1}^{n} \frac{1}{(k-1)!} f^{(k)}(x_0)(x - x_0)^{k-1} \\
(\textit{substitute } j = k-1) \qquad &= \sum_{j=0}^{n-1} \frac{1}{j!} f^{(j+1)}(x_0)(x - x_0)^j \\
&= \sum_{j=0}^{n-1} \frac{1}{j!} (f')^{(j)}(x_0)(x - x_0)^j \\
&= (Tf')_{n-1,x_0}(x)
\end{aligned}
$$

and the proof is complete. □

We may take additional derivatives, to arrive at the formula, true for any $k = 0, \dots, n$:

$$(Tf)^{(k)}_{n,x_0}(x) = (Tf^{(k)})_{n-k,x_0}(x).$$

Therefore:

$$(Tf)^{(k)}_{n,x_0}(x_0) = (Tf^{(k)})_{n-k,x_0}(x_0) = f^{(k)}(x_0)$$

since the value of the Taylor polynomial of a function at x_0 equals the value of that function at x_0.

Theorem 9.2 (Peano's remainder)**:** For a function $f : \mathbb{R} \to \mathbb{R}$ that is n-times differentiable at a point x_0, the difference between f and its Taylor polynomial at a point x close to x_0 is of order $o((x - x_0)^n)$:

$$f(x) - (Tf)_{n,x_0}(x) = o((x - x_0)^n), \qquad x \to x_0.$$

This is called **Peano's approximation of the remainder**.

Proof. We must show that

$$\lim_{x \to x_0} \frac{f(x) - (Tf)_{n,x_0}(x)}{(x - x_0)^n} = 0.$$

This is an expression of the form $\frac{0}{0}$. Applying de l'Hôpital's Theorem we have

$$\begin{aligned} \lim_{x \to x_0} \frac{f(x) - (Tf)_{n,x_0}(x)}{(x - x_0)^n} &= \lim_{x \to x_0} \frac{f'(x) - (Tf)'_{n,x_0}(x)}{n(x - x_0)^{n-1}} \\ \textit{(using Lemma 9.1)} \quad &= \lim_{x \to x_0} \frac{f'(x) - (Tf')_{n-1,x_0}(x)}{n(x - x_0)^{n-1}} \end{aligned}$$

which is still of the form $\frac{0}{0}$ (if $n \geq 2$). Applying de l'Hôpital's Theorem and Lemma 9.1 $n-1$ times we eventually arrive at

$$\begin{aligned} \lim_{x \to x_0} \frac{f(x) - (Tf)_{n,x_0}(x)}{(x - x_0)^n} &= \lim_{x \to x_0} \frac{f^{(n-1)}(x) - (Tf^{(n-1)})_{1,x_0}(x)}{n!(x - x_0)} \\ &= \frac{1}{n!} \lim_{x \to x_0} \frac{f^{(n-1)}(x) - f^{(n-1)}(x_0) - f^{(n)}(x_0)(x - x_0)}{x - x_0} \\ &= \frac{1}{n!} \lim_{x \to x_0} \left(\frac{f^{(n-1)}(x) - f^{(n-1)}(x_0)}{x - x_0} - f^{(n)}(x_0) \right) = 0 \end{aligned}$$

by the definition of the nth derivative at x_0. This justifies having applied de l'Hôpital's Theorem repeatedly, and proves the theorem. □

Theorem 9.3 (Lagrange's remainder)**:** Let $f : \mathbb{R} \to \mathbb{R}$ be a function that is n-times differentiable at x_0 and $n+1$ times differentiable in a neighborhood $I_r(x_0) \setminus \{x_0\}$. Then for $x \in I_r(x_0)$ the difference between f and its Taylor polynomial at x can be written as:

$$f(x) - (Tf)_{n,x_0}(x) = \frac{1}{(n+1)!} f^{(n+1)}(\overline{x})(x - x_0)^{n+1},$$

where $\overline{x}$ is a point between x and x_0. This is called **Lagrange's formula for the remainder**.

Proof. Denote the remainder term $\varphi(x)$:

$$\begin{aligned}\varphi(x) &= f(x) - (Tf)_{n,x_0}(x) \\ &= f(x) - \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x-x_0)^k\end{aligned}$$

and define

$$\psi(x) = (x-x_0)^{n+1}.$$

By our construction of the Taylor polynomial, it holds that for all $j = 0, \ldots, n$,

$$\begin{aligned}&\varphi^{(j)}(x_0) = 0, \\ &\psi^{(j)}(x_0) = 0 \qquad \text{and} \qquad \psi^{(j)}(x) \neq 0, \text{ if } x \neq x_0.\end{aligned}$$

Applying Cauchy's Theorem (Theorem 8.12) to φ and ψ on the interval between x and x_0, we have

$$\frac{\varphi(x)}{\psi(x)} = \frac{\varphi(x) - \varphi(x_0)}{\psi(x) - \psi(x_0)} = \frac{\varphi'(x_1)}{\psi'(x_1)}$$

where x_1 is some point between x and x_0. Repeating this argument with the derivatives (considering now the interval between x_0 and x_1), we have

$$\frac{\varphi'(x_1)}{\psi'(x_1)} = \frac{\varphi'(x_1) - \varphi'(x_0)}{\psi'(x_1) - \psi'(x_0)} = \frac{\varphi''(x_2)}{\psi''(x_2)}$$

where x_2 is some point between x_0 and x_1. Looking now between x_0 and x_2, we repeat the argument with the second derivatives to find

$$\frac{\varphi''(x_2)}{\psi''(x_2)} = \frac{\varphi''(x_2) - \varphi''(x_0)}{\psi''(x_2) - \psi''(x_0)} = \frac{\varphi'''(x_3)}{\psi'''(x_3)}$$

where x_3 is some point between x_0 and x_2. Repeating this argument $n+1$ times, we arrive at

$$\frac{\varphi(x)}{\psi(x)} = \frac{\varphi'(x_1)}{\psi'(x_1)} = \frac{\varphi''(x_2)}{\psi''(x_2)} = \cdots = \frac{\varphi^{(n+1)}(x_{n+1})}{\psi^{(n+1)}(x_{n+1})}$$

where x_{n+1} is a point lying between x_0 and x_n, x_n is a point lying between x_0 and x_{n-1}, and so on... However, observe that

$$\begin{aligned}\varphi^{(n+1)}(x) &= f^{(n+1)}(x) \\ \psi^{(n+1)}(x) &= (n+1)!\end{aligned}$$

So, by defining $\overline{x} = x_{n+1}$ the proof is complete since we have found that

$$\frac{f(x) - (Tf)_{n,x_0}(x)}{(x-x_0)^{n+1}} = \frac{\varphi(x)}{\psi(x)} = \frac{\varphi^{(n+1)}(\overline{x})}{\psi^{(n+1)}(\overline{x})} = \frac{f^{(n+1)}(\overline{x})}{(n+1)!}.$$

□

Remark: Recalling the definition of the Taylor polynomial, and using Lagrange's formula for the remainder, we have

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x-x_0)^k + \frac{1}{(n+1)!} f^{(n+1)}(\overline{x})(x-x_0)^{n+1}.$$

So we see that the function f is nearly equal to the Taylor polynomial, and that the remainder term has a form that is very similar to the $n+1$ terms of the polynomial. The only (important) difference, is that the derivative term within the remainder is evaluated at a different point $\overline{x}$ which lies somewhere between x_0 and x (and we do not know where exactly).

Proposition 9.4: Any Maclaurin polynomial of an even function contains only even powers. Any Maclaurin polynomial of an odd function contains only odd powers.

Proof. This is a simple proof that is left as an exercise. □

9.2 Expanding the elementary functions

We now derive the Taylor expansions of some of the most commonly used functions. Recall that a function f, the Taylor polynomial of order n at x_0 is given by

$$(Tf)_{n,x_0}(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k.$$

Hence, to construct this polynomial for a given function one needs to know the first n derivatives at a given point x_0.

The exponential function

For the function $f(x) = e^x$ we know that $f^{(k)}(x) = e^x$ for all $k \in \mathbb{N}$ and for all $x \in \mathbb{R}$. Plugging this into the expression for the Taylor polynomial, we find that

$$(Tf)_{n,x_0}(x) = \sum_{k=0}^{n} \frac{1}{k!} e^{x_0}(x - x_0)^k.$$

For simplicity, we choose $x_0 = 0$, so that $e^{x_0} = 1$ and we're left with the *Maclaurin expansion*

$$(Tf)_{n,0}(x) = \sum_{k=0}^{n} \frac{x^k}{k!}$$

Hence, with **Peano**'s approximation of the remainder we have that

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + o(x^n), \quad x \to 0.$$

Lagrange's formula for the remainder gives us

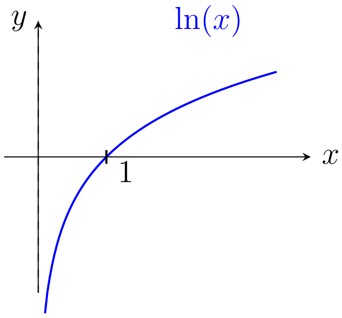
$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{e^{\overline{x}}}{(n+1)!} x^{n+1}$$

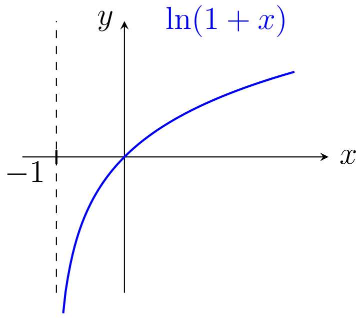
where $\overline{x}$ is some point between 0 and x.

Observe that by taking $x = 1$ we obtain an approximation of the number e. In fact, this approximation of e converges extremely fast. We can bound the error by bounding the error term in Lagrange's formulation.

The logarithm

We now want to write an expansion of the natural logarithm $\ln x$. It is only defined for $x > 0$, so the point x_0 where we perform the expansion must be greater than 0. As we shall see, it will be convenient to choose $x_0 = 1$. However, it will be even more convenient to translate the function and consider $\ln(1 + x)$, so that the vertical asymptote will be at -1; then we will choose to expand around 0.





Let's start with $f(x) = \ln x$. We can write the first few derivatives:

$$f'(x) = \frac{1}{x} \qquad f''(x) = -\frac{1}{x^2} \qquad f'''(x) = \frac{2}{x^3} \qquad f^{(4)}(x) = -\frac{3 \cdot 2}{x^4} \qquad \cdots$$

and the general formula is:

$$f^{(k)}(x) = (-1)^{k+1}\frac{(k-1)!}{x^k}, \qquad \forall k \in \mathbb{N}_+.$$

We therefore find that

$$\begin{aligned}(Tf)_{n,x_0}(x) &= \ln x_0 + \sum_{k=1}^{n} \frac{1}{k!}(-1)^{k-1}\frac{(k-1)!}{x_0^k}(x - x_0)^k \\ &= \ln x_0 + \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k}\frac{(x - x_0)^k}{x_0^k}\end{aligned}$$

For simplicity, we take $x_0 = 1$, so that $\ln x_0 = 0$, and obtain

$$(Tf)_{n,1}(x) = \sum_{k=1}^{n}(-1)^{k-1}\frac{(x-1)^k}{k}$$

$$
\begin{aligned}
\ln x &= (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots + (-1)^{n-1}\frac{(x-1)^n}{n} + o((x-1)^n) \\
&= \sum_{k=1}^{n} (-1)^{k-1}\frac{(x-1)^k}{k} + o((x-1)^n), \qquad x \to 0.
\end{aligned}
$$

Now let's try $g(x) = \ln(1+x)$. Then

$$
g'(x) = \frac{1}{1+x} \qquad g''(x) = -\frac{1}{(1+x)^2} \qquad g'''(x) = \frac{2}{(1+x)^3} \qquad \cdots
$$

and the general formula is:

$$
g^{(k)}(x) = (-1)^{k+1}\frac{(k-1)!}{(1+x)^k} \qquad \forall k \in \mathbb{N}_+.
$$

We therefore find that

$$
\begin{aligned}
(Tg)_{n,x_0}(x) &= \ln(1+x_0) + \sum_{k=1}^{n} \frac{1}{k!}(-1)^{k-1}\frac{(k-1)!}{(1+x_0)^k}(x-x_0)^k \\
&= \ln(1+x_0) + \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k}\frac{(x-x_0)^k}{(1+x_0)^k}
\end{aligned}
$$

Taking $x_0 = 0$, so that $g(x_0) = 0$, we obtain

$$
(Tg)_{n,0}(x) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} x^k.
$$

Hence we find that the nth degree Taylor polynomial for $\ln(1+x)$ around the point $x_0 = 0$ (hence it is also a Maclaurin polynomial) is:

$$
x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - + \cdots
$$

$$
\begin{aligned}
\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \cdots + (-1)^{n-1}\frac{x^n}{n} + o(x^n) \\
&= \sum_{k=1}^{n} (-1)^{k-1}\frac{x^k}{k} + o(x^n), \qquad x \to 0.
\end{aligned}
$$

The trigonometric functions $\sin x$ and $\cos x$

We know that $\sin x$ is an odd function, so by Proposition 9.4 we expect its Maclaurin polynomial to only have odd powers. Taking the first few derivatives, we have

$$\sin' x = \cos x \quad \sin'' x = -\sin x \quad \sin''' x = -\cos x \quad \sin^{(4)} x = \sin x \quad \cdots$$

and at $x_0 = 0$ we get

$$\sin' 0 = 1 \quad \sin'' 0 = 0 \quad \sin''' 0 = -1 \quad \sin^{(4)} 0 = 0 \quad \cdots$$

With **Peano**'s approximation of the remainder we have that with $n = 2m+2$

$$\begin{aligned}\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} - \cdots + (-1)^m \frac{x^{2m+1}}{(2m+1)!} + o(x^{2m+2}) \\ &= \sum_{k=1}^{m} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + o(x^{2m+2}), \qquad x \to 0.\end{aligned}$$

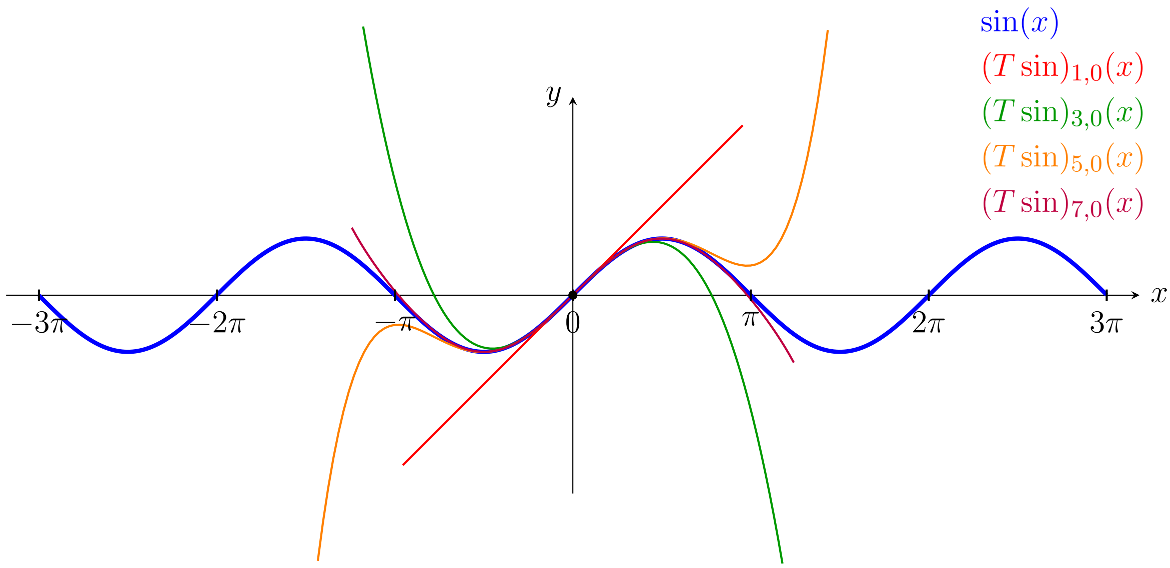


Figure 9.1: The function $\sin(x)$ (blue) and its Maclaurin polynomials $(T\sin)_{n,0}(x)$ of orders $n = 2m+1 = 1, 3, 5, 7$. Higher order polynomials provide better approximations.

We know that $\cos x$ is an even function, so by Proposition 9.4 we expect its Maclaurin polynomial to only have even powers. Taking the first few derivatives, we have

$$\cos' x = -\sin x \quad \cos'' x = -\cos x \quad \cos''' x = \sin x \quad \cos^{(4)} x = \cos x \quad \cdots$$

and at $x_0 = 0$ we get

$$\cos' 0 = 0 \quad \cos'' 0 = -1 \quad \cos''' 0 = 0 \quad \cos^{(4)} 0 = 1 \quad \cdots$$

With **Peano**'s approximation of the remainder we have that with $n = 2m + 1$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} - \cdots + (-1)^m \frac{x^{2m}}{(2m)!} + o(x^{2m+1})$$
$$= \sum_{k=0}^{m} (-1)^k \frac{x^{2k}}{(2k)!} + o(x^{2m+1}), \qquad x \to 0.$$

Power functions

We would want to study functions of the form x^α for $\alpha \in \mathbb{R}$. However, since power functions of this form are defined only for $x > 0$, we translate the function by 1 and then study the Maclaurin polynomials (just like we did for the natural logarithm). We therefore define

$$f(x) = (1 + x)^\alpha$$

so that

$$f'(x) = \alpha(1+x)^{\alpha-1} \quad f''(x) = \alpha(\alpha-1)(1+x)^{\alpha-2} \quad f'''(x) = \alpha(\alpha-1)(\alpha-2)(1+x)^{\alpha-3} \quad \cdots$$

and we get

$$f'(0) = \alpha \quad f''(0) = \alpha(\alpha - 1) \quad f'''(0) = \alpha(\alpha - 1)(\alpha - 2) \quad \cdots$$

Thus, the coefficients that will appear in the Maclaurin polynomial will be:

$$f(0) = 1 \qquad \text{and} \qquad \frac{f^{(k)}(0)}{k!} = \frac{\alpha(\alpha - 1)(\alpha - 2) \cdots (\alpha - k + 1)}{k!} \quad \text{for } k \geq 1.$$

To abbreviate these expressions, we use their similarity to the binomial coefficients, to define:

$$\binom{\alpha}{0} = 1 \qquad \text{and} \qquad \binom{\alpha}{k} = \frac{\alpha(\alpha - 1)(\alpha - 2) \cdots (\alpha - k + 1)}{k!} \quad \text{for } k \geq 1.$$

Therefore:

$$(1 + x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha - 1)}{2!} x^2 + \cdots + \binom{\alpha}{n} x^n + o(x^n)$$
$$= \sum_{k=0}^{n} \binom{\alpha}{k} x^k + o(x^n)$$

Remark: If $\alpha \in \mathbb{N}$ is an integer, then $(1 + x)^\alpha$ can be expanded as a polynomial using the standard binomial formula. In this case, the Maclaurin polynomials will coincide with this polynomial if their order is at least α, otherwise they will only contain the first α terms of the polynomial. For example, consider

$$f(x) = (1 + x)^5 = 1 + 5x + 10x^2 + 10x^3 + 5x^4 + x^5.$$

Then it expansion of order 3 is

$$(1+x)^5 = \underbrace{1+5x+10x^2+10x^3}_{(Tf)_{3,0}(x)} + o(x^3)$$

but its expansion of any order ≥ 5 is identical to f itself:

$$(Tf)_{n,0}(x) = 1+5x+10x^2+10x^3+5x^4+x^5 = f(x) \qquad \forall n \geq 5.$$

Remark: If $\alpha \notin \mathbb{N}$, then, there will be nontrivial Maclaurin polynomials of arbitrary order (i.e. there will be arbitrarily high powers of x that won't vanish). This is in contrast to the case $\alpha \in \mathbb{N}$.

Let us highlight some special cases:

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 - \cdots + (-1)^n x^n + o(x^n) = \sum_{k=0}^{n} (-1)^k x^k + o(x^n)$$

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 + o(x^3)$$

We stop at x^3, because the next terms aren't as nice: the coefficient of x^4 is $-\frac{5}{128}$ and the coefficient of x^5 is $\frac{7}{256}$.

9.3 Operations on Taylor expansions

Uniqueness of the expansion

Sometimes we may be able to identify a polynomial that approximates a function f by other means, and then we may wonder what is its relationship to the function's Taylor or Maclaurin expansions. The following proposition answers this question, but telling us that an expansion of a given order is unique. Therefore, if we've found one (no matter how), we're done!

Proposition 9.5 (Uniqueness of the Taylor expansion): Let $f : \mathbb{R} \to \mathbb{R}$ be n times differentiable at x_0. Then $(Tf)_{n,x_0}(x)$ is the only polynomial of degree $\leq n$ satisfying

$$f(x) - (Tf)_{n,x_0}(x) = o((x-x_0)^n) \qquad \text{as } x \to x_0.$$

Proof. By contradiction, suppose we can write

$$f(x) - P_n(x) = o((x-x_0)^n) \qquad \text{as } x \to x_0$$

where $P_n(x)$ is another polynomial of degree $\leq n$. We'll show that $P_n(x) = (Tf)_{n,x_0}(x)$. On the one hand, subtracting the above two expressions, we have

$$P_n(x) - (Tf)_{n,x_0}(x) = o((x - x_0)^n) \qquad \text{as } x \to x_0.$$

On the other hand, since both $P_n(x)$ and $(Tf)_{n,x_0}(x)$ are both polynomials of degree $\leq n$, it must hold that so is their difference:

$$P_n(x) - (Tf)_{n,x_0}(x) = \sum_{k=0}^{n} c_k (x - x_0)^k.$$

Let m be the smallest index such that $c_m \neq 0$. Then

$$P_n(x) - (Tf)_{n,x_0}(x) = \sum_{k=m}^{n} c_k (x - x_0)^k$$

and we can write

$$\frac{P_n(x) - (Tf)_{n,x_0}(x)}{(x - x_0)^m} = c_m + \sum_{k=m+1}^{n} c_k (x - x_0)^{k-m}.$$

Left hand side: Using the fact that $n \geq m$, we have

$$\lim_{x\to x_0} \frac{P_n(x) - (Tf)_{n,x_0}(x)}{(x - x_0)^m} = \lim_{x\to x_0} \frac{P_n(x) - (Tf)_{n,x_0}(x)}{(x - x_0)^n} (x - x_0)^{n-m} = 0$$

since $P_n(x) - (Tf)_{n,x_0}(x) = o((x - x_0)^n)$.

Right hand side: It obviously holds that

$$\lim_{x\to x_0} \left(c_m + \sum_{k=m+1}^{n} c_k (x - x_0)^{k-m} \right) = c_m + \lim_{x\to x_0} \left(\sum_{k=m+1}^{n} c_k (x - x_0)^{k-m} \right) = c_m.$$

This implies that $c_m = 0$, a contradiction. $\square$

Example 9.1: Suppose that we know that a function $f(x)$ satisfies:

$$f(x) = 9 - 3(x-2) + (x-2)^2 - \frac{1}{4}(x-2)^3 + o((x-2)^3) \qquad \text{as } x \to 2.$$

Then by the uniqueness of the Taylor polynomial, we know that

$$(Tf)_{n=3,x_0=2}(x) = 9 - 3(x-2) + (x-2)^2 - \frac{1}{4}(x-2)^3$$

implying that

$$f(2) = 9 \qquad f'(2) = -3 \qquad f''(2) = 2 \qquad f'''(2) = -3 \cdot 2 \cdot \frac{1}{4} = -\frac{3}{2}.$$

For simplicity, in what follows we consider the Maclaurin expansion, i.e. we take $x_0 = 0$. In general, this can always be achieved by translating a function to bring the point x_0 to 0.

