Real Analysis: Functions of a real variable

Genival da Silva*

November 19, 2024

Contents

*email: gdasilva@tamusa.edu, website: www.gdasilvajr.com

I Naive set theory

1 Sets

A **set** X is a collection of objects, also called the *elements* of the set. If 'a' is an element of X, we write $a \in X$. On the other hand, if 'a' isn't an element of X, we write $a \notin X$.

A set X is *well defined* when there is a rule that allows us to say if an arbitrary element 'a' is or isn't an element of X.

Example 1. *The set X of all right triangles is well-defined. Indeed, given any object 'a', if 'a' is not a triangle or doesn't have a right angle then $a \notin X$. If 'a' is a right triangle then $a \in X$.*

Example 2. *The set X of all tall people is not well-defined. The notion of 'tall' is not universally defined, hence given any element a we can't say if $a \in X$ or $a \notin X$.*

Usually one uses the notation

$$X = \{a,\, b,\, c,\, \ldots\}$$

to represent the set X whose elements are $a, b, c, \ldots$, and if a set has no elements we denote it by $\emptyset$ and call it the **empty set**.

The set of *natural numbers* $1, 2, 3, \ldots$ will be represented by

$$\mathbb{N} = \{1, 2, 3,\, \ldots\}$$

The set of *integers* will be represented by

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3,\, \ldots\}$$

The set of *rational numbers*, that is, fractions $\frac{a}{b}$, where $a, b \in \mathbb{Z}$ and $b \neq 0$, will be denoted by

$$\mathbb{Q} = \{\, \frac{a}{b} \,|\, a, b \in \mathbb{Z},\, b \neq 0 \,\}$$

The vast majority of sets in mathematics are not defined by specifying its elements one by one. What usually happens is a set being defined by some property its elements satisfy, i.e. if a has property P than $a \in X$, whereas if a doesn't have property P then $a \notin X$. One writes

$$X = \{a \,|\, a \text{ has property } P\}$$

For example, the set

$$X = \{a \in \mathbb{N} \,|\, a > 10\},$$

consists of all natural numbers bigger than 10.

Given two sets A, B, one says that A is a **subset** of B or that A is *included* in B (B *contains* A), represented by $A \subseteq B$, if every element of A is an element of B.

Example 3. *We have the obvious inclusion of sets:*

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}.$$

Example 4. *Let X be the set of all squares and Y be the set of all rectangles. Then $X \subseteq Y$, since every square is a rectangle.*

When one writes $X \subseteq Y$, it's possible that $X = Y$. In case $X \neq Y$, we say X is a *proper subset*, the notation $X \subsetneq Y$ is sometimes used to indicate that X is a proper subset of Y.

Notice that to write $a \in X$ is equivalent to say $\{a\} \subseteq X$. Also, by definition, it's always true that $\emptyset \subseteq X$ for every set X.

It's easy to see that the inclusion of sets has the following properties:

1. *Reflexive*, $X \subseteq X$ for every set X;
2. *Anti-symmetric*, if $X \subseteq Y$ and $Y \subseteq X$ then $X = Y$;
3. *Transitive*, if $X \subseteq Y$ and $Y \subseteq Z$ then $X \subseteq Z$.

It follows that two sets X and Y are the same if and only if $X \subseteq Y$ and $Y \subseteq X$, that is to say, they have the same elements.

Given a set X, we define the *power set* of X, $\mathcal{P}(X)$ as

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

The set $\mathcal{P}(X)$ is the set of all subsets of X, in particular it's never empty, it has at least $\emptyset$ and X itself as elements.

Example 5. *Let $X = \{1, 2, 3\}$ then*

$$\mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}\}.$$

Notice that by using the Fundamental Counting Principle, any set with n elements has 2^n subsets. Therefore, the number of elements of $\mathcal{P}(X)$ is 2^n.