We will consider two functions f and g and their Maclaurin expansions:

$$f(x) = \underbrace{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n}_{p_n(x)} + o(x^n)$$
$$g(x) = \underbrace{b_0 + b_1 x + b_2 x^2 + \cdots + b_n x^n}_{q_n(x)} + o(x^n).$$

Sums of expansions

$$\begin{aligned} f(x) \pm g(x) &= [p_n(x) + o(x^n)] \pm [q_n(x) + o(x^n)] \\ &= [p_n(x) \pm q_n(x)] + [o(x^n) \pm o(x^n)] \\ &= p_n(x) \pm q_n(x) + o(x^n). \end{aligned}$$

Example 9.2: We have seen that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + o(x^n).$$

Changing x to $-x$, we obtain

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots + (-1)^n \frac{x^n}{n!} + o(x^n).$$

Let's take n to be an even number given by $n = 2m$ where $m \in \mathbb{N}_+$, then we have

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^{2m}}{(2m)!} + o(x^{2m}) \\ e^{-x} &= 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots + \frac{x^{2m}}{(2m)!} + o(x^{2m}) \end{aligned}$$

where the last term in the Maclaurin polynomial of e^{-x} is positive because it is of even order. Then we have

$$\begin{aligned} e^x + e^{-x} &= 2 + \frac{2x^2}{2!} + \frac{2x^4}{4!} + \frac{2x^6}{6!} + \cdots + \frac{2x^{2m}}{(2m)!} + o(x^{2m}) \\ e^x - e^{-x} &= 2x + \frac{2x^3}{3!} + \frac{2x^5}{5!} + \frac{2x^7}{7!} + \cdots + \frac{2x^{2m-1}}{(2m-1)!} + o(x^{2m}) \end{aligned}$$

Dividing these expressions by 2 we have:

$$\begin{aligned} \frac{e^x + e^{-x}}{2} &= 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots + \frac{x^{2m}}{(2m)!} + o(x^{2m}) \\ \frac{e^x - e^{-x}}{2} &= x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots + \frac{x^{2m-1}}{(2m-1)!} + o(x^{2m}) \end{aligned}$$

These two functions are called the **hyperbolic cosine** and **hyperbolic sine**, respectively, and are denoted

$$\cosh(x) = \frac{e^x + e^{-x}}{2} \qquad \text{and} \qquad \sinh(x) = \frac{e^x - e^{-x}}{2}$$

While the regular sine and cosine satisfy $\sin^2 x + \cos^2 x = 1$, the hyperbolic versions satisfy a hyperbolic version of this relation:

$$\cosh^2 x - \sinh^2 x = 1.$$

These functions have many applications in engineering.

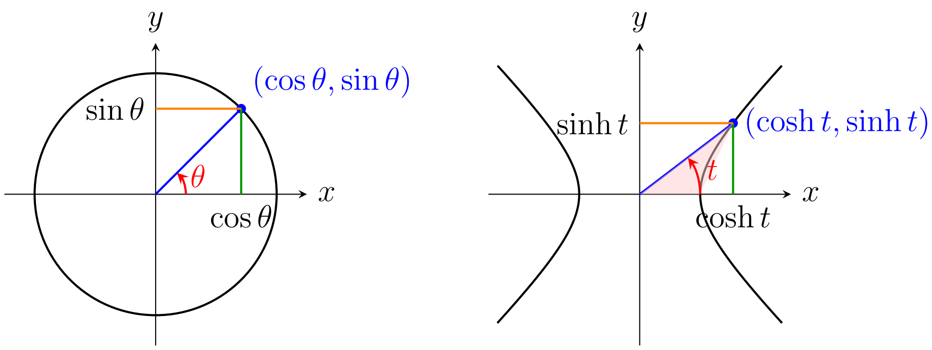


Figure 9.2: Geometric interpretation of trigonometric functions. *Left:* Circular functions $(\cos\theta, \sin\theta)$ on the unit circle $x^2 + y^2 = 1$, where θ is the angle. *Right:* Hyperbolic functions $(\cosh t, \sinh t)$ on the unit hyperbola $x^2 - y^2 = 1$, where t is twice the shaded area.

Example 9.3: Consider the function

$$h(x) = e^x - \sqrt{1+2x}$$

which vanishes as $x \to 0$. We want to understand its order at 0. We have

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + o(x^n) \\ \sqrt{1+2x} &= 1 + x - \frac{x^2}{2} + \frac{x^3}{2} + o(x^3) \end{aligned}$$

Observe that if we stop at order 1 for both expansions there's complete cancellation and we get

$$h(x) = (1 + x + o(x)) - (1 + x + o(x)) = o(x) \quad \text{as } x \to 0.$$

However, taking a further order we have:

$$h(x) = \left(1 + x + \frac{x^2}{2!} + o(x^2)\right) - \left(1 + x - \frac{x^2}{2} + o(x^2)\right) = x^2 + o(x^2) \quad \text{as } x \to 0.$$

Taking another order:

$$\begin{aligned} h(x) &= \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + o(x^3)\right) - \left(1 + x - \frac{x^2}{2} + \frac{x^3}{2} + o(x^3)\right) \\ &= x^2 - \frac{x^3}{3} + o(x^3) \quad \text{as } x \to 0. \end{aligned}$$

Products of expansions

$$\begin{aligned}
f(x)\cdot g(x) &= [p_n(x)+o(x^n)]\cdot[q_n(x)+o(x^n)] \\
&= p_n(x)\cdot q_n(x)+p_n(x)o(x^n)+q_n(x)o(x^n)+o(x^n)o(x^n) \\
&= p_n(x)\cdot q_n(x)+o(x^n)+o(x^n)+o(x^{2n}) \\
&= p_n(x)\cdot q_n(x)+o(x^n).
\end{aligned}$$

Remark: A couple of remarks about the above computation. First, the fact that $p_n(x)o(x^n) = q_n(x)o(x^n) = o(x^n)$ follows from the fact that both functions p_n and q_n are bounded near $x = 0$, and therefore when they multiply something small of order $o(x^n)$ they do not affect its order. Second, we have $o(x^n) + o(x^n) + o(x^{2n}) = o(x^n)$ since the $o(x^{2n})$ is of higher order and therefore is negligible with respect to $o(x^n)$. Then, we can sum as many $o(x^n)$ as we want, and they will still result in an element that is of order $o(x^n)$.

Observe that the product of the two polynomials p_n and q_n will include monomials up to x^{2n} coming from the product of $a_n x^n$ with $b_n x^n$:

$$\begin{aligned}
p_n(x)\cdot q_n(x) &= \left(\sum_{i=0}^{n} a_i x^i\right)\left(\sum_{j=0}^{n} b_j x^j\right) \\
&= \sum_{k=0}^{2n}\left(\sum_{m=0}^{k} a_m b_{k-m}\right)x^k \\
&= \sum_{k=0}^{n}\left(\sum_{m=0}^{k} a_m b_{k-m}\right)x^k + \underbrace{\sum_{k=n+1}^{2n}\left(\sum_{m=0}^{k} a_m b_{k-m}\right)x^k}_{o(x^n)}
\end{aligned}$$

However, since our our order of accuracy is $o(x^n)$ there is no point in maintaining terms that are of that order. Hence we can write:

$$p_n(x)\cdot q_n(x) = \sum_{k=0}^{n}\left(\sum_{m=0}^{k} a_m b_{k-m}\right)x^k + o(x^n)$$

Example 9.4: Expand $h(x) = \sqrt{1+x}\ \cdot e^x$ to second order. To second order, these two functions have the expansions:

$$f(x) = \sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + o(x^2)$$

and

$$g(x) = e^x = 1 + x + \frac{x^2}{2} + o(x^2)$$

$$
\begin{aligned}
f(x)\cdot g(x) &= \left(1+\frac{x}{2}-\frac{x^2}{8}+o(x^2)\right)\left(1+x+\frac{x^2}{2}+o(x^2)\right)\\
&= 1\cdot\left(1+x+\frac{x^2}{2}\right)+\frac{x}{2}\cdot\left(1+x+\frac{x^2}{2}\right)-\frac{x^2}{8}\cdot\left(1+x+\frac{x^2}{2}\right)+o(x^2)\\
&= 1+\left(x+\frac{x}{2}\right)+\left(\frac{x^2}{2}+\frac{x^2}{2}-\frac{x^2}{8}\right)+\underbrace{\left(\frac{x^3}{4}-\frac{x^3}{8}\right)+\left(-\frac{x^4}{16}\right)}_{o(x^2)}+o(x^2)\\
&= 1+\frac{3}{2}x+\frac{7}{8}x^2+o(x^2) \qquad x\to 0.
\end{aligned}
$$

Example 9.5: Let us give an example of a different flavor. Suppose we want to approximate the product $\pi\cdot e$. Here are the expressions for both number to 5 decimal places:

$$\pi = 3.14159 \pm 5\cdot 10^{-6} \qquad e = 2.71828 \pm 5\cdot 10^{-6}$$

Let's compare the product of these two approximation to the actual product of π and e:

$$
\begin{aligned}
3.14159\times 2.71828 &= 8.5397\mathbf{212652}\\
\pi\times e &= 8.5397\mathbf{342226}\cdots
\end{aligned}
$$

We see that they already disagree starting from the 5th decimal place. This demonstrates that when multiplying approximations we need to exercise caution! *But please do note that the type of error here is different from our 'little o' errors. The example here has to do with rounding errors, so there's no exact comparison. However this example serves as a warning to be careful with approximations: just because we can get a longer sequence of numbers when multiplying two approximations, doesn't mean it's correct.*

9.4 Local behavior of a function via its Taylor expansion

In this section we consider a function $f:\mathbb{R}\to\mathbb{R}$ and suppose that near the point $x_0\in\mathbb{R}$ it can be written as

$$f(x) = a_0 + a_1(x-x_0)+a_2(x-x_0)^2+\cdots+a_n(x-x_0)^n+o((x-x_0)^n)$$

Order and principal part of infinitesimal functions

Suppose that in the expansion of f near x_0, all coefficients a_i, $i=0,\ldots,m-1$, are 0, and $m>0$ is the first one such that $a_m\neq 0$. Then, near x_0, f can be expressed as

$$f(x) = a_m(x-x_0)^m + a_{m+1}(x-x_0)^{m+1}+\cdots+a_n(x-x_0)^n+o((x-x_0)^n)$$

and since $m > 0$ it means that $\lim_{x \to x_0} f(x) = 0$, so that f is infinitesimal at x_0. The above expression for f can be written even more crudely as

$$f(x) = \underbrace{a_m(x - x_0)^m}_{p(x)} + o((x - x_0)^m)$$

which immediately reveals to us that f is of order m at x_0 with respect to $\varphi(x) = x - x_0$, and has principal part $p(x) = a_m(x - x_0)^m$.

Example 9.6: Consider the function

$$f(x) = \sin x - x \cos x - \frac{1}{3}x^3.$$

Let us study this function around $x_0 = 0$. The relevant Maclaurin polynomials are:

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} - \cdots + (-1)^m \frac{x^{2m+1}}{(2m+1)!} + o(x^{2m+2})$$
$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} - \cdots + (-1)^m \frac{x^{2m}}{(2m)!} + o(x^{2m+1})$$

so that

$$x \cos x = x - \frac{x^3}{2!} + \frac{x^5}{4!} - \frac{x^7}{6!} - \cdots + (-1)^m \frac{x^{2m+1}}{(2m)!} + o(x^{2m+2})$$

It follows that

$$\begin{aligned} f(x) &= \sin x - x \cos x - \frac{1}{3}x^3 \\ &= \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} + o(x^6)\right) - \left(x - \frac{x^3}{2!} + \frac{x^5}{4!} + o(x^6)\right) - \frac{1}{3}x^3 \\ &= (1-1)x + \left(\frac{1}{2} - \frac{1}{6} - \frac{1}{3}\right) x^3 + \left(\frac{1}{5!} - \frac{1}{4!}\right) x^5 + o(x^6) \\ &= -\frac{1}{30}x^5 + o(x^6) \qquad \text{as } x \to 0. \end{aligned}$$

So f is infinitesimal of order 5 at $x_0 = 0$ with respect to $\varphi(x) = x$, and its principal part is $p(x) = -\frac{1}{30}x^5$.

Local behavior of a function

Since

$$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n + o((x - x_0)^n)$$

we can deduce that

$$\begin{aligned}
f(x_0) &= a_0 \\
f'(x_0) &= a_1 \\
f''(x_0) &= 2 \cdot a_2 \\
f'''(x_0) &= 3! \cdot a_3 \\
&\vdots \\
f^{(k)}(x_0) &= k! \cdot a_k \\
&\vdots
\end{aligned}$$

up to the nth derivative. In particular, knowing $f(x_0), f'(x_0), f''(x_0)$ means that we already know a lot about the function:

- $f(x_0)$ determines the sign of f near x_0 (if $f(x_0) \neq 0$)
- $f'(x_0)$ determines the monotonicity type of f near x_0 (if $f'(x_0) \neq 0$) or if x_0 is a critical point (if $f'(x_0) = 0$)
- $f''(x_0)$ determines convexity/concavity of f near x_0 (if $f''(x_0) \neq 0$)

Critical points

Theorem 9.6: Let f be differentiable $n \geq 2$ times at x_0 and suppose the for some $2 \leq m \leq n$

$$f'(x_0) = \cdots = f^{(m-1)}(x_0) = 0, \quad f^{(m)}(x_0) \neq 0$$

Then:

m even $\Rightarrow$ x_0 is an extremum point (maximum if $f^{(m)}(x_0) < 0$ and minimum if $f^{(m)}(x_0) > 0$).

m odd $\Rightarrow$ x_0 is an inflection point with horizontal tangent.

Proof. We skip the proof of this theorem. □

Example 9.7: Suppose that a function $f : \mathbb{R} \to \mathbb{R}$ can be written as

$$f(x) = 2 - 3(x-1)^4 + 2(x-1)^5 + o((x-1)^5), \qquad \text{as } x \to 1.$$

Then

$$f(1) = 2 \qquad f'(1) = f''(1) = f'''(1) = 0 \qquad f^{(4)}(1) = -3 \cdot 4! = -72 < 0$$

so that $x_0 = 1$ is a local maximum.

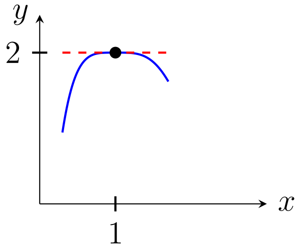


Figure 9.3: The Taylor polynomial $2 - 3(x-1)^4 + 2(x-1)^5$ near $x = 1$. It is extremely flat near $x = 1$.

Inflection points

In Corollary 8.20 we had the following statement which we did not prove since we lacked the tools:

Claim (Corollary 8.20)**:** Let f be a twice-differentiable function in a neighborhood of x_0. Then:

$$x_0 \text{ is an inflection point} \qquad \Rightarrow \qquad f''(x_0) = 0,$$

and

$$f''(x_0) = 0 \text{ and } f'' \text{ changes sign at } x_0 \qquad \Rightarrow \qquad x_0 \text{ is an inflection point.}$$

Moreover,

$$f''(x_0) = 0 \text{ and } f'' \text{ doesn't change sign at } x_0 \qquad \Rightarrow \qquad x_0 \text{ isn't an inflection point.}$$

Proof. Let us only prove the first implication. Assume that x_0 is an inflection point. By definition, this means that f is above the tangent line at x_0 on one side of x_0, and below on the other side. First we recall the first- and second-order Taylor polynomials at x_0:

$$\begin{aligned}(Tf)_{1,x_0}(x) &= f(x_0) + f'(x_0)(x - x_0)\\ (Tf)_{2,x_0}(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2\\ &= (Tf)_{1,x_0}(x) + \frac{1}{2}f''(x_0)(x - x_0)^2\end{aligned}$$

and we keep in mind that $(Tf)_{1,x_0}(x)$ **is, in fact, the equation of the tangent** at x_0.

We will use Peano's estimate of the remainder for a second-order Taylor polynomial at x_0:

$$f(x) = (Tf)_{2,x_0}(x) + o((x - x_0)^2), \qquad x \to x_0.$$

The error can be written more explicitly:

$$f(x) = (Tf)_{2,x_0}(x) + \varphi(x)$$

where $\varphi(x) = o((x - x_0)^2)$ near x_0. That is, φ is a function that satisfies:

$$\lim_{x \to x_0} \frac{\varphi(x)}{(x - x_0)^2} = 0.$$

Define

$$\psi(x) = \frac{\varphi(x)}{(x - x_0)^2}.$$

Then

$$\varphi(x) = \psi(x)(x - x_0)^2 \qquad \text{and} \qquad \lim_{x \to x_0} \psi(x) = 0.$$

Now we can write a precise expression for f:

$$\begin{aligned}f(x) &= (Tf)_{2,x_0}(x) + \varphi(x)\\ &= (Tf)_{2,x_0}(x) + \psi(x)(x - x_0)^2\\ &= (Tf)_{1,x_0}(x) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \psi(x)(x - x_0)^2\\ &= (Tf)_{1,x_0}(x) + \left[\frac{1}{2}f''(x_0) + \psi(x)\right](x - x_0)^2\end{aligned}$$

Moving $(Tf)_{1,x_0}(x)$ over to the left-hand-side will give us the difference between the function and the tangent:

$$f(x) - (Tf)_{1,x_0}(x) = \left[\frac{1}{2}f''(x_0) + \psi(x)\right](x - x_0)^2.$$

Recall our assumption: x_0 is an inflection point. By definition, this means that the right-hand-side should have different signs on either side of x_0. Since ψ is an infinitesimal function at x_0 (it is very small), if $f''(x_0) \neq 0$ then the term in the square brackets would have the same sign in a neighborhood of x_0. Since $(x - x_0)^2 \geq 0$, the entire right-hand-side would have the same sign in a neighborhood of x_0 (both sides of x_0). But since this is not the case, then necessarily $f''(x_0) = 0$. $\square$

Chapter 10

Intergral calculus

In this chapter we will discuss two fundamental problems for a given function f:

(1) finding another function F satisfying $F' = f$, and

(2) computing the area under its graph.

Through the *Fundamental Theorem of Integral Calculus* we will see that these two problems are, in fact, two aspects of the same problem.

10.1 Primitive functions and indefinite integrals

Primitive function

Given a function $f : \mathbb{R} \to \mathbb{R}$, defined on some interval $I \subseteq \mathbb{R}$, any function $F : \mathbb{R} \to \mathbb{R}$ satisfying

$$F'(x) = f(x), \qquad \forall x \in I,$$

is called a **primitive function (antiderivative)** of f. If $F'(x) = f(x)$ then also $\frac{\mathrm{d}}{\mathrm{d}x}(F(x) + C) = f(x)$ for any constant C, since the derivative of a constant is 0. Hence the antiderivative *isn't unique.* We therefore denote

$$\int f(x)\,\mathrm{d}x = F(x) + C$$

where the right hand side is an infinite set of functions.

$$\begin{tikzcd} \boxed{F(x)+C} \arrow[r, bend left, "\text{Derivative } \frac{\mathrm{d}}{\mathrm{d}x}"] & \boxed{f(x)} \arrow[l, bend left, "\text{Antiderivative } \int \ \mathrm{d}x"] \end{tikzcd}$$

Example 10.1: 1. Consider $f(x) = x$. We know that $\frac{\mathrm{d}}{\mathrm{d}x}(x^2) = 2x$ so that $\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}x}(x^2) = x$. It follows that

$$\int x\,\mathrm{d}x = \frac{1}{2}x^2 + C.$$

2. Consider $f(x) = x^2$. We know that $\frac{\mathrm{d}}{\mathrm{d}x}(x^3) = 3x^2$ so that $\frac{1}{3}\frac{\mathrm{d}}{\mathrm{d}x}(x^3) = x^2$. It follows that
$$\int x^2 \,\mathrm{d}x = \frac{1}{3}x^3 + C.$$

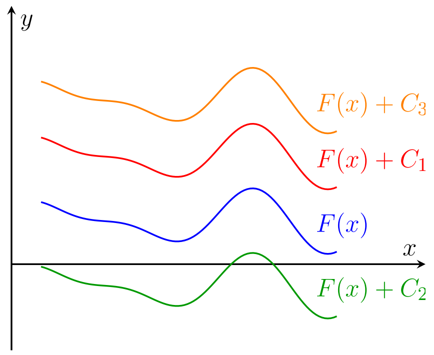
3. More generally, consider $f(x) = x^n$. We know that $\frac{\mathrm{d}}{\mathrm{d}x}(x^{n+1}) = (n+1)x^n$ so that $\frac{1}{n+1}\frac{\mathrm{d}}{\mathrm{d}x}(x^{n+1}) = x^n$. It follows that
$$\int x^n \,\mathrm{d}x = \frac{1}{n+1}x^{n+1} + C.$$

4. Consider $f(x) = e^x$. We know that $\frac{\mathrm{d}}{\mathrm{d}x}(e^x) = e^x$ so that
$$\int e^x \,\mathrm{d}x = e^x + C.$$

Following these ideas, we can list some of the fundamental antiderivatives:

$$\int x^\alpha \,\mathrm{d}x = \frac{x^{\alpha+1}}{\alpha+1} + C \quad (\alpha \neq -1)$$
$$\int \frac{1}{x} \,\mathrm{d}x = \log|x| + C \quad (\text{for } x > 0 \text{ or } x < 0)$$
$$\int \sin x \,\mathrm{d}x = -\cos x + C$$
$$\int \cos x \,\mathrm{d}x = \sin x + C$$
$$\int e^x \,\mathrm{d}x = e^x + C$$
$$\int \frac{1}{1+x^2} \,\mathrm{d}x = \arctan x + C$$
$$\int \frac{1}{\sqrt{1-x^2}} \,\mathrm{d}x = \arcsin x + C$$
$$\int \sinh x \,\mathrm{d}x = \cosh x + C$$
$$\int \cosh x \,\mathrm{d}x = \sinh x + C$$

It must be emphasized that there is no unique primitive function: if F is an antiderivative of f, then any vertical translation of F is also an antiderivative. There are *infinitely many antiderivatives*!



We can pinpoint which of these antiderivatives we are looking for by requiring that at some given x_0 the antiderivative has a particular value y_0: $F(x_0) = y_0$.

Example 10.2:

1. Identify the antiderivative $F(x)$ of $f(x) = \cos x$ that satisfies $F(2\pi) = 5$.

   The antiderivative is $\int f(x)\,\mathrm{d}x = \sin x + C = F(x)$. If we require that $F(2\pi) = 5$ then we find that

   $$\sin 2\pi + C = 5 \qquad \Rightarrow \qquad C = 5.$$

   This has forced a choice of C.

2. Find the value at $x_1 = 3$ of the antiderivative of $f(x) = 6x^2 + 5x$ that vanishes at the point $x_0 = 1$.

   First we identify the antiderivative:

   $$\int f(x)\,\mathrm{d}x = \int \left(6x^2 + 5x\right)\,\mathrm{d}x = \frac{6x^3}{3} + \frac{5x^2}{2} + C = 2x^3 + \frac{5}{2}x^2 + C = F(x).$$

   Plugging in the given condition we have:

   $$0 = F(1) = 2 + \frac{5}{2} + C \qquad \Rightarrow \qquad C = -\frac{9}{2}.$$

   Hence our antiderivative is:

   $$F(x) = 2x^3 + \frac{5}{2}x^2 - \frac{9}{2}$$

   and

   $$F(3) = 2 \cdot 3^3 + \frac{5}{2} \cdot 3^2 - \frac{9}{2} = 2 \cdot 27 + \frac{5}{2} \cdot 9 - \frac{9}{2} = 72.$$

3. Find the antiderivative of $f(x) = \sin 3x$. We see that $\frac{\mathrm{d}}{\mathrm{d}x}(\cos 3x) = -3\sin 3x$, so that

   $$F(x) = \int \sin 3x\,\mathrm{d}x = -\frac{1}{3}\cos 3x + C.$$

10.2 Rules of indefinite integration

Theorem 10.1 (Linearity): If $\int f(x)\,\mathrm{d}x$ and $\int g(x)\,\mathrm{d}x$ are defined, then for any $\alpha, \beta \in \mathbb{R}$, the antiderivative of $\alpha f(x) + \beta g(x)$ is also defined and satisfies:

$$\int \left(\alpha f(x) + \beta g(x)\right)\,\mathrm{d}x = \alpha \int f(x)\,\mathrm{d}x + \beta \int g(x)\,\mathrm{d}x.$$

Proof. Suppose that $F(x)$ is an antiderivative of $f(x)$ and that $G(x)$ is an antiderivative of $g(x)$. Then, using the linearity of the derivative, we have:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\alpha F(x) + \beta G(x)\right) = \alpha F'(x) + \beta G'(x) = \alpha f(x) + \beta g(x)$$

so that $\alpha F(x) + \beta G(x)$ is an antiderivative of $\alpha f(x) + \beta g(x)$ and the claimed equality is satisfied. $\square$

Example 10.3: Let us compute $\int \cos^2 x \, \mathrm{d}x$. We use the relation $\cos^2 x = \frac{1}{2}(1+\cos(2x))$ to write

$$\begin{aligned}
\int \cos^2 x \, \mathrm{d}x &= \int \frac{1}{2}(1 + \cos(2x)) \, \mathrm{d}x \\
&= \frac{1}{2}\int 1 \, \mathrm{d}x + \frac{1}{2}\int \cos(2x) \, \mathrm{d}x \\
&= \frac{x}{2} + \frac{1}{2}\frac{\sin(2x)}{2} + C \\
&= \frac{x}{2} + \frac{\sin(2x)}{4} + C.
\end{aligned}$$

Integration by parts

Let $f(x)$ and $g(x)$ be differentiable functions over an interval $I \subseteq \mathbb{R}$. Then

$$\int f(x)g'(x) \, \mathrm{d}x = f(x)g(x) - \int f'(x)g(x) \, \mathrm{d}x$$

if the right hand side is defined (i.e. if $f'(x)g(x)$ is integrable on I).

Proof. Since $f'(x)g(x)$ is integrable on I, we let $H(x)$ be an antiderivative, i.e. $H'(x) = f'(x)g(x)$. Then, using the linearity of the derivative and the product rule, we have:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\big(f(x)g(x) - H(x)\big) &= \frac{\mathrm{d}}{\mathrm{d}x}\big(f(x)g(x)\big) - H'(x) \\
&= f'(x)g(x) + f(x)g'(x) - f'(x)g(x) = f(x)g'(x)
\end{aligned}$$

which implies that $f(x)g(x) - H(x)$ is an antiderivative of $f(x)g'(x)$, as has been claimed. $\square$

An easy way to remember this is as follows, starting from the product rule:

$$\frac{\mathrm{d}}{\mathrm{d}x}\big(f(x)g(x)\big) = f'(x)g(x) + f(x)g'(x)$$
$$\Downarrow$$
$$f(x)g'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\big(f(x)g(x)\big) - f'(x)g(x)$$
$$\Downarrow$$
$$\int f(x)g'(x)\,\mathrm{d}x = \int \frac{\mathrm{d}}{\mathrm{d}x}\big(f(x)g(x)\big)\,\mathrm{d}x - \int f'(x)g(x)\,\mathrm{d}x$$
$$\Downarrow$$
$$\int f(x)g'(x)\,\mathrm{d}x = f(x)g(x) - \int f'(x)g(x)\,\mathrm{d}x$$

A fundamental concept

Integration by parts is a fundamental concept in mathematical physics and beyond, as it allows us to transfer a derivative from one function to another:

$$\int f(x)\mathbf{g}'(\mathbf{x})\,\mathrm{d}x = f(x)\mathbf{g}(\mathbf{x}) - \int f'(x)\mathbf{g}(\mathbf{x})\,\mathrm{d}x$$

Observe that on the left hand side, g has a derivative and f does not, while on the right hand side g doesn't have a derivative and f does. So going from left to right, we have 'moved' a derivative from g onto f. In equations coming from mathematical physics (for instance, the wave equation, the heat equation or Schrödinger's equation) we often use this idea to prove certain properties of solutions. For instance, this is used to prove that the energy of an electromagnetic wave is conserved, while the energy carried by heat dissipates over time.

Example 10.4: 1. Determine $\int xe^x\,\mathrm{d}x$. We have two functions — x and e^x — one of which will need to take the role of $f(x)$ and the other will take the role of $g'(x)$. We observe that the following choice can work quite nicely:

$$\begin{aligned} f(x) &= x & \qquad f'(x) &= 1 \\ g'(x) &= e^x & \qquad g(x) &= e^x \end{aligned}$$

so that we get

$$\int xe^x\,\mathrm{d}x = xe^x - \int e^x\,\mathrm{d}x = xe^x - e^x + C = e^x(x-1) + C.$$

Observe that if we had made the other choice we would have ended up with a more complicated integral. Indeed, choose

$$f(x) = e^x \qquad f'(x) = e^x$$
$$g'(x) = x \qquad g(x) = \frac{1}{2}x^2$$

so that

$$\int xe^x \,\mathrm{d}x = \frac{1}{2}x^2e^x - \frac{1}{2}\int x^2e^x \,\mathrm{d}x = ?$$

2. Determine $\int \ln x \,\mathrm{d}x$. Here our two functions are $\ln x$ and the constant function 1! Here is our choice for f and g:

$$f(x) = \ln x \qquad f'(x) = \frac{1}{x}$$
$$g'(x) = 1 \qquad g(x) = x$$

which implies

$$\int \ln x \,\mathrm{d}x = x\ln x - \int \frac{x}{x}\,\mathrm{d}x = x\ln x - x + C.$$

3. Determine $I_1 = \int e^x \sin x \,\mathrm{d}x$. We choose

$$f(x) = e^x \qquad f'(x) = e^x$$
$$g'(x) = \sin x \qquad g(x) = -\cos x$$

so that

$$I_1 = \int e^x \sin x \,\mathrm{d}x = -e^x\cos x + \underbrace{\int e^x \cos x \,\mathrm{d}x}_{I_2}.$$

We still cannot determine the integral on the right hand side, so we perform another integration by parts with

$$f(x) = e^x \qquad f'(x) = e^x$$
$$g'(x) = \cos x \qquad g(x) = \sin x$$

hence

$$I_2 = \int e^x \cos x \,\mathrm{d}x = e^x \sin x - \underbrace{\int e^x \sin x \,\mathrm{d}x}_{I_1}$$

and we observe that we're back to the original integral I_1. Putting everything together, we found that

$$I_1 = -e^x\cos x + I_2 = -e^x \cos x + e^x \sin x - I_1.$$

Rearranging, we finally get

$$I_1 = \frac{1}{2}e^x(\sin x - \cos x) + C.$$

Integration by substitution

Let $f(y)$ be integrable on an interval J and let $F(y)$ be an antiderivative. Suppose that $\varphi(x) : I \to J$ is differentiable. Then $f(\varphi(x))\varphi'(x)$ is integrable on I and

$$\int f(\varphi(x))\varphi'(x)\,\mathrm{d}x = F(\varphi(x)) + C.$$

A simpler way to remember this formula is by writing $y = \varphi(x)$ to get:

$$\int f(\varphi(x))\varphi'(x)\,\mathrm{d}x = \int f(y)\frac{\mathrm{d}y}{\mathrm{d}x}\,\mathrm{d}x = \int f(y)\,\mathrm{d}y.$$

Proof. From the chain rule (Theorem 8.3) we know that

$$\frac{\mathrm{d}}{\mathrm{d}x}\big(F(\varphi(x))\big) = F'(\varphi(x))\varphi'(x) = f(\varphi(x))\varphi'(x).$$

By definition of the antiderivative the result follows. □

Example 10.5: 1. Determine $\int \sin(k(x-x_0))\,\mathrm{d}x$, where $0 \neq k \in \mathbb{R}$ and $x_0 \in \mathbb{R}$. We let

$$y = \varphi(x) = k(x-x_0) \qquad \text{so that} \qquad \varphi'(x) = k.$$

Then we have

$$\begin{aligned}\int \sin(k(x-x_0))\,\mathrm{d}x = \int \sin(\varphi(x))\,\mathrm{d}x &= \int \sin(\varphi(x))\frac{k}{k}\,\mathrm{d}x\\ &= \frac{1}{k}\int \sin(\varphi(x))\varphi'(x)\,\mathrm{d}x\\ &= \frac{1}{k}\int \sin y\,\mathrm{d}y\\ &= -\frac{1}{k}\cos y + C\\ &= -\frac{1}{k}\cos(k(x-x_0)) + C.\end{aligned}$$

2. Determine $\int xe^{x^2}\,\mathrm{d}x$. We let

$$y = \varphi(x) = x^2 \qquad \text{so that} \qquad \varphi'(x) = 2x.$$

Then we have

$$\begin{aligned}\int xe^{x^2}\,\mathrm{d}x = \frac{1}{2}\int 2xe^{x^2}\,\mathrm{d}x &= \frac{1}{2}\int \varphi'(x)e^{\varphi(x)}\,\mathrm{d}x\\ &= \frac{1}{2}\int e^y\,\mathrm{d}y\\ &= \frac{1}{2}e^y + C\\ &= \frac{1}{2}e^{x^2} + C.\end{aligned}$$

3. Determine $\int \tan x \, \mathrm{d}x$. We let

$$y = \varphi(x) = \cos x \qquad \text{so that} \qquad \varphi'(x) = -\sin x.$$

Then we have

$$\begin{aligned}\int \tan x \, \mathrm{d}x = \int \frac{\sin x}{\cos x} \, \mathrm{d}x &= \int \frac{-\varphi'(x)}{\varphi(x)} \, \mathrm{d}x \\ &= -\int \frac{1}{y} \, \mathrm{d}y \\ &= -\ln|y| + C \\ &= -\ln|\cos x| + C.\end{aligned}$$

4. Determine $\int \frac{\varphi'(x)}{\varphi(x)} \, \mathrm{d}x$. This generalizes the previous example. Let

$$y = \varphi(x) \qquad \text{so that} \qquad \frac{\mathrm{d}y}{\mathrm{d}x} = \varphi'(x).$$

Then we have

$$\int \frac{\varphi'(x)}{\varphi(x)} \, \mathrm{d}x = \int \frac{1}{y} \frac{\mathrm{d}y}{\mathrm{d}x} \, \mathrm{d}x = \int \frac{1}{y} \, \mathrm{d}y = \ln|y| + C = \ln|\varphi(x)| + C.$$

5. Determine $\int \frac{1}{\sqrt{1+x^2}} \, \mathrm{d}x$. Here's a trick. Let

$$y = \varphi(x) = \sqrt{1+x^2} - x.$$

It follows that

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \varphi'(x) = \frac{x}{\sqrt{1+x^2}} - 1 = \frac{x - \sqrt{1+x^2}}{\sqrt{1+x^2}} = \frac{-\varphi(x)}{\sqrt{1+x^2}} = \frac{-y}{\sqrt{1+x^2}}$$

which means that

$$\frac{1}{\sqrt{1+x^2}} = -\frac{1}{y} \frac{\mathrm{d}y}{\mathrm{d}x}.$$

Integrating, we get

$$\int \frac{1}{\sqrt{1+x^2}} \, \mathrm{d}x = -\int \frac{1}{y} \frac{\mathrm{d}y}{\mathrm{d}x} \, \mathrm{d}x = -\int \frac{1}{y} \, \mathrm{d}y = -\ln|y| + C = -\ln|\sqrt{1+x^2} - x| + C.$$

Since $\sqrt{1+x^2} - x > 0$ we don't need the absolute value, and we have

$$\int \frac{1}{\sqrt{1+x^2}} \, \mathrm{d}x = -\ln(\sqrt{1+x^2} - x) + C.$$

Remembering how to integrate by substitution

When asked to integrate $\int g(x)\,\mathrm{d}x$ there are two options:

(a) We are able identify that there exists $y = y(x)$ such that $g(x)$ has the form

$$g(x) = f(y(x))y'(x)$$

as we had done in the examples above. In this case, since $\frac{\mathrm{d}y}{\mathrm{d}x} = y'(x)$ we write $\mathrm{d}y = y'(x)\,\mathrm{d}x$ to get

$$\int g(x)\,\mathrm{d}x = \int f(y(x))y'(x)\,\mathrm{d}x = \int f(y)\,\mathrm{d}y.$$

This approach might work if g is a complicated function.

(b) If we cannot identify $y = y(x)$ as above, we try to go about it the other way around: identify x as a function of y: $x = x(y)$, compute $\frac{\mathrm{d}x}{\mathrm{d}y} = x'(y)$ and write $\mathrm{d}x = x'(y)\,\mathrm{d}y$ to get:

$$\int g(x)\,\mathrm{d}x = \int g(x(y))x'(y)\,\mathrm{d}y.$$

This approach might work if g is a simple function.

6. Determine $\int \sqrt{1-x^2}\,\mathrm{d}x$. Let's use the second approach from above. We start with

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

Now we choose $x(y) = \sin y$ because it will simplify g, since $1 - \sin^2 y = \cos^2 y$. So we have

$$x(y) = \sin y \qquad \text{so that} \qquad \frac{\mathrm{d}x}{\mathrm{d}y} = \cos y \quad \Rightarrow \quad \mathrm{d}x = \cos y\,\mathrm{d}y$$

and we get

$$\begin{aligned}
\int \sqrt{1-x^2}\,\mathrm{d}x &= \int \sqrt{1-\sin^2 y}\cos y\,\mathrm{d}y \\
&= \int \cos^2 y\,\mathrm{d}y \\
&= \frac{y}{2} + \frac{\sin(2y)}{4} + C \\
&= \frac{y}{2} + \frac{1}{2}\sin y\cos y + C \\
&= \frac{y}{2} + \frac{1}{2}\sin y\sqrt{1-\sin^2 y} + C \\
\textit{(using that } y = \arcsin x\textit{)} \qquad &= \frac{1}{2}\arcsin x + \frac{1}{2}x\sqrt{1-x^2} + C.
\end{aligned}$$

7. Determine $\int \frac{1}{e^x+e^{-x}}\,\mathrm{d}x$. Using the second approach, we have

$$g(x) = \frac{1}{e^x + e^{-x}}.$$

Choose $x(y) = \ln y$ to undo the exponents. Then we have

$$x(y) = \ln y \qquad \text{so that} \qquad \frac{\mathrm{d}x}{\mathrm{d}y} = \frac{1}{y} \qquad \Rightarrow \qquad \mathrm{d}x = \frac{1}{y}\,\mathrm{d}y$$

and we get

$$\begin{aligned}\int \frac{1}{e^x + e^{-x}}\,\mathrm{d}x &= \int \frac{1}{e^{\ln y} + e^{-\ln y}}\frac{1}{y}\,\mathrm{d}y \\ &= \int \frac{1}{y + \frac{1}{y}}\frac{1}{y}\,\mathrm{d}y \\ &= \int \frac{1}{1+y^2}\,\mathrm{d}y \\ &= \arctan y + C = \arctan e^x + C.\end{aligned}$$

Rational functions

It turns out that we can integrate rational functions such as

$$r(x) = \frac{p_n(x)}{q_m(x)}$$

The general theory is beyond the scope of this course, so we will stick to simple examples. However, we note that the key is the (deep) algebraic fact that every polynomial can be factorized into simple factors of the following forms:

$$(x+d)^s \qquad \text{or} \qquad (x^2+bx+c)^t, \text{ with discriminant } \Delta = b^2 - 4c < 0$$

In particular, we can write

$$q_m(x) = (x+d_1)^{s_1}(x+d_2)^{s_2}\cdots(x^2+b_1x+c_1)^{t_1}(x^2+b_2x+c_2)^{t_2}\cdots$$

where $m = s_1+s_2+\cdots+2(t_1+t_2+\cdots)$ is the degree of q_m, and hence the total number of roots (counting multiplicity). Consequently, the rational function $r(x)$ can be rewritten (roughly, there are a few missing terms) as

$$r(x) \approx \frac{A_1}{(x+d_1)^{s_1}} + \frac{A_2}{(x+d_2)^{s_2}} + \cdots + \frac{B_1x+C_1}{(x^2+b_1x+c_1)^{t_1}} + \frac{B_2x+C_2}{(x^2+b_2x+c_2)^{t_2}} + \cdots$$

where the constants A_i, B_i, C_i are determined by multiplying all the fractions on the right hand side and equating the resulting numerator to $p_n(x)$. The conclusion is that it is enough to know how to integrate rational functions of the forms

$$\frac{A_1}{(x+d_1)^{s_1}} \qquad \text{and} \qquad \frac{B_1x+C_1}{(x^2+b_1x+c_1)^{t_1}}$$

Let us try to see how this might work.

Example 10.6: 1. The simplest case is easy:

$$\int \frac{1}{x+c}\,\mathrm{d}x = \ln|x+c| + C.$$

2. The next case follows from what we know for power functions ($s > 1$):

$$\int \frac{1}{(x+c)^s}\,\mathrm{d}x = \frac{1}{1-s}\frac{1}{(x+c)^{s-1}} + C.$$

3. Next we consider $\int g(x)\,\mathrm{d}x$ where

$$g(x) = \frac{1}{x^2+bx+c} \qquad \text{with discriminant } \Delta = b^2 - 4c < 0$$

so that the polynomial in the denominator is strictly positive (no real roots). We can write

$$x^2+bx+c = x^2+bx+\underbrace{c-\frac{1}{4}b^2}_{-\frac{1}{4}\Delta}+\frac{1}{4}b^2 = \underbrace{x^2+bx+\frac{1}{4}b^2}_{(x+\frac{1}{2}b)^2}-\frac{1}{4}\Delta = (x+\frac{1}{2}b)^2 - \frac{1}{4}\Delta$$

$$= -\frac{1}{4}\Delta\left[1+\left(\frac{x+\frac{1}{2}b}{\frac{1}{2}\sqrt{-\Delta}}\right)^2\right]$$

Make the substitution

$$y(x) = \frac{x+\frac{1}{2}b}{\frac{1}{2}\sqrt{-\Delta}}$$

whose derivative is

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y'(x) = \frac{1}{\frac{1}{2}\sqrt{-\Delta}} \qquad \Rightarrow \qquad \mathrm{d}x = \frac{1}{2}\sqrt{-\Delta}\,\mathrm{d}y$$

We therefore have

$$\begin{aligned}
\int \frac{1}{x^2+bx+c}\,\mathrm{d}x &= -\frac{1}{\frac{1}{4}\Delta}\int \frac{1}{1+\left(\frac{x+\frac{1}{2}b}{\frac{1}{2}\sqrt{-\Delta}}\right)^2}\,\mathrm{d}x \\
&= -\frac{1}{\frac{1}{4}\Delta}\int \frac{1}{1+y^2}\frac{1}{2}\sqrt{-\Delta}\,\mathrm{d}y \\
&= \frac{1}{\frac{1}{2}\sqrt{-\Delta}}\int \frac{1}{1+y^2}\,\mathrm{d}y \\
&= \frac{1}{\frac{1}{2}\sqrt{-\Delta}}\arctan y + C \\
&= \frac{1}{\frac{1}{2}\sqrt{-\Delta}}\arctan\left(\frac{x+\frac{1}{2}b}{\frac{1}{2}\sqrt{-\Delta}}\right) + C
\end{aligned}$$

4. Determine $\int \frac{4x-5}{x^2-2x+10}\,\mathrm{d}x$. We write

$$\int \frac{4x-5}{x^2-2x+10}\,\mathrm{d}x = 2\int \frac{2x-2}{x^2-2x+10} - \int \frac{1}{x^2-2x+10}\,\mathrm{d}x$$

The first integral is of the form $\int \frac{\varphi'(x)}{\varphi(x)}\,\mathrm{d}x = \ln|\varphi(x)| + C$, while for the second integral we see that

$$\Delta = (-2)^2 - 4 \cdot 10 = -36 < 0$$

so that, using our previous computation, we have

$$\int \frac{1}{x^2-2x+10}\,\mathrm{d}x = \frac{1}{\frac{1}{2}\sqrt{36}} \arctan\left(\frac{x-1}{\frac{1}{2}\sqrt{36}}\right) = \frac{1}{3}\arctan\left(\frac{x-1}{3}\right)$$

and we therefore conclude that

$$\int \frac{4x-5}{x^2-2x+10}\,\mathrm{d}x = 2\ln(x^2-2x+10) - \frac{1}{3}\arctan\left(\frac{x-1}{3}\right) + C.$$

5. Determine $\int \frac{2x^3+x^2-4x+7}{x^2+x-2}\,\mathrm{d}x$. First we simplify the rational function:

$$\frac{2x^3+x^2-4x+7}{x^2+x-2} = \frac{2x(x^2+x-2)-(x^2+x-2)+x+5}{x^2+x-2} = 2x-1+\frac{x+5}{x^2+x-2}$$

The next step is to see if we can factor the denominator in the last term. In this case we can indeed see that

$$x^2+x-2 = (x+2)(x-1)$$

so that we write

$$\begin{aligned}\frac{x+5}{x^2+x-2} &= \frac{A_1}{x-1} + \frac{A_2}{x+2} \\ &= \frac{A_1(x+2)+A_2(x-1)}{(x-1)(x+2)} \\ &= \frac{(A_1+A_2)x+(2A_1-A_2)}{(x-1)(x+2)}\end{aligned}$$

So we find that A_1 and A_2 solve the system of equations

$$A_1 + A_2 = 1 \qquad \text{and} \qquad 2A_1 - A_2 = 5$$

so that

$$A_1 = 2 \qquad \text{and} \qquad A_2 = -1.$$

So we finally have:

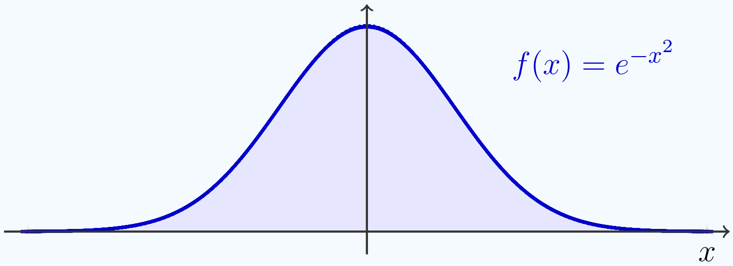
$$\begin{aligned}\int \frac{2x^3+x^2-4x+7}{x^2+x-2}\,\mathrm{d}x &= \int (2x-1)\,\mathrm{d}x + \int \frac{2}{x-1}\,\mathrm{d}x - \int \frac{1}{x+2}\,\mathrm{d}x \\ &= x^2 - x + 2\ln|x-1| - \ln|x+2| + C.\end{aligned}$$

Not all functions can be integrated

We have seen that integrating functions is not easy. It is important to understand that it is not even guaranteed to be possible. There exist functions that do not have an antiderivative that can be expressed using any of the functions we know. An example of this is the function

$$f(x) = e^{-x^2}$$

which is known as a **Gaussian** or a **bell curve**. This function is extremely important in probability theory, and yet we cannot express its antiderivatives using any of the functions that we know.

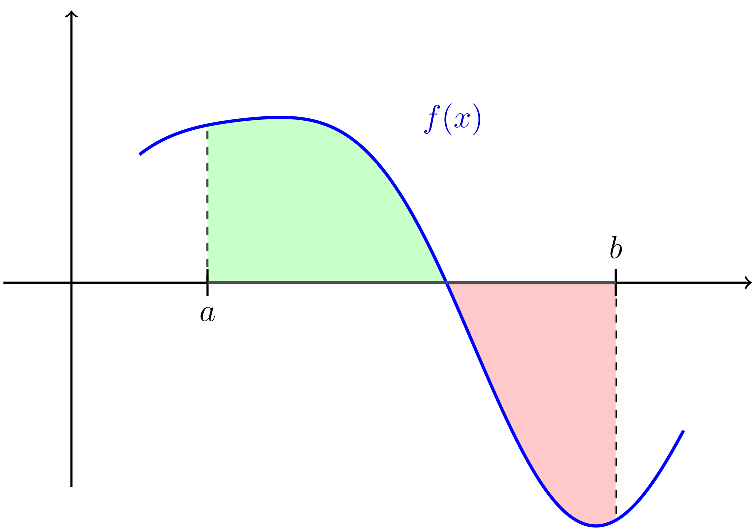


10.3 Definite integrals

The *definite* integral, denoted

$$\int_a^b f(x)\,\mathrm{d}x$$

is the measurement of the area under the graph of f between $x = a$ and $x = b$, where whenever the graph lies below the x-axis we count the corresponding area as a *negative* area.