2 Operation with sets

We given two sets X and Y, one can build many other sets. For example, the **union** of X and Y, denoted by $X \cup Y$ is the of elements that are in X or Y, more precisely:

$$X \cup Y = \{ a \,|\, a \in X \text{ or } a \in Y \}.$$

Similarly, the **intersection** of X and Y, denoted by $X \cap Y$ is the of elements that are common to both X and Y:

$$X \cap Y = \{ a \,|\, a \in X \text{ and } a \in Y \}.$$

If $X \cap Y = \emptyset$, then X and Y are said to be *disjoint*.

Example 6. *Let* $X = \{a \in \mathbb{N} \,|\, a \leq 100\}$ *and* $Y = \{a \in \mathbb{N} \,|\, a > 50\}$ *then*

$$X \cup Y = \mathbb{N} \textit{ and } X \cap Y = \{a \in \mathbb{N} \,|\, 50 < a \leq 100\}$$

Example 7. *The sets* $X = \{a \in \mathbb{N} \,|\, a > 1\}$ *and* $Y = \{a \in \mathbb{N} \,|\, a < 2\}$ *are disjoint, i.e.* $X \cap Y = \emptyset$ *since there is no natural number between* 1 *and* 2.

The **difference** between X and Y, denoted by $X - Y$ is the set of elements that are in X but not in Y, more precisely:

$$X - Y = \{ a \,|\, a \in X \text{ and } a \notin Y \}.$$

Given an inclusion of sets $X \subseteq Y$, the **complement** of X in Y is the set $Y - X$, the notation X^c sometimes is used if there is no confusion about who the set Y is.

Example 8. *Consider the sets* $X = \{a \in \mathbb{N} \,|\, a \textit{ is even}\}$ *and* $Y = \mathbb{N}$. *Then* $X \subseteq Y$ *and* $X^c = \{a \in \mathbb{N} \,|\, a \textit{ is odd}\}$.

Proposition 9. *Given sets* A, B, C, D *the following properties are true:*

1. $A \cup \emptyset = A$; $A \cap \emptyset = \emptyset$
2. $A \cup A = A$; $A \cap A = A$
3. $A \cup B = B \cup A$; $A \cap B = B \cap A$
4. $A \cup (B \cup C) = (A \cup B) \cup C$; $A \cap (B \cap C) = (A \cap B) \cap C$

5. $A \cup B = A \Leftrightarrow B \subseteq A$; $A \cap B = A \Leftrightarrow A \subseteq B$

6. *if* $A \subseteq B$ *and* $C \subseteq D$ *then* $A \cup C \subseteq B \cup D$ *and* $A \cap C \subseteq B \cap D$

7. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$; $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

8. $(A^c)^c = A$

9. $(A \cup B)^c = A^c \cap B^c$; $(A \cap B)^c = A^c \cup B^c$

Proof. The last property, $(A \cup B)^c = A^c \cap B^c$, will be demonstrated below, the others are trivial or can be proved in a similar way.

We prove that $(A \cup B)^c \subseteq A^c \cap B^c$. Let $a \in (A \cup B)^c$, then $a \notin A \cup B$, in particular, $a \notin A$ and $a \notin B$, hence $a \in A^c \cap B^c$.

Conversely, take $a \in A^c \cap B^c$. Then $a \notin A$ and $a \notin B$, so $a \notin A \cup B$ and it follows that $a \in (A \cup B)^c$. □

An *ordered pair* (a, b) is formed by two objects a and b, such that for any other such pair (c, d):

$$(a, b) = (c, d) \Leftrightarrow a = c \text{ and } b = d.$$

The elements a and b are called *coordinates* of (a, b), a is the first coordinate and b the second one.

The **cartesian product** $X \times Y$ of two sets X and Y is the set of all ordered pairs (x, y) such that $x \in X$ and $y \in Y$:

$$X \times Y = \{\, (x, y) \,|\, x \in X \text{ and } y \in Y \,\}.$$

Remark 1. *An ordered pair is not the same as a set, i.e.* $(a, b) \neq \{a, b\}$. *Notice that* $\{a, b\} = \{b, a\}$ *but* $(a, b) \neq (b, a)$ *in general.*

Example 10. *Consider the sets* $X = \{1, 2, 3\}$ *and* $Y = \{a, b\}$, *then*

$$X \times Y = \{\, (1, a), (1, b), (2, a), (2, b), (3, a), (3, b) \,\}.$$

3 Functions

A **function** $f : X \to Y$ consists of three components: a set X, *the domain*, a set Y, the *co-domain*, and a rule that associates each element $a \in X$ an unique element in $f(a) \in Y$, $f(a)$ is called the *value* of $f(x)$ at a, or the image of a under $f(x)$.

Another common notation to denote a function is $x \mapsto f(x)$. In this case the domain and codomain can be identified by the context.

Example 11. *The function $f : \mathbb{N} \to \mathbb{N}$ given by $f(n) = n + 1$ is called the successor function.*

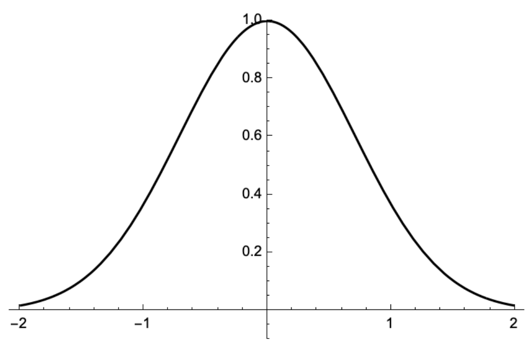
Example 12. *Let X be the set of all triangles. One can define a function $f : X \to \mathbb{R}$ by $f(x) =$ area of x.*

Example 13. *(Relation that is not a function) The correspondence that associates to each real number x, all y satisfying $y^2 = x$ is not a function because any $x \neq 0$ will be associated to two values, namely $\pm\sqrt{x}$, and in order to be a function every x has to have exactly one image $y = f(x)$.*

The graph of a function $f : X \to Y$ is a subset of $X \times Y$ defined by

$$\Gamma(f) = \{\, (x, f(x)) \,|\, x \in X \,\}.$$

Example 14. *Consider the function $f(x) = e^{-x^2}$, its graph is given below:*

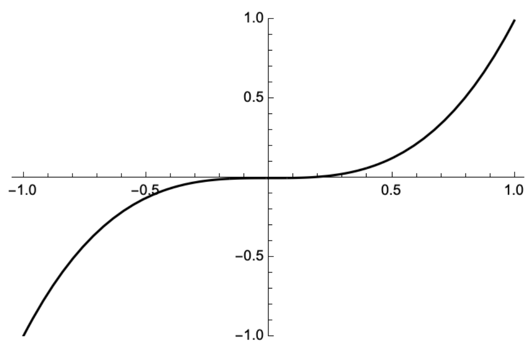


A function $f : X \to Y$ is said to be *injective or one-to-one* if for every x, y such that $f(x) = f(y)$ then $x = y$. Suppose $X \subseteq Y$, then inclusion $i : X \to Y$ given by $i(x) = x$ is a typical example of injective function.

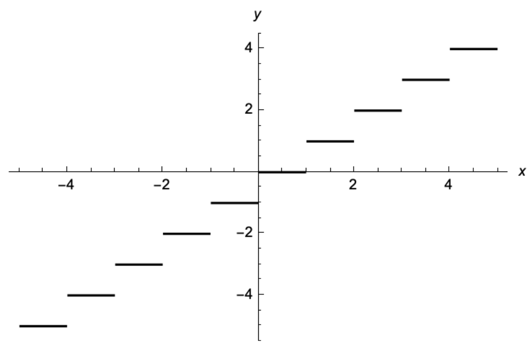
A function $f : X \to Y$ is said to be *surjective or onto* if for every $y \in Y$ there is $x \in X$ such that $y = f(x)$. The projection $p : X \times Y \to X$ in the first coordinate, given by $p(x, y) = x$ is a typical example of surjection.