In the next sections we will give precise definitions, one due to Augustin-Louis Cauchy (18th-19th centuries, France) and one due to Bernhard Riemann (19th century, Germany). Cauchy's integral is simpler to understand, and lends itself better to computations (including on a computer), but Riemann's definition is the more 'correct' mathematical one. They lead to the same result.

10.4 Cauchy integral

Let $f : \mathbb{R} \to \mathbb{R}$ be continuous on $[a, b]$ where $a < b$. Divide the interval $[a, b]$ into n equal subintervals of width

$$\Delta x = \frac{b-a}{n}$$

with the partition points x_k satisfying

$$x_0 = a, \quad x_1 = x_0 + \Delta x, \ldots, \quad x_k = x_0 + k\Delta x, \ldots, \quad x_n = x_0 + n\Delta x = b.$$

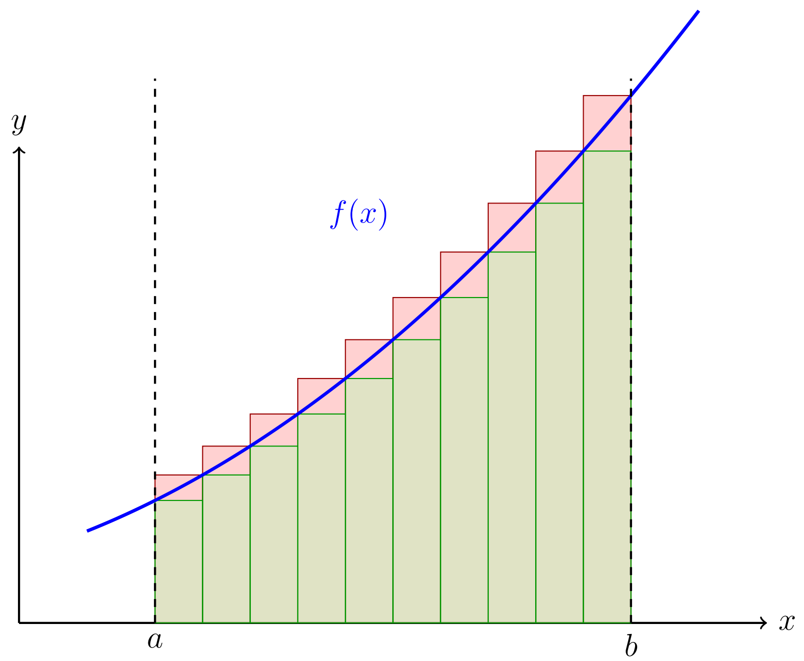
This is called an equal **partition** of $[a, b]$ into n subintervals, or an n-**partition**. Denote the n intervals by I_k, $k = 1, \ldots, n$, so that the kth interval is

$$I_k = [x_{k-1}, x_k], \qquad k = 1, \ldots, n.$$

On each of these intervals we can define the minimum and the maximum

$$m_k = \min_{I_k} f(x) \qquad \text{and} \qquad M_k = \max_{I_k} f(x)$$

(their existence is guaranteed thanks to Weierstrass' Theorem).



Lower and upper sums

The area under the graph of f between a and b can be approximated from below by rectangles of heigh m_k over each interval I_k, and from above by rectangles of height M_k. We therefore define the **lower and upper sums**, respectively:

$$s_n = \sum_{k=1}^{n} m_k \Delta x \qquad \text{and} \qquad S_n = \sum_{k=1}^{n} M_k \Delta x$$

Since $\Delta x > 0$ and $m_k \leq M_k$, we have that $s_n \leq S_n$.

Definition 10.2: A partition of a partition is called a **refinement**.

Lemma 10.3: A refinement is in itself a partition of $[a, b]$. In particular, if an m-partition is refined by subdividing into ℓ further subintervals, the result is an $(m \cdot \ell)$-partition.

Proof. An m-partition is given by

$$a = x_0 < x_1 < \cdots < x_m = b$$

where

$$x_k = x_0 + k\frac{b-a}{m}.$$

A refinement into ℓ sub-subintervals partitions the kth subinterval into

$$x_{k-1} = x_{k-1,0} < x_{k-1,1} < \cdots < x_{k-1,\ell} = x_k$$

where

$$x_{k-1,j} = x_{k-1} + j\frac{x_k - x_{k-1}}{\ell}.$$

But, using the fact that $x_k - x_{k-1} = \frac{b-a}{m}$ we have

$$x_{k-1,j} - x_{k-1,j-1} = \frac{x_k - x_{k-1}}{\ell} = \frac{\frac{b-a}{m}}{\ell} = \frac{b-a}{m \cdot \ell}.$$

This proves that the refinement is an $(m \cdot \ell)$-partition with partition points

$$a = x_0 = x_{0,0} < x_{0,1} < \cdots < x_{0,\ell} = x_1 = x_{1,0} < x_{1,1} \cdots x_{m-1,\ell-1} < x_{m-1,\ell} = x_m = b.$$

□

Lemma 10.4: Under a refinement, the lower sum cannot decrease, and the upper sum cannot increase.

Proof. Let the refinement be as in the previous proof. Denote

$$I_{k,j} = [x_{k-1,j-1}, x_{k-1,j}], \qquad k = 1, \ldots, m, \; j = 1, \ldots, \ell,$$

and observe that

$$I_{k,j} \subset I_k = [x_{k-1}, x_k].$$

Then, since the minimum over a larger set can only be smaller, we have

$$m_k = \min_{I_k} f(x) \le \min_{I_{k,j}} f(x) = m_{k,j}, \qquad k = 1, \dots, m,\ j = 1, \dots, \ell.$$

It follows that

$$s_m = \sum_{k=1}^{m} m_k \frac{b-a}{m} \le \sum_{k=1}^{m} \sum_{j=1}^{\ell} m_{k,j} \frac{b-a}{m \cdot \ell} = s_{m \cdot \ell}.$$

This proves that the lower sum cannot decrease. Similarly, the upper sum cannot increase. □

Theorem 10.5: The sequences $\{s_n\}_{n \in \mathbb{N}_+}$ and $\{S_n\}_{n \in \mathbb{N}_+}$ are both convergent, and both have the same limit.

Definite integral

The limit of s_n and S_n is called the **definite integral of f on** $[a, b]$ and is denoted

$$\int_a^b f(x)\, \mathrm{d}x = \lim_{n \to \infty} s_n = \lim_{n \to \infty} S_n.$$

Proof. **Step 1. Compare two different partitions.** We know that $s_n \le S_n$, but we want to be able to compare any s_m to any S_ℓ. The first corresponds to a partition into m subintervals and the second corresponds to a partition into ℓ subintervals. These have nothing to do with one another, but we can refine them to create a common $m \cdot \ell$ partition. From Lemma 10.4, we have

$$s_m \le s_{m \cdot \ell} \qquad \text{and} \qquad S_{m \cdot \ell} \le S_\ell.$$

But since we know that for any n, $s_n \le S_n$, we have

$$s_m \le s_{m \cdot \ell} \le S_{m \cdot \ell} \le S_\ell$$

and we therefore have a comparison of s_m and S_ℓ for any m and ℓ. It follows that the sequence $\{s_n\}_{n \in \mathbb{N}_+}$ has an upper bound, and the sequence $\{S_n\}_{n \in \mathbb{N}_+}$ has a lower bound. So we can define their supremum and infimum, respectively, which are finite, and satisfy:

$$s = \sup_n s_n \qquad \text{is less than or equal to} \qquad S = \inf_n S_n$$

Step 2. Prove that $s = S$. Since f is continuous on the closed interval $[a, b]$, the Heine-Cantor Theorem (Theorem 7.11) implies that it is *uniformly continuous* there. That is, for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|x_1 - x_2| < \delta$ implies $|f(x_1) - f(x_2)| < \varepsilon$.

Fix $\varepsilon > 0$. Let N_ε be big enough so that

$$\frac{b-a}{N_\varepsilon} < \delta.$$

Let $n > N_\varepsilon$. Then

$$\Delta x = \frac{b-a}{n} < \frac{b-a}{N_\varepsilon} < \delta.$$

Consider the kth subinterval, I_k. There are $\xi_k \in I_k$ and $\eta_k \in I_k$ such that

$$f(\xi_k) = m_k = \min_{I_k} f \qquad \text{and} \qquad f(\eta_k) = M_k = \max_{I_k} f.$$

Since both points belong to I_k, their distance is less than δ. Uniform continuity implies that

$$M_k - m_k = f(\eta_k) - f(\xi_k) < \varepsilon.$$

We therefore have that

$$\begin{aligned} S_n - s_n &= \sum_{k=1}^{n} M_k \Delta x - \sum_{k=1}^{n} m_k \Delta x \\ &= \sum_{k=1}^{n} (M_k - m_k) \Delta x \\ &< \varepsilon \sum_{k=1}^{n} \Delta x \\ &= \varepsilon(b-a). \end{aligned}$$

By definition of s and of S, we have

$$0 \le S - s \le S_n - s_n < \varepsilon(b-a).$$

Since ε was arbitrary, we find that $s = S$ (by the Squeeze Theorem, for instance).

Step 3. Prove that the limits exist. Since $S_n \ge s = S$, we also have that

$$0 \le S - s_n \le S_n - s_n < \varepsilon(b-a)$$

so that the limit of $\{s_n\}_{n\in\mathbb{N}_+}$ exists and

$$\lim_{n\to\infty} s_n = S.$$

Similarly, since $s_n \le s = S$ we have

$$0 \le S_n - S \le S_n - s_n < \varepsilon(b-a)$$

so that the limit of $\{S_n\}_{n\in\mathbb{N}_+}$ exists and

$$\lim_{n\to\infty} S_n = S.$$

$\square$

Piecewise continuous functions

Definition 10.6: A function $f : \mathbb{R} \to \mathbb{R}$ is said to be **piecewise continuous on** $[a,b]$ if it is continuous everywhere in $[a,b]$, except for finitely many points where it might have either a removable discontinuity or a jump discontinuity.

Let $y_0 = a < y_1 < \cdots < y_{m-1} < y_m = b$ be the points in $[a, b]$ where f is discontinuous. Define $f_k(x)$ to be the restriction of f to the interval $[y_{k-1}, y_k]$, where at the endpoints f_k is defined to be the limit of its values from the interior points:

$$f_k(x) = \begin{cases} \lim_{x \to y_{k-1}^+} f(x) & \text{when } x = y_{k-1}, \\ f(x) & \text{when } x \in (y_{k-1}, y_k), \\ \lim_{x \to y_k^-} f(x) & \text{when } x = y_k. \end{cases}$$

Then we can define

$$\begin{aligned} \int_a^b f(x)\,\mathrm{d}x &= \int_a^{y_1} f_1(x)\,\mathrm{d}x + \int_{y_1}^{y_2} f_2(x)\,\mathrm{d}x + \cdots + \int_{y_m}^{b} f_m(x)\,\mathrm{d}x \\ &= \sum_{k=1}^{m} \int_{y_{k-1}}^{y_k} f_k(x)\,\mathrm{d}x. \end{aligned}$$

and, thus, can still define the definite (Cauchy) integral despite the discontinuities.

Example 10.7: Consider the function $f(x) = x$ on $[0, 1]$. We know that the area under the curve is a triangle with area $\frac{1}{2}$. Let us show that the Cauchy integral gives us the same result. We take an n-partition of $[0, 1]$, giving us partition points $x_k = \frac{k}{n}$, $k = 0, 1, \ldots, n$, so that $\Delta x = \frac{1}{n}$. On each I_k, the function $f(x) = x$ will attain its minimum on the left endpoint, and the maximum on the right end point. We therefore have

$$\begin{aligned} s_n &= \sum_{k=1}^{n} x_{k-1}\Delta x = \frac{1}{n}\left(0 + \frac{1}{n} + \frac{2}{n} + \cdots + \frac{n-1}{n}\right) = \frac{1}{n^2}\left(1 + 2 + \cdots + (n-1)\right) \\ &= \frac{1}{n^2}\frac{n(n-1)}{2} = \frac{1}{2}\frac{n-1}{n} \xrightarrow{n\to\infty} \frac{1}{2} \end{aligned}$$

and

$$\begin{aligned} S_n &= \sum_{k=1}^{n} x_k \Delta x = \frac{1}{n}\left(\frac{1}{n} + \frac{2}{n} + \cdots + \frac{n}{n}\right) = \frac{1}{n^2}\left(1 + 2 + \cdots + n\right) \\ &= \frac{1}{n^2}\frac{(n+1)n}{2} = \frac{1}{2}\frac{n+1}{n} \xrightarrow{n\to\infty} \frac{1}{2} \end{aligned}$$

which confirms that we obtain the correct result:

$$\int_0^1 x\,\mathrm{d}x = \frac{1}{2}.$$

10.5 Riemann integral

The problem with the Cauchy integral

In the previous section we saw that the limits of the lower and upper sums of a continuous function f over an interval $[a, b]$ exist and are equal to one another. Thus it was natural to conclude that this limit is indeed the area, which is called the definite integral of f over $[a, b]$ and is denoted $\int_a^b f(x)\,\mathrm{d}x$. The mathematical problem with this process is the following:

> The partitions we used were very specific: we split $[a, b]$ into n subintervals of equal length. *Could it be that the outcome was influenced by this?*

The solution

The solution to this question is to allow ***un*even partitions**. Instead of choosing an even n-partition with $x_0 = a$ and $x_k = x_0 + k\frac{b-a}{n}$, we take some points

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$$

which might *not* be equally distributed within the interval $[a, b]$. We denote, as before, the kth subinterval

$$I_k = [x_{k-1}, x_k], \qquad k = 1, \dots, n.$$

As before, a **refinement** is a further partition of an existing partition (only that now we do not require an even partition). In what follows we denote

$$I = [a, b].$$

Step functions

Definition 10.7: A function $f : I \to \mathbb{R}$ is called **step function** if it is constant on subintervals of I. We say that it is **adapted** to a particular partition $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ if it is constant on the interior of any of the subintervals I_k. We denote

$$\mathcal{S}(I) = \text{the set of all step functions on } I.$$

Definition 10.8: Let $f \in \mathcal{S}(I)$ and let $a = x_0 < x_1 < \cdots < x_n = b$ be an adapted partition. Denote by c_k the value of f on I_k. Then we define the **definite integral** of f on I to be

$$\int_I f = \sum_{k=1}^n c_k(x_k - x_{k-1}).$$

Proposition 10.9: Let $g, h \in \mathcal{S}(I)$ be two step functions on I and assume that $g(x) \le h(x)$ for all $x \in I$. Then

$$\int_I g \le \int_I h.$$

Proof. We take a partition $\{x_k\}_{k=0}^n$ that is adapted to both g and h (i.e., a partition that includes all points where either g or h have a jump). Then on any interval I_k there exist constants c_k, d_k such that

$$c_k = g(x) \le h(x) = d_k, \qquad \forall x \in I_k.$$

Then we have

$$\int_I g = \sum_{k=1}^n c_k(x_k - x_{k-1}) \le \sum_{k=1}^n d_k(x_k - x_{k-1}) = \int_I h.$$

□

Bounded functions

Let $f : I \to \mathbb{R}$ be a bounded function (not necessarily continuous or piecewise continuous). Define the sets $\mathcal{S}_f^+$ and $\mathcal{S}_f^-$ containing all step functions greater than f and less than f, respectively:

$$\mathcal{S}_f^+ = \{h \in \mathcal{S}(I) \;\mid\; f(x) \leq h(x),\ \forall x \in I\}$$
$$\mathcal{S}_f^- = \{g \in \mathcal{S}(I) \;\mid\; g(x) \leq f(x),\ \forall x \in I\}$$

These sets are not empty, since they contain, respectively, any constant function that is greater than the upper bound of f on I and any constant function that is smaller than the lower bound of f on I. We can therefore define:

Lower and upper integral

For $f : I \to \mathbb{R}$ that is bounded, we define the **upper integral of f on I** to be

$$\overline{\int_I} f = \inf\left\{\int_I h \;\mid\; h \in \mathcal{S}_f^+\right\},$$

and the **lower integral of f on I** to be

$$\underline{\int_I} f = \sup\left\{\int_I g \;\mid\; g \in \mathcal{S}_f^-\right\}.$$

Proposition 10.10: If f is bounded on I, then

$$\underline{\int_I} f \leq \overline{\int_I} f$$

Proof. Let $g \in \mathcal{S}_f^-$ and $h \in \mathcal{S}_f^+$. Then:

$$g(x) \leq f(x) \leq h(x), \qquad \forall x \in I.$$

It follows that

$$\int_I g \leq \int_I h.$$

Taking the infimum over all $h \in \mathcal{S}_f^+$ in this inequality gives us

$$\int_I g \leq \overline{\int_I} f$$

and now taking the supremum over all $g \in \mathcal{S}_f^-$ gives us

$$\underline{\int_I} f \leq \overline{\int_I} f.$$

□

The lower and upper integrals aren't necessarily equal

Consider the bounded function (called the *Dirichlet function*)

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

Regardless of the partition $\{x_k\}_{k=0}^n$ of I, any subinterval I_k will include both rational and irrational points, and therefore f will attain the values 0 and 1 on any such subinterval. Hence $\mathcal{S}_f^+$ will contain step functions whose values are at least 1 and $\mathcal{S}_f^-$ will contain step functions whose values are at most 0. It is not hard to conclude that

$$\underline{\int_I} f = 0 \qquad \text{and} \qquad \overline{\int_I} f = 1.$$

This example demonstrates that it is not evident that the lower and upper integral should be equal. In fact, it motivates the following definition:

Riemann integrable functions

A bounded function $f : I \to \mathbb{R}$ is said to be (Riemann) **integrable** on I if

$$\underline{\int_I} f = \overline{\int_I} f.$$

This value, called the definite integral, is denoted $\int_a^b f(x)\,\mathrm{d}x$ or $\int_I f(x)\,\mathrm{d}x$.

Theorem 10.11: The following functions are (Riemann) integrable on I:

1. Continuous functions on I.
2. Piecewise-continuous functions on I.
3. Functions that are continuous on (a, b) and bounded on $[a, b]$.
4. Monotone functions on $[a, b]$.

Proof. We skip this proof. □

Example 10.8: The function $f(x) = x$ is Riemann integrable (it was also Cauchy integrable). We saw that the result of the Cauchy integral was $\int_0^1 x\,\mathrm{d}x = \frac{1}{2}$. The Riemann integral will give the same result, and to see that one could take the step functions

$$h_n(x) = \begin{cases} 0 & x = 0 \\ \frac{k+1}{n} & \frac{k}{n} < x \le \frac{k+1}{n}, \ k = 0, \dots, n-1 \end{cases}$$

$$g_n(x) = \begin{cases} 0 & x = 0 \\ \frac{k}{n} & \frac{k}{n} < x \le \frac{k+1}{n},\ k = 0, \dots, n-1 \end{cases}$$

which satisfy $g_n(x) \le f(x) \le h_n(x)$ for any $x \in [0,1]$, and hence $h_n \in \mathcal{S}_f^+$ and $g_n \in \mathcal{S}_f^-$. It can easily be shown (can you show it?) that

$$\int_I h_n = \frac{1}{2} + \frac{1}{2n}$$

and

$$\int_I g_n = \frac{1}{2} - \frac{1}{2n}$$

so that

$$\overline{\int_I} f \le \inf_n \int_I h_n = \frac{1}{2} \qquad \text{and} \qquad \underline{\int_I} f \ge \sup_n \int_I g_n = \frac{1}{2}$$

hence

$$\overline{\int_I} f \le \frac{1}{2} \le \underline{\int_I} f.$$

But we know that

$$\underline{\int_I} f \le \overline{\int_I} f.$$

It must therefore hold that

$$\overline{\int_I} f = \frac{1}{2} = \underline{\int_I} f$$

and f is Riemann integrable on $I = [0,1]$ (and $\int_I f(x)\,\mathrm{d}x = \frac{1}{2}$).

Example 10.9: Here are two examples of functions that are Riemann integrable but not Cauchy integrable:

1. The function

$$f(x) = \begin{cases} \sin\frac{1}{x} & 0 < x \le 1, \\ 0 & x = 0, \end{cases}$$

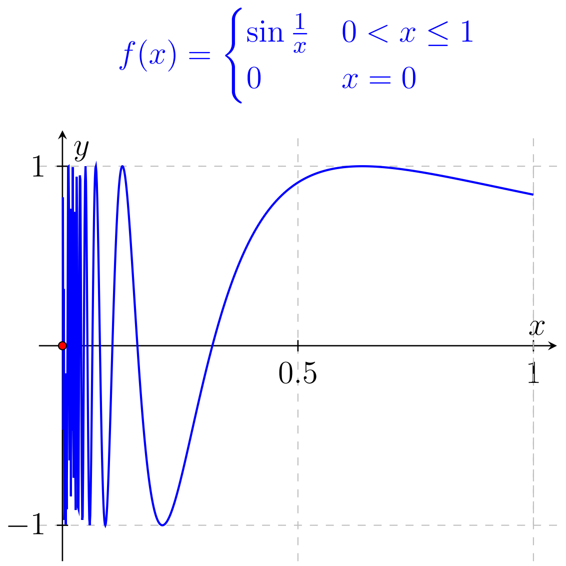
   is Riemann integrable (despite having infinitely many oscillations as $x \to 0$) as it falls under the third category of functions within the theorem.

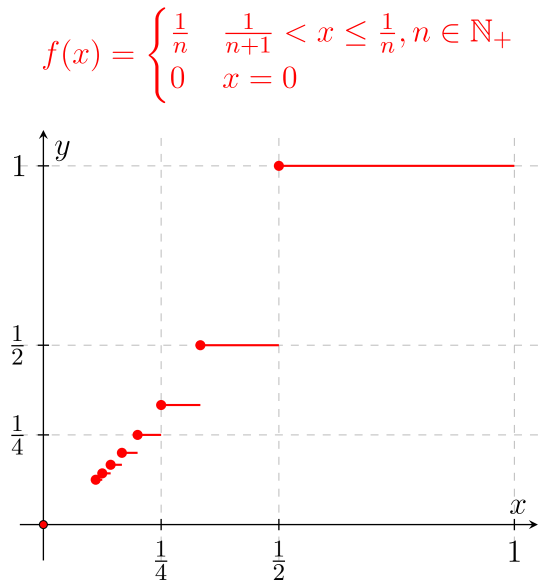
2. The function

$$f(x) = \begin{cases} \frac{1}{n} & \frac{1}{n+1} < x \le \frac{1}{n},\ n \in \mathbb{N}_+ \\ 0 & x = 0 \end{cases}$$

   is Riemann integrable (despite having infinitely many points with a jump discontinuity) as it falls under the fourth category of functions within the theorem.

Here are figures of both functions. Note that for the figure on the right, there are infinitely many ‘steps’ descending all the way to 0, but they become difficult to draw.





10.6 Properties of definite integrals

In this section we give some basic properties of the definite integral, without proof. Denote

$$\mathcal{R}([a,b]) = \text{the set of all integrable functions on } [a,b]$$

Definition 10.12: We define

$$\int_b^a f(x)\,\mathrm{d}x = -\int_a^b f(x)\,\mathrm{d}x$$

and if $a = b$ then

$$\int_a^a f(x)\,\mathrm{d}x = 0.$$

Properties (definite) integrable functions I

1. If $f \in \mathcal{R}([a,b])$ then it is also integrable on any subinterval of $[a,b]$.
2. If $f \in \mathcal{R}([a,b])$ then also $|f| \in \mathcal{R}([a,b])$.