Finally, a function $f : X \to Y$ is *bijective or a bijection* if it is both surjective and injective.

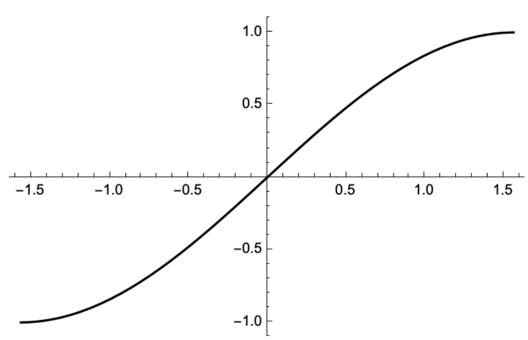
Example 15. *The function given by $f(x) = x^3$ is injective.*



Example 16. *The floor function* $\lfloor x \rfloor = \max\{\, n \in \mathbb{Z} \,|\, n \leq x \,\}$ *is not injective.*



Example 17. *The function* $f(x) = \sin x$ *is a bijection if we consider* $f : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$.



Given a function $f : X \to Y$, the *image of a set* $A \subseteq X$ is defined by

$$f(A) = \{\, y \in Y \,|\, y = f(a),\, a \in A \,\}.$$

Conversely, the *inverse image of a set* (sometimes called *pre-image*) $B \subseteq Y$ is given by

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

Proposition 18. *Given $f : X \to Y$ and subsets $A, B \subseteq X$, we have:*

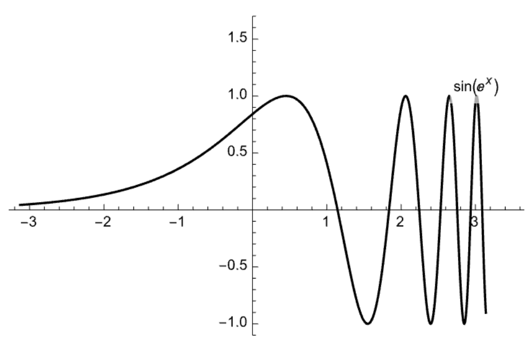
1. $f(A \cup B) = f(A) \cup f(B)$; $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$
2. $f(A \cap B) \subseteq f(A) \cap f(B)$; $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$
3. *if* $A \subseteq B$ *then* $f(A) \subseteq f(B)$ *and* $f^{-1}(A) \subseteq f^{-1}(B)$
4. $f(\emptyset) = \emptyset$; $f^{-1}(\emptyset) = \emptyset$
5. $f^{-1}(Y) = X$
6. $f^{-1}(A^c) = (f^{-1}(A))^c$

Example 19. *Consider the function $f(x, y) = x^2 + y^2$, the inverse image $f^{-1}(\{1\})$ is a circle of radius 1. Similarly, any line $ax + by = c$ can be seen as $g^{-1}(\{c\})$, where $g(x, y) = ax + by$.*

Given two functions $f : X \to Y$ and $g : Y \to Z$, *the composition* $g \circ f$ of g and f is defined as the function:

$$(g \circ f)(x) = g(f(x))$$

Example 20. *The composition of the functions $g(x) = \sin x$ and $f(x) = e^x$ is the function $(g \circ f)(x) = \sin e^x$ depicted below.*



Given a function $f : X \to Y$ and a subset $A \subseteq X$, *the restriction* of $f(x)$ to A, denoted by $f|_A : A \to Y$, is defined by $f|_A(x) = f(x)$. Similarly, if $X \subseteq Z$, *a extension* of $f(x)$ to Z is any function $g : Z \to Y$ such that $g|_X(x) = f(x)$.