Properties (definite) integrable functions II

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x$$

$$\int_a^b (\alpha f(x) + \beta g(x))\,\mathrm{d}x = \alpha\int_a^b f(x)\,\mathrm{d}x + \beta\int_a^b g(x)\,\mathrm{d}x$$

Properties (definite) integrable functions III

$$\begin{aligned}
f \geq 0,\ a < b \qquad &\Rightarrow \qquad \int_a^b f(x)\,\mathrm{d}x \geq 0 \quad (\text{equality iff } f \equiv 0)\\
f \leq g,\ a < b \qquad &\Rightarrow \qquad \int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x\\
a < b \qquad &\Rightarrow \qquad \left|\int_a^b f(x)\,\mathrm{d}x\right| \leq \int_a^b |f(x)|\,\mathrm{d}x
\end{aligned}$$

10.7 Integral mean value

Average (mean value)

Let $f \in \mathcal{R}([a,b])$. The **average (mean value) of f on** $[a,b]$ is the number

$$m(f;a,b) = \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x$$

it is often denoted

$$⨍_a^b f(x)\,\mathrm{d}x.$$

This can be rewritten as

$$(b-a)m(f;a,b) = \int_a^b f(x)\,\mathrm{d}x.$$

Theorem 10.13 (Integral Mean Value Theorem)**:** Let $f \in \mathcal{R}([a,b])$. Then

$$\inf_{x\in[a,b]} f(x) \leq m(f;a,b) \leq \sup_{x\in[a,b]} f(x).$$

If $f \in \mathcal{C}^0([a,b])$ (i.e., if f is continuous) then there exists $z \in [a,b]$ such that

$$m(f;a,b) = f(z).$$

Proof. We skip this proof. □

Example 10.10: Let f be the following *continuous* function on $[0,2]$:

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1,\\ 2 & 1 < x \leq 2.\end{cases}$$

Then:

$$\begin{aligned}\int_0^2 f(x)\,\mathrm{d}x &= \int_0^1 f(x)\,\mathrm{d}x + \int_1^2 f(x)\,\mathrm{d}x \\ &= \int_0^1 2x\,\mathrm{d}x + \int_1^2 2\,\mathrm{d}x = 2\int_0^1 x\,\mathrm{d}x + 2\int_1^2 1\,\mathrm{d}x = 1 + 2 = 3.\end{aligned}$$

Therefore

$$m(f;0,2) = \frac{3}{2}$$

and since f is continuous, it follows that there exists $z \in [0,2]$ such that $f(z) = \frac{3}{2}$. Indeed, this is true for $z = \frac{3}{4}$.

Remark: We can change the order of a and b in the definition of the average and the sign remains (since we get two minuses)

$$m(f;a,b) = \frac{1}{b-a}\int_a^b f(x)\,\mathrm{d}x = \frac{1}{a-b}\int_b^a f(x)\,\mathrm{d}x = m(f;b,a).$$

This makes sense: the average of a function shouldn't depend on whether we measure it going on the x axis to the right or to the left.

10.8 Fundamental Theorem of Integral Calculus

Integral function

Let $f : \mathbb{R} \to \mathbb{R}$ be integrable on any finite interval. Fix some $x_0 \in \mathbb{R}$. We define the **integral function of** f to be

$$F(x) = \int_{x_0}^{x} f(y)\,\mathrm{d}y.$$

Observe that the integral function depends on the choice of x_0, and $F(x_0) = 0$.

Fundamental Theorem of Integral Calculus

Let $f : \mathbb{R} \to \mathbb{R}$ be defined and continuous on an interval (possibly unbounded) $I \subseteq \mathbb{R}$. Then for any $x_0 \in I$, $F(x)$ is differentiable on I and

$$F'(x) = f(x), \qquad \forall x \in I.$$

Proof. Fix x in the interior of I (i.e. it is not on the boundary of I). Let Δx be small enough (positive or negative) so that $x + \Delta x$ belongs to I. Then, using the definition of F and the fact that

$$F(x+\Delta x) = \int_{x_0}^{x+\Delta x} f(y)\,\mathrm{d}y$$

we have

$$\begin{aligned}\frac{F(x+\Delta x)-F(x)}{\Delta x} &= \frac{1}{\Delta x}\left(\int_{x_0}^{x+\Delta x} f(y)\,\mathrm{d}y - \int_{x_0}^{x} f(y)\,\mathrm{d}y\right)\\ &= \frac{1}{\Delta x}\int_{x}^{x+\Delta x} f(y)\,\mathrm{d}y\\ &= m(f; x, x+\Delta x).\end{aligned}$$

Since f is continuous, the Integral Mean Value Theorem (Theorem 10.13) implies that there exists z between x and $x+\Delta x$ such that $m(f; x, x+\Delta x) = f(z)$. So we have

$$\frac{F(x+\Delta x)-F(x)}{\Delta x} = f(z).$$

Observe that z depends on the choice of Δx, so we should write $z = z(\Delta x)$. Necessarily

$$\lim_{\Delta x\to 0} z(\Delta x) = x$$

(by the Squeeze Theorem). Since f is continuous, we have

$$\lim_{\Delta x\to 0} f(z(\Delta x)) = f(\lim_{\Delta x\to 0} z(\Delta x)) = f(x).$$

Hence we have:

$$F'(x) = \lim_{\Delta x\to 0}\frac{F(x+\Delta x)-F(x)}{\Delta x} = \lim_{\Delta x\to 0} f(z(\Delta x)) = f(x).$$

This proves the theorem for the case where x lies in the interior of I. If x is on the boundary then we must take one-sided limits, but the details are very similar. □

This theorem tells us how to define a (specific) antiderivative (depending on our choice of x_0): $F(x) = \int_{x_0}^{x} f(y)\,\mathrm{d}y$. Now we can state a result that links this to any other antiderivative:

Corollary 10.14: If we define $F_{x_0}(x) = \int_{x_0}^{x} f(y)\,\mathrm{d}y$, then

$$F_{x_0}(x) = G(x) - G(x_0)$$

for any other G which is an antiderivative of f.

Proof. The proof is immediate by plugging in $x = x_0$ in the above expression, since we know that all antiderivatives differ by a constant □

Important corollary

Let $f : \mathbb{R} \to \mathbb{R}$ be continuous on $[a, b]$ and let G be any antiderivative. Then

$$\int_a^b f(x)\,\mathrm{d}x = G(b) - G(a).$$

Proof. Let F_a be the antiderivative defined with the choice $x_0 = a$. Then

$$\int_a^b f(x)\,\mathrm{d}x = F_a(b).$$

By Corollary 10.14 we then further have

$$\int_a^b f(x)\,\mathrm{d}x = F_a(b) = G(b) - G(a)$$

for any antiderivative G. □

Notation

Instead of $G(b) - G(a)$, we often write

$$\big[G(x)\big]_a^b \qquad \text{or} \qquad G(x)\big|_a^b$$

Example 10.11:

$$\int_0^1 x^2\,\mathrm{d}x = \left[\frac{1}{3}x^3\right]_0^1 = \frac{1}{3}\cdot 1^3 - \frac{1}{3}\cdot 0^3 = \frac{1}{3}$$

$$\int_\pi^{2\pi} \sin x\,\mathrm{d}x = [-\cos x]_\pi^{2\pi} = -\cos 2\pi - (-\cos\pi) = -1 - 1 = -2$$

$$\int_2^6 \frac{1}{x}\,\mathrm{d}x = [\ln x]_2^6 = \ln 6 - \ln 2 = \ln 3$$

$$\int_{-1}^1 e^{2x}\,\mathrm{d}x = \left[\frac{1}{2}e^{2x}\right]_{-1}^1 = \frac{1}{2}e^2 - \frac{1}{2}e^{-2} = \frac{e^2 - e^{-2}}{2} = \sinh(2)$$

$$\int_0^1 \frac{1}{1+x^2}\,\mathrm{d}x = [\arctan x]_0^1 = \arctan 1 - \arctan 0 = \frac{\pi}{4} - 0 = \frac{\pi}{4}$$

$$\int_0^4 \sqrt{x}\,\mathrm{d}x = \left[\frac{2}{3}x^{3/2}\right]_0^4 = \frac{2}{3}(4^{3/2}) - \frac{2}{3}(0) = \frac{2}{3}(8) = \frac{16}{3}$$

$$\int_1^3 \frac{x}{x^2+1}\,\mathrm{d}x = \left[\frac{1}{2}\ln(x^2+1)\right]_1^3 = \frac{1}{2}\ln(10) - \frac{1}{2}\ln(2) = \frac{1}{2}\ln(5)$$

$$\int_0^\pi \cos x\,\mathrm{d}x = [\sin x]_0^\pi = \sin\pi - \sin 0 = 0 - 0 = 0$$

Corollary 10.15: If $f \in \mathcal{C}^1(I)$ (that is, f is differentiable on I and f' is continuous on I), then for any $x_0 \in I$,

$$f(x) = f(x_0) + \int_{x_0}^x f'(y)\,\mathrm{d}y.$$

Proof. This follows immediately from the previous corollary, and the proof is left as an exercise. □

Application to Maclaurin expansions

The following lemma will help us express the Maclaurin polynomials of some functions whose derivative is easy to study.

Lemma 10.16: Integration increases by 1 the order of decay of an infinitesimal function:

$$\int_0^x o(y^\alpha)\,\mathrm{d}y = o(x^{\alpha+1}), \qquad \text{as } x \to 0.$$

Proof. We skip this proof. □

The function $f(x) = \arctan x$.

We recall that

$$\frac{\mathrm{d}}{\mathrm{d}x}\arctan(x) = \frac{1}{1+x^2}$$

so that

$$\arctan x = \int_0^x \frac{1}{1+y^2}\,\mathrm{d}y.$$

The Maclaurin polynomial of $\frac{1}{1+y^2}$ is known to be (see Section 9.2):

$$\frac{1}{1+y^2} = 1 - y^2 + y^4 - y^6 - \cdots + (-1)^n y^{2n} + o(y^{2n+1}) = \sum_{k=0}^{n}(-1)^k y^{2k} + o(y^{2n+1}).$$

Combining these two facts, and using the lemma, we find that as $x \to 0$,

$$\boxed{\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + o(x^{2n+2}) = \sum_{k=0}^{n}(-1)^k \frac{x^{2k+1}}{2k+1} + o(x^{2n+2})}$$

The function $f(x) = \arcsin x$.

We recall that

$$\frac{\mathrm{d}}{\mathrm{d}x}\arcsin(x) = \frac{1}{\sqrt{1-x^2}}$$

so that

$$\arcsin x = \int_0^x \frac{1}{\sqrt{1-y^2}}\,\mathrm{d}y.$$

The Maclaurin polynomial of $\frac{1}{\sqrt{1-y^2}}$ is known to be (see Section 9.2):

$$\frac{1}{\sqrt{1-y^2}} = 1 + \frac{1}{2}y^2 + \frac{3}{8}y^4 + \frac{5}{16}y^6 + \cdots + \left|\binom{-\frac{1}{2}}{n}\right| y^{2n} + o(y^{2n+1}) = \sum_{k=0}^{n}\left|\binom{-\frac{1}{2}}{k}\right| y^{2k} + o(y^{2n+1}).$$

Combining these two facts, and using the lemma, we find that as $x \to 0$,

$$
\begin{aligned}
\arcsin x &= x + \frac{1}{6}x^3 + \frac{3}{40}x^5 + \frac{5}{112}x^7 + \cdots + \left|\binom{-\frac{1}{2}}{n}\right| \frac{x^{2n+1}}{2n+1} + o(x^{2n+2}) \\
&= \sum_{k=0}^{n} \left|\binom{-\frac{1}{2}}{k}\right| \frac{x^{2k+1}}{2k+1} + o(x^{2n+2})
\end{aligned}
$$

Application to the remainder of a Taylor expansion

We have seen Peano's remainder:

$$f(x) - (Tf)_{n,x_0}(x) = o((x - x_0)^n), \qquad x \to x_0.$$

and Lagrange's remainder:

$$f(x) - (Tf)_{n,x_0}(x) = \frac{1}{(n+1)!} f^{(n+1)}(\overline{x})(x - x_0)^{n+1},$$

where $\overline{x}$ is some point between x and x_0. These lead us to:

Taylor formula with integral remainder

Let $n \in \mathbb{N}$ and suppose that $f \in \mathcal{C}^{n+1}(I)$ in some neighborhood I of x_0. Then

$$f(x) - (Tf)_{n,x_0}(x) = \frac{1}{n!} \int_{x_0}^{x} f^{(n+1)}(y)(x - y)^n \,\mathrm{d}y.$$

Proof. The proof, which we skip here, relies on induction. □

Remark: Observe that for $n = 0$ this result is precisely Corollary 10.15:

$$f(x) - f(x_0) = \int_{x_0}^{x} f'(y) \,\mathrm{d}y.$$

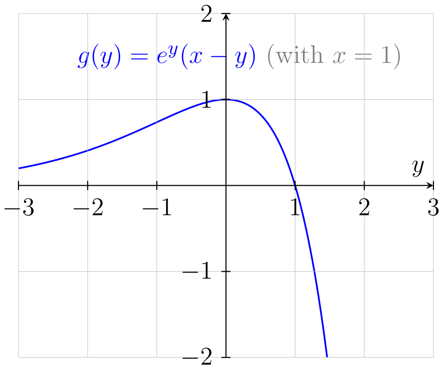
Example 10.12: Let us compare approximation of the number e using Lagrange's remainder and the integral remainder. Taking order $n = 1$, we have

$$
\begin{aligned}
\text{Lagrange:} \qquad & e^x = 1 + x + \frac{1}{2}e^{\overline{x}}x^2, \\
\text{Integral:} \qquad & e^x = 1 + x + \int_0^x e^y (x - y) \,\mathrm{d}y.
\end{aligned}
$$

Lagrange: Since the exponential function is strictly increasing, we can deduce that

$$0 < e^x - (1 + x) = \frac{1}{2}e^{\overline{x}}x^2 < \frac{1}{2}e^x x^2$$

Integral: consider the integrand $g(y) = e^y(x - y)$. We have $g'(y) = e^y(x - y - 1)$. Searching for extrema for $x \geq 1$ we impose $g'(y) = 0$ to find $y = x - 1$. For $y < x - 1$, $g'(y) > 0$ and for $y > x - 1$, $g'(y) < 0$. Hence $y = x - 1$ is a global maximum.



Therefore:

$$0 < \int_0^x e^y(x-y)\,\mathrm{d}y < e^{x-1}\int_0^x \mathrm{d}y = e^{x-1}x = \frac{1}{e}e^x x$$

and it follows that

$$0 < e^x - (1+x) = \int_0^x e^y(x-y)\,\mathrm{d}y < \frac{1}{e}e^x x, \quad \forall x \geq 1.$$

Comparing the two bounds we've obtained, we have

$$\begin{aligned} \text{Lagrange:} \qquad & 0 < e^x - (1+x) < \frac{1}{2}e^x x^2 \\ \text{Integral:} \qquad & 0 < e^x - (1+x) < \frac{1}{e}e^x x \end{aligned}$$

We see that for $x \geq 1$, the error with the integral remainder is smaller, since $\frac{1}{e} < \frac{1}{2}$ and $x \leq x^2$.

10.9 Rules of definite integration

Even and odd functions

Proposition 10.17: Let f be integrable on the interval $[-a, a]$ (where $a > 0$). Then

$$\begin{aligned} f \text{ even} \qquad & \Rightarrow \qquad \int_{-a}^{a} f(x)\,\mathrm{d}x = 2\int_0^a f(x)\,\mathrm{d}x, \\ f \text{ odd} \qquad & \Rightarrow \qquad \int_{-a}^{a} f(x)\,\mathrm{d}x = 0. \end{aligned}$$

Proof. This is left as an exercise. □

Integration by parts and by substitution

We now want to write the formulas for integration by parts and for integration by substitution for definite integrals:

Integration by parts (definite integrals)

Let $f, g \in \mathcal{C}^1([a,b])$. Then

$$\int_a^b f(x)g'(x)\,\mathrm{d}x = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x)\,\mathrm{d}x.$$

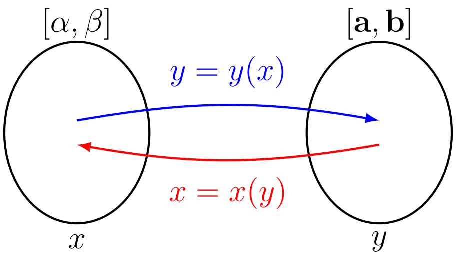
Integration by substitution (definite integrals)

Let $f(y)$ be continuous on an interval $[a,b]$ and let $y(x) : [\alpha,\beta] \to [a,b]$ belong to $\mathcal{C}^1([\alpha,\beta])$. Then

$$\int_\alpha^\beta f(y(x))y'(x)\,\mathrm{d}x = \int_{y(\alpha)}^{y(\beta)} f(y)\,\mathrm{d}y.$$

If $y(x)$ is 1-1 then we also have

$$\int_a^b f(y)\,\mathrm{d}y = \int_{y^{-1}(a)}^{y^{-1}(b)} f(y(x))y'(x)\,\mathrm{d}x.$$



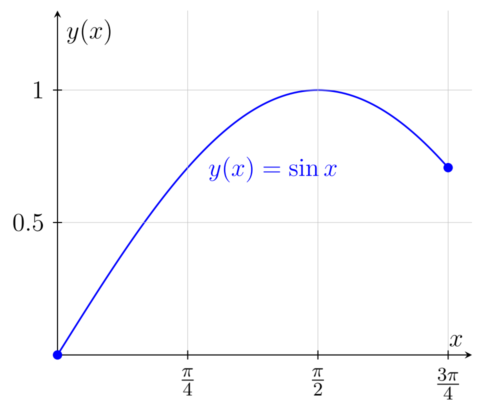
Remembering how to integrate by substitution

When asked to evaluate the definite integral $\int_\alpha^\beta g(x)\,\mathrm{d}x$ there are two options:

1. We are able identify that there exists $y = y(x)$ such that $g(x)$ has the form
$$g(x) = f(y(x))y'(x).$$
In this case, since $\frac{\mathrm{d}y}{\mathrm{d}x} = y'(x)$ we write $\mathrm{d}y = y'(x)\,\mathrm{d}x$ to get
$$\int_\alpha^\beta g(x)\,\mathrm{d}x = \int_\alpha^\beta f(y(x))y'(x)\,\mathrm{d}x = \int_a^b f(y)\,\mathrm{d}y.$$
This approach might work if g is a complicated function.

2. If we cannot identify $y = y(x)$ as above, we try to go about it the other way around: identify x as a function of y: $x = x(y)$, compute $\frac{\mathrm{d}x}{\mathrm{d}y} = x'(y)$ and write $\mathrm{d}x = x'(y)\,\mathrm{d}y$ to get:
$$\int_\alpha^\beta g(x)\,\mathrm{d}x = \int_a^b g(x(y))x'(y)\,\mathrm{d}y.$$
This approach might work if g is a simple function. **In this second case, we need to be sure to check that $x = x(y)$ is invertible. This means that $x(y)$ must be strictly monotone on $[a, b]$.**

Example 10.13: Compute
$$\int_0^{\frac{3\pi}{4}} \sin^3 x \cos x\,\mathrm{d}x.$$
Define
$$y(x) = \sin x \qquad \text{so that} \qquad \frac{\mathrm{d}y}{\mathrm{d}x} = \cos x \qquad \text{and} \qquad [0, \frac{3\pi}{4}] \to [0, \frac{\sqrt{2}}{2}]$$



and we have
$$\int_0^{\frac{3\pi}{4}} \sin^3 x \cos x\,\mathrm{d}x = \int_0^{\frac{\sqrt{2}}{2}} y^3\,\mathrm{d}y = \left[\frac{1}{4}y^4\right]_0^{\frac{\sqrt{2}}{2}} = \frac{1}{16}.$$

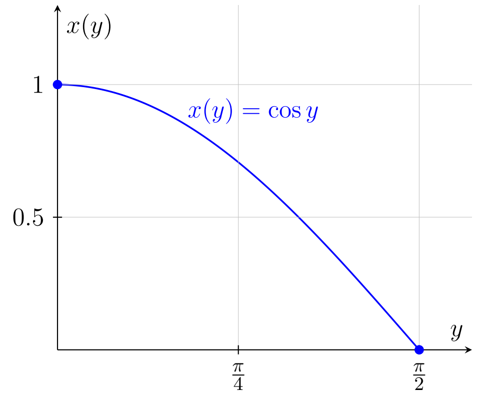
Here we observe that we used the substitution $y = y(x)$, but y is not 1-1, so it is not invertible, and we *wouldn't* have been able to use the second method.

Example 10.14: Compute

$$\int_0^1 \arcsin\sqrt{1-x^2}\,\mathrm{d}x.$$

Define

$$x(y) = \cos y \qquad \text{so that} \qquad \frac{\mathrm{d}x}{\mathrm{d}y} = -\sin y \qquad \text{and} \qquad [0,1] \to [\frac{\pi}{2}, 0]$$

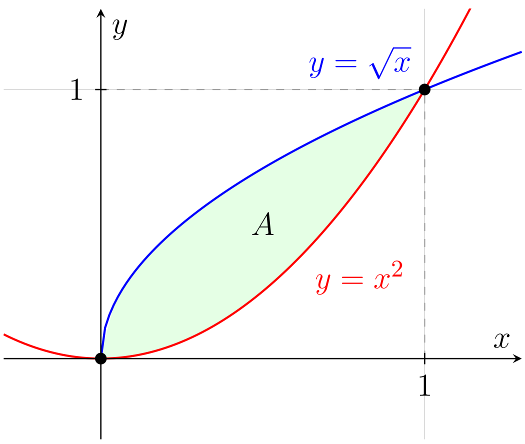


which is a *bijective* function and therefore *permitted.* So we have

$$\begin{aligned}
\int_0^1 \arcsin\sqrt{1-x^2}\,\mathrm{d}x &= \int_{\frac{\pi}{2}}^0 \arcsin\sqrt{1-\cos^2 y}\,(-\sin y)\,\mathrm{d}y \\
&= \int_0^{\frac{\pi}{2}} y\sin y\,\mathrm{d}y \\
\textit{(integration by parts)} \qquad &= [-y\cos y]_0^{\frac{\pi}{2}} + \int_0^{\frac{\pi}{2}} \cos y\,\mathrm{d}y \\
&= 0 + [\sin y]_0^{\frac{\pi}{2}} \\
&= 1.
\end{aligned}$$

Computation of areas

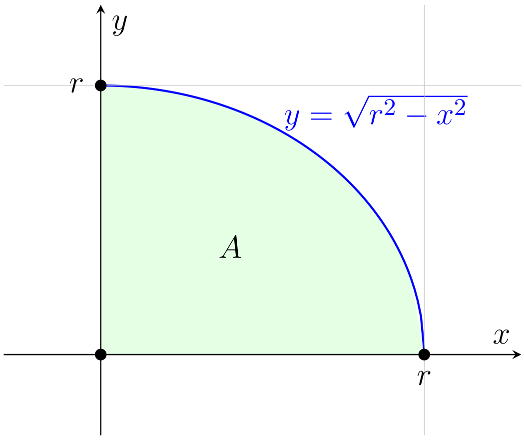
Example 10.15: Compute the area enclosed A between the graphs of $y = \sqrt{x}$ and $y = x^2$.



We identify that the intersection points are at $x = 0, 1$ and that the graph of $y = \sqrt{x}$ lies above the graph of $y = x^2$. We therefore have

$$\begin{aligned} A &= \int_0^1 \sqrt{x}\,\mathrm{d}x - \int_0^1 x^2\,\mathrm{d}x \\ &= \left[\frac{2}{3}x^{\frac{3}{2}}\right]_0^1 - \left[\frac{1}{3}x^3\right]_0^1 \\ &= \frac{2}{3} - \frac{1}{3} = \frac{1}{3}. \end{aligned}$$

Example 10.16: Prove that the area of a disc of radius r is $A(r) = \pi r^2$.



The area is four times the area of a quarter disc:

$$A(r) = 4\int_0^r \sqrt{r^2 - x^2}\,\mathrm{d}x$$

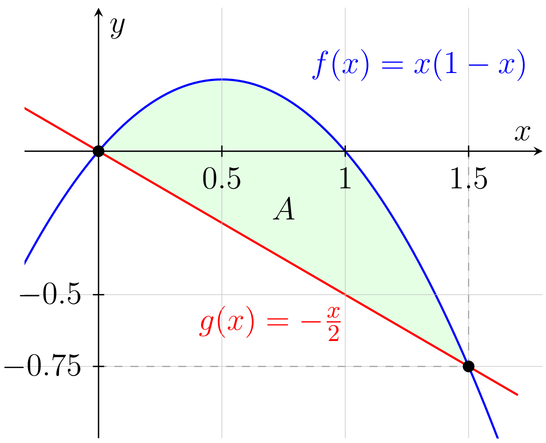
Make the (very simple, linear) change of variables:

$$x(z) = rz \qquad \text{so that} \qquad \frac{\mathrm{d}x}{\mathrm{d}z} = r \qquad \text{and} \qquad [0, r] \to [0, 1]$$

is the inverse of the invertible function $z(x) = \frac{x}{r}$ (so the change of variables is allowed) and we have

$$\begin{aligned} A &= 4\int_0^r \sqrt{r^2 - x^2}\,\mathrm{d}x \\ &= 4r^2 \int_0^1 \sqrt{1 - z^2}\,\mathrm{d}z \\ \text{(see Example 10.5(6))} \qquad &= 4r^2\left[\frac{1}{2}z\sqrt{1 - z^2} + \frac{1}{2}\arcsin z\right]_0^1 \\ &= 4r^2\left[(0 - 0) + \left(\frac{\pi}{4} - 0\right)\right] \\ &= \pi r^2. \end{aligned}$$

Example 10.17: Compute the area enclosed between the graphs of $f(x) = x(1 - x)$ and $g(x) = -\frac{x}{2}$.



Even though some of the area lies below the x axis, it is still given by

$$A = \int_0^{\frac{3}{2}} \left[x(1-x) - \left(-\frac{x}{2}\right)\right] \, \mathrm{d}x$$

Can you think why? In the above expression we have already computed the intersection points to be $x = 0$ and $x = \frac{3}{2}$. So we compute:

$$\begin{aligned} A &= \int_0^{\frac{3}{2}} \left(-x^2 + \frac{3}{2}x\right) \, \mathrm{d}x \\ &= \left[-\frac{1}{3}x^3 + \frac{3}{4}x^2\right]_0^{\frac{3}{2}} \\ &= \frac{9}{16}. \end{aligned}$$

10.10 Differentiation of integrals with functional limits

The standard form of the Fundamental Theorem of Integral Calculus provides the derivative of the integral function $F(x) = \int_{x_0}^{x} f(y) \, \mathrm{d}y$, where the upper limit is x and the lower limit x_0 is a constant.

We now extend this result to cases where the limits of integration are functions of x. This extension requires the application of the **Chain Rule**. Let $f : \mathbb{R} \to \mathbb{R}$ be a function continuous on its domain, and let $u(x)$ and $v(x)$ be differentiable functions. We want to find the derivative of the function $K(x)$ defined by:

$$K(x) = \int_{v(x)}^{u(x)} f(y) \, \mathrm{d}y.$$

Case 1: The Upper Limit is a Function of x

Consider $G(x)$ defined by:

$$G(x) = \int_{x_0}^{u(x)} f(y) \, \mathrm{d}y$$

where x_0 is a constant and $u(x)$ is differentiable.

We view $G(x)$ as a composite function $G(x) = F(u(x))$, where $F(u)$ is the integral function:

$$F(u) = \int_{x_0}^{u} f(y)\,\mathrm{d}y$$

By the Chain Rule, $G'(x) = F'(u(x)) \cdot u'(x)$. So we have:

Differentiation with Upper Functional Limit

The derivative is given by:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_{x_0}^{u(x)} f(y)\,\mathrm{d}y\right] = f(u(x)) \cdot u'(x)$$

Example 10.18: Find $\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_1^{\cos x} \frac{e^y}{y+2}\,\mathrm{d}y\right]$.

Here $f(y) = \frac{e^y}{y+2}$ and $u(x) = \cos x$, so $u'(x) = -\sin x$.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_1^{\cos x} \frac{e^y}{y+2}\,\mathrm{d}y\right] &= f(\cos x) \cdot \frac{\mathrm{d}}{\mathrm{d}x}(\cos x) \\ &= \frac{e^{\cos x}}{\cos x + 2} \cdot (-\sin x) \\ &= -\frac{\sin x \cdot e^{\cos x}}{\cos x + 2}\end{aligned}$$

Case 2: The Lower Limit is a Function of x

Consider $H(x)$ defined by:

$$H(x) = \int_{v(x)}^{x_0} f(y)\,\mathrm{d}y$$

where x_0 is a constant and $v(x)$ is differentiable. We use the integral property $\int_a^b f(y)\,\mathrm{d}y = -\int_b^a f(y)\,\mathrm{d}y$:

$$H(x) = -\int_{x_0}^{v(x)} f(y)\,\mathrm{d}y$$

Applying the result from Case 1 (and the linearity of the derivative):

Differentiation with Lower Functional Limit

The derivative is given by:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_{v(x)}^{x_0} f(y)\,\mathrm{d}y\right] = -f(v(x)) \cdot v'(x)$$

Case 3: Both Limits are Functions of x

To find the derivative of $K(x) = \int_{v(x)}^{u(x)} f(y)\,\mathrm{d}y$, we split the integral at any constant c in the domain of f:

$$K(x) = \int_{v(x)}^{c} f(y)\,\mathrm{d}y + \int_{c}^{u(x)} f(y)\,\mathrm{d}y$$

Rewriting the first term using the reversal property:

$$K(x) = -\int_{c}^{v(x)} f(y)\,\mathrm{d}y + \int_{c}^{u(x)} f(y)\,\mathrm{d}y$$

Differentiating term-by-term using the results from Case 1 and Case 2:

$$\frac{\mathrm{d}}{\mathrm{d}x}K(x) = -\big(f(v(x)) \cdot v'(x)\big) + \big(f(u(x)) \cdot u'(x)\big)$$

General Case: The Leibniz Integral Rule

Let $u(x)$ and $v(x)$ be differentiable functions. The derivative is:

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_{v(x)}^{u(x)} f(y)\,\mathrm{d}y\right] = f(u(x)) \cdot u'(x) - f(v(x)) \cdot v'(x)$$

Example 10.19: Find $\frac{\mathrm{d}}{\mathrm{d}x}\left[\int_{\sqrt{x}}^{x^2} \sin(y^3)\,\mathrm{d}y\right]$ for $x > 0$.

Here, $f(y) = \sin(y^3)$, $u(x) = x^2$ (so that $u'(x) = 2x$), and $v(x) = \sqrt{x}$ (so that $v'(x) = \frac{1}{2\sqrt{x}}$).

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}K(x) &= f(u(x))u'(x) - f(v(x))v'(x) \\
&= \sin((x^2)^3) \cdot (2x) - \sin((\sqrt{x})^3) \cdot \left(\frac{1}{2\sqrt{x}}\right) \\
&= 2x\sin(x^6) - \frac{1}{2\sqrt{x}}\sin(x^{3/2})
\end{aligned}$$

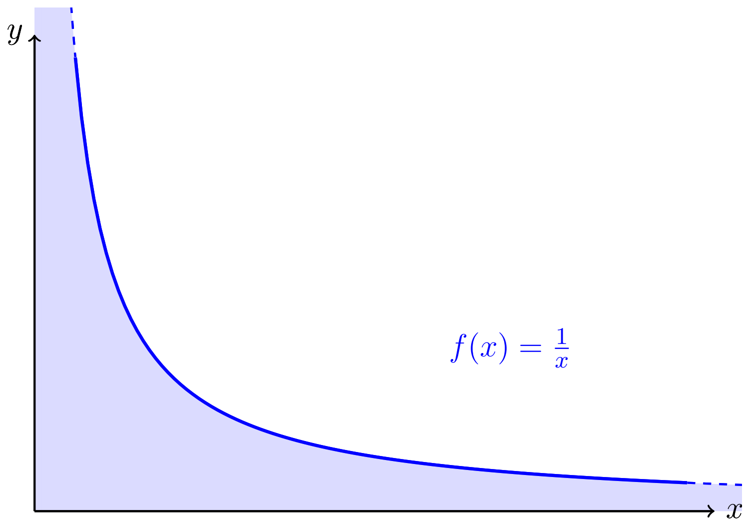
Chapter 11

Improper integrals and numerical series

In this chapter we will extend the notion of a definite integral $\int_a^b f(x)\,\mathrm{d}x$ to cases where either

- one (or both) of the endpoints is at $\pm\infty$,
- $f(x)$ has a vertical asymptote at a and/or b.

These are called **improper integrals**.



This will allow us to ask what is the area under the graph of functions such as $f(x) = \frac{1}{x}$ between 0 and $+\infty$:

$$\int_0^{+\infty} \frac{1}{x}\,\mathrm{d}x = ?$$

A topic directly related to this is that of (infinite) **numerical series**: we could ask (very much related to the previous question):

$$\sum_{n=1}^{\infty} \frac{1}{n} = ?$$

11.1 Improper integrals

11.1.1 Type I improper integrals

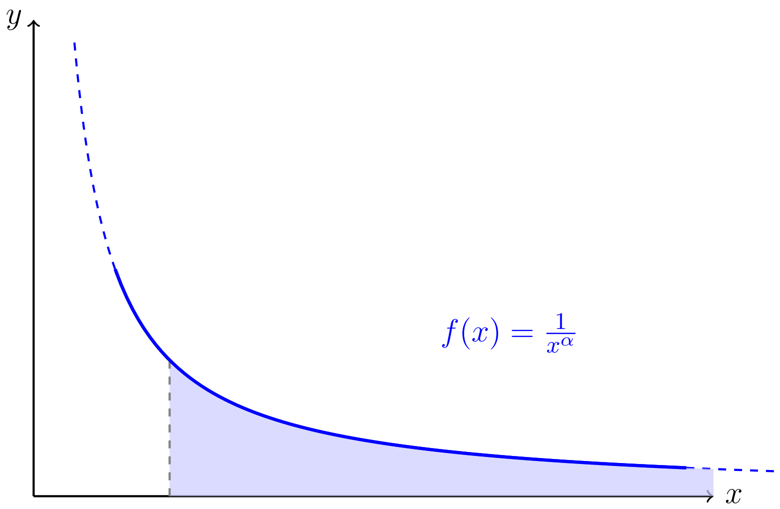
We start with improper integrals that have a limit at infinity:

Improper integral (type I)

Let $f : \mathbb{R} \to \mathbb{R}$ be a function integrable on $[a, b]$ for any $b > a$. We define the **improper integral** (type I) of f on $[a, +\infty)$ to be

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \lim_{b \to +\infty} \int_a^b f(x)\,\mathrm{d}x.$$

It will *converge*, *diverge* or be *indeterminate* depending on the limit on the right hand side.



Example 11.1: Investigate the convergence of the integral

$$\int_1^{+\infty} \frac{1}{x^\alpha}\,\mathrm{d}x,$$

where $\alpha > 0$. We use $a = 1$ to avoid issues at $x = 0$.

We must analyze two main cases: $\alpha = 1$ and $\alpha \neq 1$.

Case 1: $\alpha = 1$**.** In this case, the integrand is $f(x) = \frac{1}{x}$. We compute the definite integral first:

$$\int_1^b \frac{1}{x}\,\mathrm{d}x = [\ln|x|]_1^b = \ln|b| - \ln|1| = \ln b$$

Now we take the limit as $b \to +\infty$:

$$\int_1^{+\infty} \frac{1}{x}\,\mathrm{d}x = \lim_{b \to +\infty} (\ln b) = +\infty.$$

Therefore, the integral *diverges* for $\alpha = 1$.

Case 2: $\alpha \neq 1$**.** We have:

$$\int_1^b \frac{1}{x^\alpha}\,\mathrm{d}x = \left[\frac{1}{1-\alpha}x^{1-\alpha}\right]_1^b = \frac{1}{1-\alpha}\left(b^{1-\alpha}-1\right)$$

Now we analyze the limit $\lim_{b\to+\infty} b^{1-\alpha}$ based on the sign of the exponent $1-\alpha$:

- If $\mathbf{1}-\alpha<\mathbf{0}$ (i.e., $\alpha>\mathbf{1}$): Since $1-\alpha$ is negative, we rewrite $b^{1-\alpha}$ by moving it to the denominator with a positive exponent:

$$b^{1-\alpha} = \frac{1}{b^{-(1-\alpha)}} = \frac{1}{b^{\alpha-1}}$$

  Since $\alpha-1>0$, as $\lim_{b\to+\infty}\frac{1}{b^{\alpha-1}} = 0$. The limit is:

$$\int_1^{+\infty} \frac{1}{x^\alpha}\,\mathrm{d}x = \lim_{b\to+\infty}\frac{1}{1-\alpha}\left(\frac{1}{b^{\alpha-1}}-1\right) = \frac{1}{1-\alpha}(0-1) = \frac{-1}{1-\alpha} = \frac{1}{\alpha-1}.$$

  The integral *converges* to $\frac{1}{\alpha-1}$ when $\alpha>\mathbf{1}$.

- If $\mathbf{1}-\alpha>\mathbf{0}$ (i.e., $\alpha<\mathbf{1}$): Since $1-\alpha$ is positive, as $\lim_{b\to+\infty} b^{1-\alpha} = +\infty$. The limit is:

$$\int_1^{+\infty} \frac{1}{x^\alpha}\,\mathrm{d}x = \lim_{b\to+\infty}\frac{1}{1-\alpha}(b^{1-\alpha}-1) = +\infty \quad (\text{since } 1-\alpha\neq 0).$$

  The integral *diverges* for $\alpha<\mathbf{1}$.

The choice of lower bound 1 was convenient but arbitrary. We could choose any $a>0$. We therefore conclude:

$$\int_a^{+\infty} \frac{1}{x^\alpha}\,\mathrm{d}x \qquad \begin{cases}\text{converges} & \text{if } \alpha>1,\\ \text{diverges} & \text{if } 0<\alpha\leq 1.\end{cases}$$

Convergence tests

Now we give some criteria for the convergence/divergence of improper integrals of type I.

Comparison Test

Fix $a \in \mathbb{R}$. Let f, g be integrable on $[a, b]$ for any $b > a$. Assume that $0 \leq f(x) \leq g(x)$ for all $x \in [a, +\infty)$. Then

$$0 \leq \int_a^{+\infty} f(x)\,\mathrm{d}x \leq \int_a^{+\infty} g(x)\,\mathrm{d}x$$

Therefore

$$\begin{aligned} \int_a^{+\infty} g(x)\,\mathrm{d}x < +\infty \ \textit{(converges)} \quad &\Rightarrow \quad \int_a^{+\infty} f(x)\,\mathrm{d}x < +\infty \ \textit{(converges)} \\ \int_a^{+\infty} f(x)\,\mathrm{d}x = +\infty \ \textit{(diverges)} \quad &\Rightarrow \quad \int_a^{+\infty} g(x)\,\mathrm{d}x = +\infty \ \textit{(diverges)} \end{aligned}$$

Example 11.2: We determine whether the integrals

$$\int_1^{+\infty} \frac{\arctan x}{x^2}\,\mathrm{d}x \qquad \text{and} \qquad \int_1^{+\infty} \frac{\arctan x}{x}\,\mathrm{d}x$$

converge or diverge. Recall that $\arctan 1 = \frac{\pi}{4}$ and $\arctan x$ is a strictly increasing function, with $\lim_{x\to+\infty} \arctan x = \frac{\pi}{2}$. Hence

$$\frac{\pi}{4} \leq \arctan x \leq \frac{\pi}{2}, \qquad \forall x \in [1, +\infty).$$

It follows that for all $x \in [1, +\infty)$,

$$\frac{\arctan x}{x^2} \leq \frac{\pi}{2}\frac{1}{x^2} \qquad \text{and} \qquad \frac{\pi}{4}\frac{1}{x} \leq \frac{\arctan x}{x}$$

We therefore have that

$$\int_1^{+\infty} \frac{\arctan x}{x^2}\,\mathrm{d}x \leq \frac{\pi}{2}\int_1^{+\infty} \frac{1}{x^2}\,\mathrm{d}x < +\infty \ \textit{(converges)}$$

and

$$\int_1^{+\infty} \frac{\arctan x}{x}\,\mathrm{d}x \geq \frac{\pi}{4}\int_1^{+\infty} \frac{1}{x}\,\mathrm{d}x = +\infty \ \textit{(diverges)}$$

If $f = \mathcal{O}(g)$

If $f = \mathcal{O}(g)$ are non-negative as $x \to +\infty$, then if the improper integral (type I) of g converges, then so does the integral of f.

Example 11.3: Since $e^{-x^2} = o(x^{-2})$ as $x \to +\infty$, we deduce that $\int_0^{+\infty} e^{-x^2}\,\mathrm{d}x < +\infty$ *(converges)*.

Absolute Convergence Test

Fix $a \in \mathbb{R}$. Suppose that both f and $|f|$ are integrable on any interval $[a, b]$ for any $b > a$. Then if the improper integral (type I) of $|f|$ converges, so does the integral of f, and

$$\left|\int_a^{+\infty} f(x)\,\mathrm{d}x\right| \leq \int_a^{+\infty} |f(x)|\,\mathrm{d}x.$$

Example 11.4: Consider the function $f(x) = \frac{\cos x}{x^2}$. Then $|f(x)| \leq \frac{1}{x^2}$. So $|f(x)|$ is integrable on $[1, +\infty)$ by the Comparison Test. It follows that $f(x)$ is integrable by the Absolute Convergence Test:

$$\left|\int_1^{+\infty} \frac{\cos x}{x^2}\,\mathrm{d}x\right| \leq \int_1^{+\infty} \left|\frac{\cos x}{x^2}\right|\,\mathrm{d}x \leq \int_1^{+\infty} \frac{1}{x^2}\,\mathrm{d}x = 1.$$

Remark: The converse is not necessarily true:

$$\int_a^{+\infty} f(x)\,\mathrm{d}x < +\infty \qquad \text{does } \textit{not} \text{ imply that} \qquad \int_a^{+\infty} |f(x)|\,\mathrm{d}x < +\infty.$$

Asymptotic Comparison Test

Suppose that $f \sim \frac{1}{x^\alpha}$ as $x \to +\infty$. Then:

$$\alpha > 1 \quad \Rightarrow \quad \int_a^{+\infty} f(x)\,\mathrm{d}x < +\infty \text{ (converges)}$$

$$\alpha \leq 1 \quad \Rightarrow \quad \int_a^{+\infty} f(x)\,\mathrm{d}x = +\infty \text{ (diverges)}$$

Example 11.5: Investigate

$$\int_1^{+\infty} (\pi - 2\arctan x)\,\mathrm{d}x = \,?$$

We know that $\lim_{x\to+\infty} \arctan x = \frac{\pi}{2}$, so that the integrand $f(x) = \pi - 2\arctan x$ tends to 0. Let's determine its order at $+\infty$ using De l'Hôpital:

$$\lim_{x\to+\infty} \frac{\pi - 2\arctan x}{\frac{1}{x}} = \lim_{x\to+\infty} \frac{2x^2}{1 + x^2} = 2.$$

Therefore $f(x) \sim \frac{2}{x}$ as $x \to +\infty$ and the integral diverges using the Asymptotic Comparison Test.

Example 11.6: Investigate integrals of the form (for $\alpha, \beta > 0$)

$$\int_2^{+\infty} \frac{1}{x^\alpha (\ln x)^\beta} \, \mathrm{d}x = ?$$

Case 1: $\alpha = 1$. In this case we can make the substitution $y = \ln x$ so that $\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{1}{x}$ and we get

$$\int_2^{+\infty} \frac{1}{x(\ln x)^\beta} \, \mathrm{d}x = \int_{\ln 2}^{+\infty} \frac{1}{y^\beta} \, \mathrm{d}y$$

which converges if $\beta > 1$ and diverges if $\beta \leq 1$.

Case 2: $\alpha > 1$. Since $\ln x$ is strictly increasing, we have

$$\frac{1}{x^\alpha (\ln x)^\beta} \leq \frac{1}{x^\alpha (\ln 2)^\beta} \qquad \forall x \geq 2.$$

The Comparison Test implies that the integral will *converge* (regardless of β) since

$$\int_2^{+\infty} \frac{1}{x^\alpha (\ln 2)^\beta} \, \mathrm{d}x = \frac{1}{(\ln 2)^\beta} \int_2^{+\infty} \frac{1}{x^\alpha} \, \mathrm{d}x < +\infty.$$

Case 3: $\alpha < 1$. In this case we can write the integrand as

$$\frac{1}{x^\alpha (\ln x)^\beta} = \frac{1}{x} \frac{x^{1-\alpha}}{(\ln x)^\beta}$$

The function $\frac{x^{1-\alpha}}{(\ln x)^\beta}$ tends to $+\infty$ for any β (we've seen that x to any positive power grows faster than $\ln x$ to any positive power), so that there exists $M > 0$ such that

$$\frac{1}{x^\alpha (\ln x)^\beta} \geq \frac{M}{x}, \qquad \forall x \geq 2.$$

The Comparison Test implies that the integral *diverges* in this case.

Remark: The integral $\int_{-\infty}^b f(x) \, \mathrm{d}x$ is defined in an analogous way, and all the analysis follows analogously:

$$\int_{-\infty}^b f(x) \, \mathrm{d}x = \lim_{a \to -\infty} \int_a^b f(x) \, \mathrm{d}x.$$

11.1.2 Type II improper integrals

Improper integrals of type II occur when the integrand has a singularity (e.g., vertical asymptote) at a finite point of the integration domain.

Improper integral (type II)

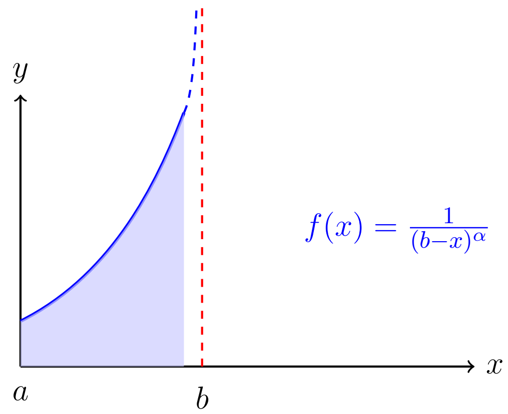
Let $f : \mathbb{R} \to \mathbb{R}$ be a function integrable on $[a, c]$ for any $c < b$, but with $\lim_{x\to b^-} |f(x)| = +\infty$ (or undefined). We define the **improper integral** (type II) of f on $[a, b)$ to be

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{t\to b^-} \int_a^t f(x)\,\mathrm{d}x.$$

Similarly, if f has a singularity at $x = a$, we define

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{t\to a^+} \int_t^b f(x)\,\mathrm{d}x.$$

They will *converge*, *diverge* or be *indeterminate* depending on the limit on the right hand side.



Example 11.7: Investigate the convergence of the integral

$$\int_0^1 \frac{1}{x^\alpha}\,\mathrm{d}x,$$

where $\alpha > 0$. The integrand has a singularity at $x = 0$ when $\alpha > 0$.

We analyze two main cases: $\alpha = 1$ and $\alpha \neq 1$.