Example 21. *Consider again the function $f(x, y) = x^2 + y^2$, and the unit circle $\mathbb{S}^1 = \{ (x, y) \mid x^2 + y^2 = 1 \}$. Then the restriction $f|_{\mathbb{S}^1}$ is the constant function $g(x) = 1$.*

Given functions $f : X \to Y$, and $g : Y \to X$, the function $g(x)$ is called *left-inverse* of $f(x)$ if

$$(g \circ f)(x) = x.$$

Similarly, the function $g(x)$ is called *right-inverse* of $f(x)$ if

$$(f \circ g)(x) = x.$$

Finally, if there is a function $f^{-1}(x)$ such that

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

$f^{-1}(x)$ is called the *inverse* of $f(x)$. Notice that any inverse, if exists, is unique. If $g(x)$ and $h(x)$ are both inverses of $f(x)$ then

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

Proposition 22. *A function $f : X \to Y$ has an inverse $f^{-1} : Y \to X \Leftrightarrow f$ is bijective.*

Proof. Suppose f has an inverse f^{-1} and $f(x) = f(y)$ for some x, y. Taking the inverse on both sides, we conclude that $x = y$ and f is injective. Similarly, take $y \in Y$ and set $x = f^{-1}(y)$, then $f(x) = y$ and it follows that f is surjective.

Conversely, suppose f bijective. If $f(x) = y$, set $f^{-1}(y) = x$. One can easily check that $(f \circ f^{-1})(x) = (f^{-1} \circ f)(x) = x$. □

Example 23. *Consider the function $f : (0, +\infty) \to (0, +\infty)$ given by $f(x) = \frac{1}{x}$, then the f is its own inverse, i.e. $(f \circ f)(x) = x$.*

4 The natural numbers $\mathbb{N}$

The natural numbers are built axiomatically. Start with a set $\mathbb{N}$, whose elements are called *natural numbers*, and a function $s : \mathbb{N} \to \mathbb{N}$, called the *successor function*. For any $n \in \mathbb{N}$, $s(n)$ is called the successor of n.

The function $s(n)$ satisfies the following axioms:

Axiom 1. $s(n)$ is injective, i.e. every number has a unique successor.

Axiom 2. The set $\mathbb{N} - s(\mathbb{N})$ has only one element, which will be denoted by 1, i.e. every number has a successor and 1 is not a successor of any number.

Axiom 3. (Principle of induction) Let $X \subseteq \mathbb{N}$ be a subset with the following property: $1 \in X$ and given $n \in X$, $s(n) \in X$ as well. Then $X = \mathbb{N}$.

Whenever axiom 3 is used to prove a result, the result is said to be proved by induction.

Proposition 24. *For any $n \in \mathbb{N}$, $s(n) \neq n$.*

Proof. The proof is by induction. Let $X \in \mathbb{N}$ be a subset defined by:

$$X = \{\, n \in \mathbb{N} \,|\, s(n) \neq n \,\}.$$

By Axiom 2, $1 \in X$. Let $n \in X$, then $s(n) \neq n$. By Axiom 1, $s(s(n)) \neq s(n)$, hence $s(n) \in X$. The proof follows by Axiom 3.

□

Given a function $f : X \to X$, its power f^n is defined inductively. More precisely, if one sets $f^1 = f$ then f^n is defined by:

$$f^{s(n)} = f \circ f^n.$$

In particular, if one sets $2 = s(1), 3 = s(2), \ldots$, then $f^2 = f \circ f, f^3 = f \circ f \circ f, \ldots$.

Now, given two natural numbers $m, n \in \mathbb{N}$, their sum $m+n \in \mathbb{N}$ is defined by:

$$m + n = s^n(m).$$

It follows that $m + 1 = s(m)$ and $m + s(n) = s(m + n)$, in particular:

$$m + (n + 1) = (m + n) + 1$$

More generally, the following can be proved using induction:

Proposition 25. *For any $m, n, p \in \mathbb{N}$:*

1. *(Associativity)* $m + (n + p) = (m + n) + p$*;*
2. *(Commutativity)* $m + n = n + m$*;*
3. *(Cancellation Law)* $m + n = m + p \Rightarrow n = p$*;*
4. *(Trichotomy) Only one of the following can occur:* $m = n$*, or* $\exists q \in \mathbb{N}$ *such that* $m = n + q$*, or* $\exists r \in \mathbb{N}$ *such that* $n = m + r$*.*

The notion of order among natural numbers can be defined in terms of addition. Namely, one writes

$$m < n,$$

if $\exists q \in \mathbb{N}$ such that $n = m + q$; in the same situation, one also writes $n > m$. Notice in particular that for every $m \in \mathbb{N}$:

$$m < s(m).$$

Finally, one writes $m \geq n$ if $m > n$ or $m = n$ and a similar definition applies to $\leq$.