Case 1: $\alpha = 1$. In this case, the integrand is $f(x) = \frac{1}{x}$. We compute:

$$\int_t^1 \frac{1}{x}\,\mathrm{d}x = [\ln|x|]_t^1 = \ln|1| - \ln|t| = -\ln t$$

Now we take the limit as $t \to 0^+$:

$$\int_0^1 \frac{1}{x}\,\mathrm{d}x = \lim_{t\to 0^+} (-\ln t) = +\infty.$$

Therefore, the integral *diverges* for $\alpha = 1$.

Case 2: $\alpha \neq 1$. We have:

$$\int_t^1 \frac{1}{x^\alpha}\,\mathrm{d}x = \left[\frac{1}{1-\alpha}x^{1-\alpha}\right]_t^1 = \frac{1}{1-\alpha}\left(1 - t^{1-\alpha}\right)$$

Now we analyze the limit $\lim_{t\to 0^+} t^{1-\alpha}$:

- If $\mathbf{1-\alpha > 0}$ (i.e., $\alpha < \mathbf{1}$): Since $1-\alpha$ is positive, $\lim_{t\to 0^+} t^{1-\alpha} = 0$. The limit is:

  $$\int_0^1 \frac{1}{x^\alpha}\,\mathrm{d}x = \lim_{t\to 0^+} \frac{1}{1-\alpha}\left(1 - t^{1-\alpha}\right) = \frac{1}{1-\alpha}(1-0) = \frac{1}{1-\alpha}.$$

  The integral *converges* to $\frac{1}{1-\alpha}$ when $\alpha < \mathbf{1}$.

- If $\mathbf{1-\alpha < 0}$ (i.e., $\alpha > \mathbf{1}$): Since $1-\alpha$ is negative, we rewrite $t^{1-\alpha} = \frac{1}{t^{\alpha-1}}$, and $\lim_{t\to 0^+} \frac{1}{t^{\alpha-1}} = +\infty$. The limit is:

  $$\int_0^1 \frac{1}{x^\alpha}\,\mathrm{d}x = \lim_{t\to 0^+} \frac{1}{1-\alpha}\left(1 - t^{1-\alpha}\right) = \frac{1}{1-\alpha}(1-\infty) = +\infty.$$

  The integral *diverges* for $\alpha > \mathbf{1}$.

We conclude:

$$\int_0^1 \frac{1}{x^\alpha}\,\mathrm{d}x \qquad \begin{cases} \text{converges} & \text{if } 0 < \alpha < 1, \\ \text{diverges} & \text{if } \alpha \geq 1. \end{cases}$$

More generally, for an integral $\int_a^b \frac{1}{(b-x)^\alpha}\,\mathrm{d}x$ with singularity at b, the same convergence conditions hold.

Convergence tests

The comparison tests for type II improper integrals are analogous to those for type I.

Comparison Test

Let f, g be integrable on $[a,t]$ for any $t < b$, with $0 \leq f(x) \leq g(x)$ for all $x \in [a,b)$. Then

$$0 \leq \int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x$$

Therefore

$$\int_a^b g(x)\,\mathrm{d}x < +\infty \ \textit{(converges)} \quad \Rightarrow \quad \int_a^b f(x)\,\mathrm{d}x < +\infty \ \textit{(converges)}$$

$$\int_a^b f(x)\,\mathrm{d}x = +\infty \ \textit{(diverges)} \quad \Rightarrow \quad \int_a^b g(x)\,\mathrm{d}x = +\infty \ \textit{(diverges)}$$

Example 11.8: Determine whether

$$\int_0^1 \frac{|\ln x|}{x^{1/3}} \, \mathrm{d}x$$

converges or diverges.

Near $x = 0$, $|\ln x| \to +\infty$ and $x^{-1/3} \to +\infty$, so the integrand is unbounded. To use the comparison test, observe that for any $\alpha > 0$, we have

$$\lim_{x \to 0^+} x^{\alpha} |\ln x| = 0,$$

which implies that there exists $\delta > 0$ such that for all $0 < x < \delta$,

$$|\ln x| \leq x^{-\alpha}$$

(you can show this using De l'Hôpital, as in Example 8.28). Take $\alpha = 1/6$. Then for $0 < x < \delta$,

$$\frac{|\ln x|}{x^{1/3}} \leq \frac{x^{-1/6}}{x^{1/3}} = x^{-1/2}.$$

Since

$$\int_0^1 x^{-1/2} \, \mathrm{d}x = 2$$

converges, we conclude by the comparison test that

$$\int_0^1 \frac{|\ln x|}{x^{1/3}} \, \mathrm{d}x$$

also converges.

Absolute Convergence Test

Suppose that both f and $|f|$ are integrable on $[a, t]$ for any $t < b$. Then if $\int_a^b |f(x)| \, \mathrm{d}x$ converges, so does $\int_a^b f(x) \, \mathrm{d}x$, and

$$\left| \int_a^b f(x) \, \mathrm{d}x \right| \leq \int_a^b |f(x)| \, \mathrm{d}x.$$

Example 11.9: Consider $\int_0^1 \frac{\sin(1/x)}{\sqrt{x}} \, \mathrm{d}x$. The integrand oscillates wildly near $x = 0$, but

$$\left| \frac{\sin(1/x)}{\sqrt{x}} \right| \leq \frac{1}{\sqrt{x}}.$$

Since $\int_0^1 \frac{1}{\sqrt{x}} \, \mathrm{d}x$ converges ($\alpha = \frac{1}{2} < 1$), the original integral converges absolutely.

Asymptotic Comparison Test

Suppose $f(x) \sim \frac{1}{(b-x)^\alpha}$ as $x \to b^-$ (or similarly near a). Then:

$$\alpha < 1 \quad \Rightarrow \quad \int_a^b f(x)\,\mathrm{d}x < +\infty \ \textit{(converges)}$$

$$\alpha \geq 1 \quad \Rightarrow \quad \int_a^b f(x)\,\mathrm{d}x = +\infty \ \textit{(diverges)}$$

Example 11.10: Investigate $\int_0^{\pi/2} \tan x \,\mathrm{d}x$.

The integrand has a singularity at $x = \pi/2$ since $\tan x \to +\infty$ as $x \to (\pi/2)^-$. We analyze the behavior near $\pi/2$:

$$\tan x = \frac{\sin x}{\cos x} \sim \frac{1}{\cos x} \quad \text{as } x \to \frac{\pi}{2}^-.$$

Using the substitution $t = \pi/2 - x$, we get $\cos x = \sin t \sim t$ as $t \to 0^+$. Thus $\tan x \sim \frac{1}{t}$ near $x = \pi/2$, i.e., $\alpha = 1$. Since $\int_0^{\pi/2} \frac{1}{(\pi/2-x)}\,\mathrm{d}x$ diverges, the original integral diverges.

Example 11.11: Investigate $\int_0^1 \frac{1}{\sqrt{x(1-x)}}\,\mathrm{d}x$.

This integrand has singularities at both endpoints $x = 0$ and $x = 1$. We split the integral at $x = 1/2$:

$$\int_0^1 \frac{1}{\sqrt{x(1-x)}}\,\mathrm{d}x = \int_0^{1/2} \frac{1}{\sqrt{x(1-x)}}\,\mathrm{d}x + \int_{1/2}^1 \frac{1}{\sqrt{x(1-x)}}\,\mathrm{d}x.$$

Near $x = 0$, $\frac{1}{\sqrt{x(1-x)}} \sim \frac{1}{\sqrt{x}}$ ($\alpha = 1/2 < 1$), so the first integral converges. Near $x = 1$, substitute $t = 1-x$ to get $\frac{1}{\sqrt{x(1-x)}} \sim \frac{1}{\sqrt{t}}$ ($\alpha = 1/2 < 1$), so the second integral converges. Thus the original integral converges.

Remark: When dealing with type II improper integrals, always:

1. Identify all points where the integrand is unbounded (singularities).
2. Split the integral at each singularity.
3. Analyze convergence separately for each part.
4. The integral converges only if *all* parts converge.

11.2 Numerical series

Improper integrals are a way of measuring the area of 'infinite' domains in the plane. Thinking of type I improper integrals, we can imagine replacing the function $f(x)$ with a step function that has the constant value $f(n)$ on any interval $[n, n+1)$, where $n \in \mathbb{N}$. Then,

$$\text{the integral} \qquad \int_N^{+\infty} f(x)\,\mathrm{d}x \qquad \text{is replaced by} \qquad \sum_{k=N}^{\infty} f(k).$$

This is called a numerical series. It can serve as an approximation for an improper integral, but it also is interesting irrespective. Thus, our goal in this section is to understand the meaning of an infinite sum of the form

$$a_0 + a_1 + a_2 + \cdots + a_n + \cdots = ?$$

Series and their partial sums

Given a sequence $\{a_n\}_{n\in\mathbb{N}}$ of real numbers, we form the **partial sums**

$$s_n = a_0 + a_1 + a_2 + \cdots + a_n = \sum_{k=0}^{n} a_k.$$

The **series** (or infinite sum) associated to $\{a_n\}_{n\in\mathbb{N}}$ is defined as the limit of the partial sums:

$$\sum_{k=0}^{\infty} a_k = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=0}^{n} a_k.$$

If the limit exists and is finite, we say the series *converges*; if the limit is infinite, the series *diverges*; if the limit does not exist, the series is *indeterminate*.

Geometric series

One of the most important examples is the geometric series.

Geometric series

For $r \in \mathbb{R}$, the geometric series is

$$\sum_{n=0}^{\infty} r^n = 1 + r + r^2 + r^3 + \cdots .$$

Its partial sums are given by

$$s_n = 1 + r + r^2 + \cdots + r^n = \frac{1 - r^{n+1}}{1 - r} \quad \text{for } r \neq 1.$$

Taking the limit as $n \to \infty$, we obtain:

$$\sum_{n=0}^{\infty} r^n = \begin{cases} \frac{1}{1-r} & \text{if } |r| < 1, \\ \text{diverges} & \text{if } |r| \geq 1. \end{cases}$$

p-Series

For $p > 0$, a **p-series** is the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$. As we shall see below, it will converge for $p > 1$ and diverge for $0 < p \leq 1$. In the case $p = 1$, the series $\sum_{n=1}^{\infty} \frac{1}{n}$ is called the **harmonic series**.

Necessary condition

Necessary condition for convergence

If $\sum_{n=0}^{\infty} a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.

Proof. Let $s_n = \sum_{k=0}^{n} a_k$ and $s = \lim_{n\to\infty} s_n$. Then

$$a_n = s_n - s_{n-1} \to s - s = 0 \quad \text{as } n \to \infty.$$

□

Remark: The converse is *false*! For example, the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges even though $\frac{1}{n} \to 0$. *(the fact that the harmonic series diverges will be a consequence of the integral test, which we will see below)*

Convergence tests for series with non-negative terms

For series with $a_n \geq 0$ for all n, the sequence of partial sums $\{s_n\}$ is non-decreasing, so it either converges to a finite limit or diverges to $+\infty$.

Comparison Test

Let $\{a_n\}$ and $\{b_n\}$ be sequences with $0 \le a_n \le b_n$ for all n. Then:

- If $\sum_{n=0}^{\infty} b_n$ converges, then $\sum_{n=0}^{\infty} a_n$ converges.
- If $\sum_{n=0}^{\infty} a_n$ diverges, then $\sum_{n=0}^{\infty} b_n$ diverges.

Example 11.12: The p-series with $p = 2$, $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges because $\frac{1}{n^2} \le \frac{2}{n(n+1)} = 2\left(\frac{1}{n} - \frac{1}{n+1}\right)$ and

$$\sum_{n=1}^{N} 2\left(\frac{1}{n} - \frac{1}{n+1}\right) = 2\left(1 - \frac{1}{N+1}\right) \to 2.$$

Ratio Test (d'Alembert's test)

Let $\sum_{n=0}^{\infty} a_n$ be a series with $a_n > 0$ for all n, and let

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

Then:

- If $L < 1$, the series converges.
- If $L > 1$, the series diverges.
- If $L = 1$, the test is inconclusive.

Example 11.13: For $\sum_{n=1}^{\infty} \frac{n}{2^n}$, we have

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)/2^{n+1}}{n/2^n} = \frac{n+1}{2n} \to \frac{1}{2} < 1,$$

so the series converges by the Ratio Test.

Root Test (Cauchy's test)

Let $\sum_{n=0}^{\infty} a_n$ be a series with $a_n \ge 0$ for all n, and let

$$L = \lim_{n\to\infty} \sqrt[n]{a_n}.$$

Then:

- If $L < 1$, the series converges.
- If $L > 1$, the series diverges.
- If $L = 1$, the test is inconclusive.

Example 11.14: For $\sum_{n=1}^{\infty}\left(\frac{n}{3n+1}\right)^n$, we have

$$\sqrt[n]{a_n} = \frac{n}{3n+1} \to \frac{1}{3} < 1,$$

so the series converges by the Root Test.

Integral Test

Let $f : [1, \infty) \to [0, +\infty)$ be a continuous, decreasing function with $f(n) = a_n$ for all $n \in \mathbb{N}$. Then

$$\sum_{n=1}^{\infty} a_n \text{ converges} \quad \Leftrightarrow \quad \int_1^{\infty} f(x)\,\mathrm{d}x \text{ converges.}$$

Moreover, we have the error estimate:

$$\int_{N+1}^{\infty} f(x)\,\mathrm{d}x \le \sum_{n=N+1}^{\infty} a_n \le \int_N^{\infty} f(x)\,\mathrm{d}x.$$

Example 11.15: Consider the p-series $\sum_{n=1}^{\infty}\frac{1}{n^p}$. Let $f(x) = \frac{1}{x^p}$, which is continuous and decreasing for $x \ge 1$. We know from improper integrals that

$$\int_1^{\infty} \frac{1}{x^p}\,\mathrm{d}x \text{ converges if and only if } p > 1.$$

Therefore, by the Integral Test:

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \text{ converges if and only if } p > 1.$$

Alternating series

A series is called **alternating** if its terms alternate in sign.

Leibniz's Alternating Series Test

Let $\{a_n\}$ be a sequence such that:

1. $a_n \ge 0$ for all n,
2. $\{a_n\}$ is decreasing: $a_{n+1} \le a_n$ for all n,
3. $\lim_{n\to\infty} a_n = 0$.

Then the alternating series $\sum_{n=0}^{\infty}(-1)^n a_n$ converges. Moreover, if $S = \sum_{n=0}^{\infty}(-1)^n a_n$ and $s_N = \sum_{n=0}^{N}(-1)^n a_n$, then the error satisfies:

$$|S - s_N| \le a_{N+1}.$$

Example 11.16: The alternating harmonic series $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ converges by Leibniz's test, since $\frac{1}{n}$ is positive, decreasing, and tends to 0.

Note that the ordinary harmonic series $\sum_{n=1}^{\infty}\frac{1}{n}$ diverges, so this shows that convergence of an alternating series does not imply absolute convergence.

Absolute convergence

Absolute Convergence Test

If $\sum_{n=0}^{\infty}|a_n|$ converges, then $\sum_{n=0}^{\infty}a_n$ converges. In this case, we say the series **converges absolutely** and we have

$$\left|\sum_{n=0}^{\infty}a_n\right| \leq \sum_{n=0}^{\infty}|a_n|.$$

Proof. Let $s_n = \sum_{k=0}^{n}a_k$ be the partial sums of $\sum_{n=0}^{\infty}a_n$, and let $t_n = \sum_{k=0}^{n}|a_k|$ be the partial sums of $\sum_{n=0}^{\infty}|a_n|$.

Since $\sum_{n=0}^{\infty}|a_n|$ converges, the sequence $\{t_n\}_{n\in\mathbb{N}}$ converges to some limit T. This means that for any $\varepsilon > 0$, there exists N such that for all $m > n \geq N$,

$$t_m - t_n = \sum_{k=n+1}^{m}|a_k| < \varepsilon.$$

Now, for the same $m > n \geq N$, using the triangle inequality:

$$|s_m - s_n| = \left|\sum_{k=n+1}^{m}a_k\right| \leq \sum_{k=n+1}^{m}|a_k| = t_m - t_n < \varepsilon.$$

Thus the partial sums $\{s_n\}$ satisfy that: for any $\varepsilon > 0$, there exists N such that $|s_m - s_n| < \varepsilon$ whenever $m > n \geq N$. In this course we haven't learned this precise condition, however this means $\{s_n\}_{n\in\mathbb{N}}$ is a convergent sequence (it is called a *Cauchy* sequence), so $\sum_{n=0}^{\infty}a_n$ converges. $\square$

Example 11.17: The series $\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}$ converges absolutely because

$$\sum_{n=1}^{\infty}\left|\frac{(-1)^n}{n^2}\right| = \sum_{n=1}^{\infty}\frac{1}{n^2}$$

converges (p-series with $p = 2 > 1$).

Remark: Absolute convergence is stronger than conditional convergence:

- *Absolutely convergent* series can be rearranged without changing the sum.
- *Conditionally convergent* series (convergent but not absolutely convergent) can be rearranged to converge to any real number or even diverge (Riemann rearrangement theorem).

Example 11.18: The alternating harmonic series $\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}$ converges conditionally but not absolutely, since $\sum_{n=1}^{\infty}\frac{1}{n}$ diverges.

Chapter 12

Ordinary differential equations

We are well acquainted with equations of a single variable, such as

$$x^2 - 1 = 0$$

whose solutions are $x = \pm 1$, as well as equations involving two variables, such as

$$x^2 + y^2 - 1 = 0$$

which describes the unit circle. In the first case, the solution was two points (± 1) in real line $\mathbb{R}$. In the second case we had a circle in the plane $\mathbb{R}^2$. As a rule of thumb, the solution to an equation of n variables is described by a set of dimension $n-1$ (but not always).

In contrast to these examples, a **differential equation** is an equation involving a *function* of the variable x and its derivatives. The **order** of an equation refers to the order of the highest derivative (so if the highest derivative appearing in the equation is the third derivative, we say that it is a third order equation).

1. The simplest example is the **first order equation**

   $$y'(x) = f(x)$$

   where f is some given (known) function. The unknown is the function $y(x)$. In this case we already know the solution: it is the antiderivative of f,

   $$y(x) = \int f(x)\, \mathrm{d}x.$$

2. A slightly more complicated first order example is the following equation, that includes both y and its own derivative:

   $$y'(x) = ky(x)$$

   where $k \neq 0$ is some fixed (known) constant. By examination we can see that a possible solution is

   $$y(x) = Ae^{kx},$$

   modeling exponential growth ($k > 0$) or decay ($k < 0$), where $A \in \mathbb{R}$ is some constant. Examples here include virus spread (exponential growth) and the decay of a radioactive material (exponential decay). We can verify the solution by plugging into the equation:

   $$y'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(Ae^{kx}\right) = kAe^{kx} = ky(x).$$

3. Another common example (which we shall discuss below) is the **second order equation**

$$y''(x) = -k^2 y(x)$$

which has the possible solutions (also, we find this by examination)

$$y_1(x) = A\cos(kx) \qquad \text{or} \qquad y_2(x) = B\sin(kx)$$

where $A, B \in \mathbb{R}$ are some constants. We can write a more general solution as

$$y(x) = A\cos(kx) + B\sin(kx).$$

Let us verify that it solves the equation:

$$\begin{aligned} y''(x) &= \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(A\cos(kx) + B\sin(kx)\right) \\ &= A\frac{\mathrm{d}^2}{\mathrm{d}x^2}\cos(kx) + B\frac{\mathrm{d}^2}{\mathrm{d}x^2}\sin(kx) \\ &= Ak^2(-\cos(kx)) + Bk^2(-\sin(kx)) \\ &= -k^2\left(A\cos(kx) + B\sin(kx)\right) \\ &= -k^2 y(x). \end{aligned}$$

Formal definitions

Ordinary differential equation

An **ordinary differential equation (ODE)** is an equation that involves the variable x, the unknown function $y = y(x)$ (this is the function we are looking for), as well as derivatives of $y(x)$ up to any finite order $n \in \mathbb{N}$. This equation can be expressed as

$$\mathcal{F}(x, y, y', \ldots, y^{(n)}) = 0.$$

The **order** of the equation is the order of the highest derivative, n in this case. The ODE is called **autonomous** if $\mathcal{F}$ does not explicitly depend on x.

Note that $\mathcal{F}$ is a real-valued function taking values in the Cartesian product of $n+2$ copies of $\mathbb{R}$:

$$\mathcal{F} : \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n+2 \text{ times}} \to \mathbb{R}$$

Solution

A function $y = y(x)$ is a **solution** of the ODE $\mathcal{F}(x, y, y', \ldots, y^{(n)}) = 0$ over the interval $I \subseteq \mathbb{R}$ if it is n-times differentiable on I, and if

$$\mathcal{F}(x, y(x), y'(x), \ldots, y^{(n)}(x)) = 0, \qquad \forall x \in I.$$

As we have seen in the examples above, solutions need not be unique. As we shall see below, solutions also *might not exist*. These two observations make the study of ODEs delicate. In this course we are only getting a glimpse of this.

Normal form

If in the ODE $\mathcal{F}(x, y, y', \ldots, y^{(n)}) = 0$ we can isolate the highest derivative of y, and rewrite it as

$$y^{(n)} = f(x, y, y', \ldots, y^{(n-1)})$$

we say that it is written in **normal form**. Here f is a real-valued function of $n+1$ real variables.

Linearity and homogeneity

The ODE $\mathcal{F}(x, y, y', \ldots, y^{(n)}) = 0$ is called **linear** if it can be written in the form

$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y - g(x) = 0$$

where the coefficients $a_i(x)$ and the *forcing function* $g(x)$ are functions of the independent variable x only. If the ODE is not linear, it is called **non-linear**. If $g(x) = 0$, the equation is said to be **homogeneous**; otherwise, it is **non-homogeneous** or **inhomogeneous**.

Example 12.1: For the ODE $\mathcal{F}(x, y, y', \ldots, y^{(n)}) = 0$, the following table characterizes the properties of various expressions for $\mathcal{F}$:

Expression for $\mathcal{F}$	Order	Linearity	Homogeneity	Autonomy
$y' - x^2$	1st	Linear	Non-homogeneous	Non-autonomous
$y' + y^2 - x^3$	1st	Non-linear	Non-homogeneous	Non-autonomous
$y'' + k^2 y$	2nd	Linear	Homogeneous	Autonomous
$y'' + k^2 \sin y$	2nd	Non-linear	Homogeneous	Autonomous
$xy'' + k^2 \sin y$	2nd	Non-linear	Homogeneous	Non-autonomous

Convention: time-dependent problems

When the independent variable is a *time* variable, we typically denote each derivative by a *dot* above the function, instead of a *prime*:

$$\dot{y}(t) \qquad \text{instead of} \qquad y'(t)$$

and

$$\ddot{y}(t) \qquad \text{instead of} \qquad y''(t).$$

Important ODEs

Basic ODEs

$$
\begin{aligned}
\text{Exponential Growth/Decay:}\quad & \frac{\mathrm{d}y}{\mathrm{d}t} = ky \\
\text{Newton's Law of Cooling:}\quad & \frac{\mathrm{d}T}{\mathrm{d}t} = -k(T - T_a) \\
\text{Logistic Growth Equation:}\quad & \frac{\mathrm{d}P}{\mathrm{d}t} = rP\left(1 - \frac{P}{K}\right) \\
\text{Non-linear Pendulum Equation:}\quad & \frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \omega^2 \sin\theta = 0 \\
\text{Simple Harmonic Motion:}\quad & \frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \omega^2\theta = 0 \\
\text{Damped Harmonic Oscillator:}\quad & m\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + c\frac{\mathrm{d}\theta}{\mathrm{d}t} + \omega^2\theta = 0
\end{aligned}
$$

First-order models, such as the Exponential Growth and Newton's Law of Cooling, utilize the relation $\frac{\mathrm{d}y}{\mathrm{d}t} = ky$ to represent physical processes like population dynamics and thermal equilibration. The Logistic equation refines these by introducing a carrying capacity K to model resource-limited growth.