Proposition 26. *For any $m, n, p \in \mathbb{N}$:*

(I) *(Transitivity)* $m < n, n < p \Rightarrow m < p$*;*

(II) *(Trichotomy) Only one of the following can occur:* $m = n$*,* $m < n$ *or* $m > n$*.*

(III) $m < n \Rightarrow m + p < n + p$*.*

The multiplication operation $m \cdot n$ will be defined in a similar way as $m+n$ was defined. Let $a_m : \mathbb{N} \to \mathbb{N}$ be the 'add m' function, $a_m(n) = n+m$. Then multiplication of two natural numbers $m \cdot n$ is defined as:

$$m \cdot 1 := m,$$
$$m \cdot (n + 1) := (a_m)^n(m).$$

So $m \cdot 2 = a_m(m) = m+m, m \cdot 3 = (a_m)^2(m) = m+m+m, \ldots$, and it follows that:

$$m \cdot (n + 1) := m \cdot n + m.$$

More generally, the following is true:

Proposition 27. *For any* $m, n, p \in \mathbb{N}$*:*

a. $m \cdot (n \cdot p) = (m \cdot n) \cdot p$*;*

b. $m \cdot n = n \cdot m$*;*

c. $m \cdot n = p \cdot n \Rightarrow m = p$*;*

d. $m \cdot (n + p) := m \cdot n + m \cdot p$*;*

e. $m < n \Rightarrow m \cdot p < n \cdot p$*.*

5 Well-ordering principle

Let $X \subseteq \mathbb{N}$. A number $m \in X$ is called **the minimum element** of X, denoted $m = \min X$, if $m \leq n$ for every $n \in X$. For example, 1 is the minimum of $\mathbb{N}$; 100 is the minimum of $\{100, 1000, 10000\}$.

Lemma 28. *If* $m = \min X$ *and* $n = \min X$ *then* $m = n$*.*

Proof. Since $m \leq p$ for every $p \in X$, $m \leq n$ in particular. Similarly, $n \leq m$ and hence $m = n$. □

The maximum element is defined similarly: $m = \max X$ if $m \geq n$, $\forall n \in X$. Notice that not all subsets $X \subseteq \mathbb{N}$ have a maximum. In fact, $\mathbb{N}$ itself doesn't have a maximum, since $m < m + 1$ for every $m \in \mathbb{N}$. The lemma above remains valid if we exchange 'minimum' by 'maximum'.

Despite not all subsets of $\mathbb{N}$ having a maximum, they do have a minimum if they are non-empty.

Theorem 29. *(Well-ordering principle) Let* $X \subseteq \mathbb{N}$ *be non-empty. Then* X *has a minimum.*

Proof. If $1 \in X$ then 1 is the minimum, so suppose $1 \notin X$. Let

$$I_n = \{\, m \in \mathbb{N} \,|\, 1 \leq m \leq n \,\},$$

and consider the set

$$L = \{\, n \in \mathbb{N} \,|\, I_n \subseteq X^c \,\}.$$

Since $1 \notin X \Rightarrow 1 \in L$. If $n \in L \Rightarrow n + 1 \in L$ then induction would imply $L = \mathbb{N}$, but $L \neq \mathbb{N}$ since $L \subseteq X^c = \mathbb{N} - X$, and $X \neq \emptyset$. We conclude that there is a m_0 such that $m_0 \in L$ but $m_0 + 1 \notin L$. It follows than $m_0 + 1$ is the minimum element of X. □

Corollary 30. *(Strong induction) Let $X \subseteq \mathbb{N}$ be a set with the following property:*

$$\forall n \in \mathbb{N}, \text{ if } X \text{ contains all } m < n \Rightarrow n \in X.$$

Then $X = \mathbb{N}$.