Second-order equations, such as the Simple Harmonic Motion and Damped Oscillator characterize mechanical and electrical vibrations, where acceleration $\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2}$ is influenced by restoring forces and resistive damping.

Specialized ODEs

$$
\begin{aligned}
\text{Bernoulli Equation:}\quad & \frac{\mathrm{d}y}{\mathrm{d}x} = p(x)y^\alpha + q(x)y, \qquad \alpha \neq 0, 1 \\
\text{Riccati Equation:}\quad & \frac{\mathrm{d}y}{\mathrm{d}x} = p(x)y^2 + q(x)y + r(x) \\
\text{Clairaut Equation:}\quad & y = x\frac{\mathrm{d}y}{\mathrm{d}x} + f\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right) \\
\text{Bessel's Equation:}\quad & x^2\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + x\frac{\mathrm{d}y}{\mathrm{d}x} + (x^2 - \nu^2)y = 0 \\
\text{Legendre's Equation:}\quad & (1 - x^2)\frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 2x\frac{\mathrm{d}y}{\mathrm{d}x} + n(n+1)y = 0 \\
\text{Hermite Equation:}\quad & \frac{\mathrm{d}^2y}{\mathrm{d}x^2} - 2x\frac{\mathrm{d}y}{\mathrm{d}x} + 2ny = 0 \\
\text{Airy Equation:}\quad & \frac{\mathrm{d}^2y}{\mathrm{d}x^2} - xy = 0 \\
\text{Van der Pol Oscillator:}\quad & \frac{\mathrm{d}^2x}{\mathrm{d}t^2} - \mu(1 - x^2)\frac{\mathrm{d}x}{\mathrm{d}t} + x = 0
\end{aligned}
$$

More specialized ODEs often arise in higher-dimensional physics and non-linear systems where solutions cannot be expressed via elementary functions.

The Riccati and Bernoulli equations are *first-order non-linear* equations that frequently appear in control theory and fluid dynamics.

Second-order equations like those of Bessel, Legendre, and Hermite are essential for solving partial differential equations in cylindrical, spherical, or quantum mechanical domains, giving rise to "Special Functions" that describe everything from electromagnetic wave propagation to atomic orbitals.

Finally, *second order non-linear* models like the Van der Pol oscillator illustrate complex behaviors such as limit cycles, which arise in self-sustaining biological and electronic systems.

12.1 First-order differential equations in normal form

A first-order ODE which can be written in normal form is

$$y' = f(x, y).$$

The function $f(x, y)$ associates to each point $(x, y) \in \mathbb{R}^2$ in the plane a number which is the slope of the solution at that point. This can be visualized as a 'direction field', see Figure 12.1.

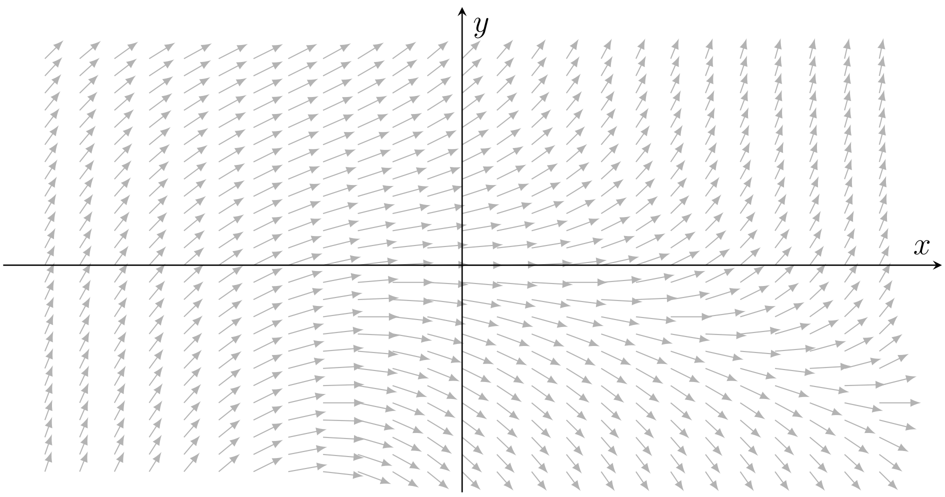


Figure 12.1: Direction field for the ODE $y' = (1 + x)y + x^2$.

By choosing a starting point $(x_0, y_0) \in \mathbb{R}^2$ in the plane and following the arrows, we construct a solution $y = y(x)$. Different starting points will lead to different solutions. This leads us to consider the *initial-value problem*.

Initial-value problem

For the normal form ODE $y' = f(x, y)$, the **initial-value problem (IVP)** (also called the **Cauchy problem**) for an interval $I \subseteq \mathbb{R}$ is

$$\begin{cases} y' = f(x, y) & \text{in } I, \\ y(x_0) = y_0, \end{cases}$$

where $x_0 \in I$ and $y_0 \in \mathbb{R}$.

For the initial-value problem presented above, the solution is obtained by following the arrows starting from the point (x_0, y_0) and drawing the resulting curve for as long as possible. But that is not always possible. Sometimes we might be led by the arrows in various directions, leading to different solutions. The following theorem tells us when the arrows will lead us in a unique direction, starting from (x_0, y_0).

Existence and uniqueness of solutions: Picard-Lindlöf Theorem

Consider the initial value problem:

$$\begin{cases} y' = f(x, y) & \text{in } I, \\ y(x_0) = y_0, \end{cases}$$

where $x_0 \in I$ and $y_0 \in \mathbb{R}$. Suppose $f(x, y)$ is a continuous function in a rectangular region R defined by $|x - x_0| \leq a$ and $|y - y_0| \leq b$ where $a, b > 0$ and $[x_0 - a, x_0 + a] \subseteq I$. If f is Lipschitz with respect to y in R, i.e. if there exists a constant $L > 0$ such that:

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2| \qquad \forall (x, y_1), (x, y_2) \in R$$

then there exists $h > 0$ such that there **exists** a **unique** solution $y(x)$ to the IVP on $[x_0 - h, x_0 + h]$. Moreover, the solution is *stable*, in the sense that a small perturbation of the initial value (x_0, y_0) will result in a small perturbation of the solution curve $y(x)$ (there's no 'butterfly effect').

We do not pursue the proof here, as it is beyond the scope of this course. This theorem generalizes to higher order ODEs that are not in normal form, but we omit this here.

Remark: The Picard-Lindelöf Theorem is crucial for modeling. By requiring the function $f(x, y)$ to be Lipschitz continuous, we ensure that the system's evolution is uniquely determined by its initial state. For linear equations, this condition is naturally satisfied if the coefficients are continuous. However, in non-linear analysis, checking the Lipschitz constant L is vital to ensure that numerical simulations – like those using Runge-Kutta methods – converge to a single, physically meaningful result rather than branching into multiple mathematical possibilities. See Example 12.5 for an example of an IVP with three possible solutions.

Equations with separable variables: $f(x, y) = h(x)g(y)$

Here is a general strategy for solving the IVP

$$\begin{cases} y' = f(x, y) & \text{in } I, \\ y(x_0) = y_0, & x_0 \in I, \end{cases}$$

in the case where $f(x, y) = h(x)g(y)$:

1. In the expression for f separate the x variable and the y variable, writing $f(x, y) = h(x)g(y)$ (note that this can only be done in special cases), so that the ODE becomes

$$\frac{\mathrm{d}y}{\mathrm{d}x} = h(x)g(y).$$

[Note: any $\tilde{y} \in \mathbb{R}$ for which $g(\tilde{y}) = 0$ will provide the (uninteresting) solution of the ODE: $y(x) = \tilde{y}$ for all x.]

2. For all y for which $g(y) \neq 0$, we can move $g(y)$ to the left hand side and the $\mathrm{d}x$ to the right hand side (this is a formal step, lacking rigor, but practical) to obtain

$$\frac{\mathrm{d}y}{g(y)} = h(x)\,\mathrm{d}x.$$

3. Compute the indefinite integrals of both sides (one involving only x and the other involving only y):

$$\int \frac{\mathrm{d}y}{g(y)} = \int h(x)\,\mathrm{d}x$$

to obtain a relation between y and x. In this step a constant C appears as well. We call it a *degree of freedom* of the equation.

4. Impose the condition $y(x_0) = y_0$ in order to determine the value of C.

Example 12.2: Let us solve the IVP

$$\begin{cases} y' = -5y \\ y(10) = 3 \end{cases}$$

in $\mathbb{R}$. Following the strategy we write

$$\frac{\mathrm{d}y}{y} = -5\,\mathrm{d}x$$

and take the antiderivatives of both sides to get

$$\ln y = -5x + C.$$

Taking the exponential of both sides we have

$$y(x) = e^{-5x+C} = e^C e^{-5x}.$$

Plugging the initial condition:

$$3 = e^C e^{-50} \qquad \Rightarrow \qquad e^C = 3e^{50}.$$

So the solution is:

$$y(x) = 3e^{50}e^{-5x}$$

which is valid for all $x \in \mathbb{R}$.

Example 12.3: Solve the IVP

$$\begin{cases} y' = \frac{e^x+1}{e^y+1} \\ y(2) = 3. \end{cases}$$

Following the strategy we have

$$(e^y + 1)\,\mathrm{d}y = (e^x + 1)\,\mathrm{d}x$$

and take the antiderivatives of both sides to get

$$e^y + y = e^x + x + C.$$

Plugging the initial condition:

$$e^3 + 3 = e^2 + 2 + C \qquad \Rightarrow \qquad C = e^3 - e^2 + 1.$$

So the solution is given implicitly by:

$$e^y + y = e^x + x + e^3 - e^2 + 1$$

which is valid for all $x \in \mathbb{R}$. Note that because the equation $e^y + y = f(x)$ is transcendental, it is not possible to express $y(x)$ explicitly in terms of elementary functions.

Example 12.4 (IVP with finite-time 'blow up')**:** Solve the IVP

$$\begin{cases} y' = y^2 \\ y(1) = 5. \end{cases}$$

Following the strategy we have

$$\frac{\mathrm{d}y}{y^2} = \mathrm{d}x.$$

Taking antiderivatives:

$$-\frac{1}{y} = x + C.$$

Rearranging we get

$$y(x) = -\frac{1}{x + C}$$

which is not defined for $x = -C$ (vertical asymptote). Plugging in the initial condition we have:

$$5 = -\frac{1}{1 + C} \qquad \Rightarrow \qquad C = -\frac{6}{5}$$

so that our solution is

$$y(x) = \frac{1}{\frac{6}{5} - x}$$

and is valid for all $x < \frac{6}{5}$ (which is the part of the domain that includes $x_0 = 1$). This solution (as well as the direction field) is visualized in Figure 12.2. In this case the solution does not exist for all $x \in \mathbb{R}$: it *blows up* as $x \to \frac{6}{5}$ from the left.

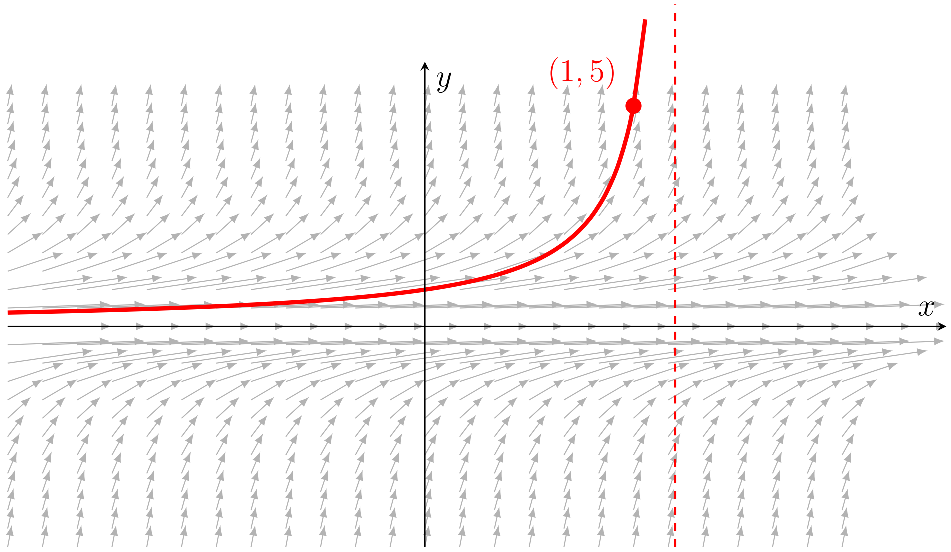


Figure 12.2: Direction field for the ODE $y' = y^2$ with particular solution $y(x) = \frac{1}{\frac{6}{5}-x}$ passing through $(1, 5)$.

Example 12.5 (IVP with more than one solution)**:** The IVP

$$\begin{cases} y' = y^{1/3} \\ y(0) = 0 \end{cases}$$

doesn't satisfy the Lipschitz requirement appearing in the Picard-Lindlöf theorem (the function $y^{1/3}$ is not Lipschitz at $y = 0$). As a result it does not have a unique solution. Let us show this:

- The trivial solution
$$y_1(x) = 0, \qquad \forall x \in \mathbb{R},$$
satisfies the IVP: $y'(x) = 0 = y^{1/3}(x)$ for all $x \in \mathbb{R}$.

- We can find another solution using our recipe from before. We rewrite the ODE as
$$\frac{\mathrm{d}y}{y^{1/3}} = \mathrm{d}x$$
and integrate, to get
$$\frac{3}{2}y^{2/3} = x + C.$$
The initial condition implies that $C = 0$ so that a second solution is
$$y_2(x) = \left(\frac{2}{3}x\right)^{3/2}, \qquad \forall x \geq 0.$$

- In fact, there's a third solution with a negative sign:
$$y_3(x) = -\left(\frac{2}{3}x\right)^{3/2}, \qquad \forall x \geq 0.$$

The three solutions y_1, y_2 and y_3 are all legitimate solutions, and we have no method to label any one of them as 'correct' or as 'incorrect'. See Figure 12.3 for a visualization.

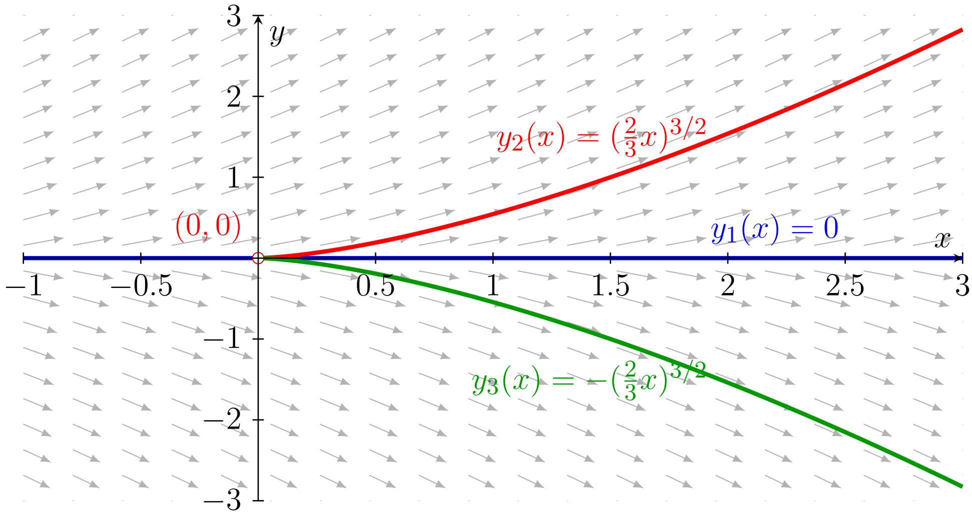


Figure 12.3: Direction field for $y' = y^{1/3}$ showing non-uniqueness at $(0, 0)$.

12.2 Second-order ODEs

12.2.1 The pendulum

In this section we demonstrate how Newton's Second Law $\mathbf{F} = m\mathbf{a}$ translates into a second-order differential equation in the case of a *physical pendulum* as in Figure 12.4.

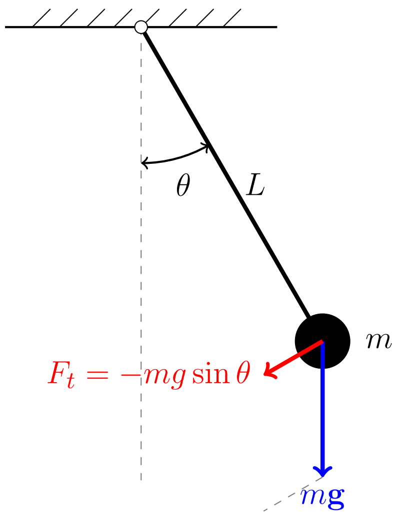


Figure 12.4: The physical pendulum. The restoring force is $mg\sin\theta$.

To derive the equation of motion for a physical pendulum of mass m and length L, we apply **Newton's Second Law** for rotation about the pivot point:

$$\tau = I\alpha$$

where τ is the *torque* (this is the force), I is the *moment of inertia* (this plays the role of the mass), and α is the *angular acceleration* (this is the acceleration).

1. Rotational Parameters

For a point mass suspended at a distance L, the moment of inertia is:

$$I = mL^2.$$

The angular acceleration α is the second derivative of the angular displacement θ with respect to time t:

$$\alpha = \frac{\mathrm{d}^2\theta}{\mathrm{d}t^2}.$$

We just need an expression for the torque.

2. The Restoring Torque

From the free-body diagram, the restoring force acting tangential to the path of motion is $F_t = -mg\sin\theta$. The resulting torque is the product of this force and the lever arm L:

$$\tau = L \cdot F_t = L \cdot (-mg\sin\theta) = -mgL\sin\theta.$$

The negative sign indicates that the torque acts in the direction opposite to the displacement, attempting to restore the pendulum to its equilibrium position ($\theta = 0$).

3. The Equation of Motion

Substituting these expressions into Newton's Second Law $I\alpha = \tau$, we obtain:

$$(mL^2)\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} = -mgL\sin\theta.$$

Dividing both sides by mL^2 yields the governing second-order non-linear ODE:

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \frac{g}{L}\sin\theta = 0.$$

Denoting the natural **frequency** of the problem

$$\omega = \sqrt{\frac{g}{L}}$$

the equation become

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \omega^2\sin\theta = 0.$$

Observe that the frequency is independent of the angle θ and of the mass m.

4. Small-Angle Approximation

For engineering applications where θ is small (typically $\theta < 10°$), we use the approximation $\sin\theta \sim \theta$ (which we've seen many times before, see Example 6.1 for instance) and the equation becomes

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \omega^2\theta = 0.$$

This procedure is called a **linearization**, as we've taken a nonlinear equation and replaced the nonlinear part ($\sin\theta$) by a linear approximation (θ).

5. Solving the Linearized Problem

The general solution is a linear combination of sines and cosines, as we've already seen:

$$\theta(t) = A\cos(\omega t) + B\sin(\omega t).$$

The motion is perfectly periodic with **period**

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{L}{g}}.$$

Indeed,

$$\begin{aligned}\theta(t+T) &= A\cos(\omega(t+T)) + B\sin(\omega(t+T)) \\ &= A\cos(\omega t + 2\pi) + B\sin(\omega t + 2\pi) \\ &= A\cos(\omega t) + B\sin(\omega t) \\ &= \theta(t).\end{aligned}$$

The two constants A and B are determined by the *initial conditions*. Let us specify the IVP, which now must have *two* conditions:

$$\begin{cases}\ddot{\theta} + \omega^2\theta = 0, \\ \theta(0) = \theta_0, \\ \dot{\theta}(0) = \omega_0.\end{cases}$$

Physically, this means that at time $t = 0$ the pendulum was

- at an angle θ_0, and
- had angular velocity ω_0.

Then we have:

$$\theta_0 = \theta(t=0) = A\cos 0 + B\sin 0 = A$$

and, taking a derivative of the expression for $\theta(t)$ and evaluating at $\theta = 0$,

$$\omega_0 = \frac{\mathrm{d}\theta}{\mathrm{d}t}(t=0) = -A\omega\sin 0 + B\omega\cos 0 = B\omega.$$

Hence these initial conditions dictate:

$$A = \theta_0 \qquad \text{and} \qquad B = \frac{\omega_0}{\omega}$$

and the solution is

$$\theta(t) = \theta_0\cos(\omega t) + \frac{\omega_0}{\omega}\sin(\omega t).$$

In particular, for a pendulum released from rest at an initial angle θ_0, the solution simplifies to

$$\theta(t) = \theta_0\cos(\omega t).$$

In this linearized regime, the period is *isochronous*—it remains constant regardless of the amplitude θ_0. This predictability is the foundation for classical timekeeping.

12.2.2 Second Order Linear Homogeneous ODEs with Constant Coefficients

Consider the second order equation with constant coefficients $a, b, c \in \mathbb{R}$:

$$ay'' + by' + cy = 0.$$

We can view the left hand side as a mapping that takes y and sends it to $ay'' + by' + cy$. This is called an **operator**, and we write it as

$$\mathcal{L} = a\frac{\mathrm{d}^2}{\mathrm{d}x^2} + b\frac{\mathrm{d}}{\mathrm{d}x} + c.$$

Thus our ODE becomes

$$\mathcal{L}y = 0.$$

We search for solutions of the form $y(x) = e^{rx}$, where $r \in \mathbb{C}$. Applying the operator to this function gives us

$$\begin{aligned}\mathcal{L}y &= \left(a\frac{\mathrm{d}^2}{\mathrm{d}x^2} + b\frac{\mathrm{d}}{\mathrm{d}x} + c\right) e^{rx} \\ &= \left(ar^2 + br + c\right) e^{rx}.\end{aligned}$$

Hence the ODE becomes

$$\left(ar^2 + br + c\right) e^{rx} = 0.$$

Since e^{rx} is always nonzero, we can divide by it, and our ODE has therefore been converted into a quadratic algebraic equation:

$$ar^2 + br + c = 0.$$

This is called the **characteristic equation**. Since we are working in $\mathbb{C}$, we distinguish three cases based on the roots r_1, r_2:

1. **Distinct Real Roots** ($b^2 - 4ac > 0$):

$$y(x) = C_1 e^{r_1 x} + C_2 e^{r_2 x}.$$

2. **Repeated Real Root** ($b^2 - 4ac = 0$): Letting $r = r_1 = r_2$,

$$y(x) = (C_1 + C_2 x) e^{rx}.$$

3. **Complex Conjugate Roots** ($b^2 - 4ac < 0$): Letting $r = \alpha \pm i\beta$, we obtain the real-valued solution

$$y(x) = e^{\alpha x}(C_1 \cos(\beta x) + C_2 \sin(\beta x)).$$

Example 12.6: Solve the IVP

$$\begin{cases} y'' + 2y' + 5y = 0 \\ y(0) = 2,\ y'(0) = 0. \end{cases}$$

The characteristic equation for the operator $\mathcal{L}$ is $r^2 + 2r + 5 = 0$. Using the quadratic formula we find

$$r = \frac{-2 \pm \sqrt{4 - 20}}{2} = -1 \pm 2i.$$

So the general solution is $y(x) = e^{-x}(C_1 \cos 2x + C_2 \sin 2x)$. Plugging the first initial condition $y(0) = 2$:

$$2 = e^0(C_1 \cos 0 + C_2 \sin 0) \qquad \Rightarrow \qquad C_1 = 2.$$

To use the second condition, we compute the derivative

$$y'(x) = -e^{-x}(C_1 \cos 2x + C_2 \sin 2x) + e^{-x}(-2C_1 \sin 2x + 2C_2 \cos 2x).$$

Plugging $y'(0) = 0$:

$$0 = -C_1 + 2C_2 \qquad \Rightarrow \qquad 2C_2 = 2 \qquad \Rightarrow \qquad C_2 = 1.$$

The solution is $y(x) = e^{-x}(2 \cos 2x + \sin 2x)$, which is valid for all $x \in \mathbb{R}$.