Proof. Set $Y = X^c$, the claim is that $Y = \emptyset$. Suppose not, that is, $Y \neq \emptyset$. By the theorem above, Y has a minimum element, say $p \in Y$. But then by hypothesis $p \in X$, a contradiction. □

Example 31. *Strong induction can be used to prove the* ***Fundamental theorem of Arithmetic****, which says that every number greater than* 1 *can written as a product of primes (a number p is* ***prime*** *if $p \neq m \cdot n$, with $m < p$ and $n < p$). Indeed, Let $X = \{ m \in \mathbb{N} \,|\, m \text{ is a product of primes} \}$ and $n \in \mathbb{N}$ a given number. If X contains all numbers m such that $m < n$, then if n is prime, $n \in X$; if n is not a prime then $n = p \cdot q$ with $p < n, q < n$, again it follows that $n \in X$. Therefore, strong induction implies $X = \mathbb{N}$.*

Let X be any set. A common way of defining a function $f : \mathbb{N} \to X$ is **by recurrence** (sometimes 'by induction' is used), namely, $f(1)$ is given and also a rule that allows one to obtain $f(m)$ knowing $f(n)$ for all $n < m$. Technically, more than one function f could exist satisfying these conditions, however it is know that such a function is unique, the proof of this fact is left as an exercise.

Example 32. *(Factorial) The factorial function $f : n \mapsto n!$ can be defined using induction. Set $f(1) = 1$ and $f(n+1) = (n+1) \cdot f(n)$. Then $f(2) = 2 \cdot 1$, $f(3) = 3 \cdot 2 \cdot 1$, $\ldots$, $f(n) = n!$.*

Example 33. *(Arbitrary sums/products) So far the definition of $m + n$ was used, what about $m + n + p$ or $m_1 + \ldots + m_n$? In order to define arbitrary sums (or products), one can use induction. Namely,*

$$m_1 + \ldots + m_n = (m_1 + \ldots + m_{n-1}) + m_n;$$

and similarly, for products:

$$m_1 \cdot \ldots \cdot m_n = (m_1 \cdot \ldots \cdot m_{n-1}) \cdot m_n.$$

6 Finite and Infinite sets

Throughout this section, I_n stands for the set of numbers less than or equal to n:

$$I_n = \{ m \in \mathbb{N} \,|\, 1 \leq m \leq n \}$$

A arbitrary set X is **finite** if $X = \emptyset$ or there is number $n \in \mathbb{N}$ and a bijection

$$f : I_n \to X.$$

In the latter case, one says that X has n elements and writes:

$$|X| = n,$$

f is said to be a counting function for X. By convention, if $X = \emptyset$ then one says X has zero elements, i.e. $|\emptyset| = 0$.

It remains to show that the number of elements is a well-defined notion, that is to say, if there are bijections $f : I_n \to X$ and $g : I_m \to X$ then $n = m$.

Theorem 34. *Let $X \subseteq I_n$. If there is a bijection $f : I_n \to X$, then $X = I_n$.*

Proof. The proof is by induction on n. The case $n = 1$ is obvious, suppose the result true for n, the proof follows if one can prove the result for $n + 1$.

Suppose $X \subseteq I_{n+1}$ and there is a bijection $f : I_{n+1} \to X$. Let $a = f(n+1)$ and consider the restriction $f : I_n \to X - \{a\}$.

If $X - \{a\} \subseteq I_n$ then $X - \{a\} = I_n$, $a = n + 1$ and $X = I_{n+1}$.

Suppose $X - \{a\} \not\subseteq I_n$, then $n + 1 \in X - \{a\}$ and one can find b such that $f(b) = n + 1$. Let $g : I_{n+1} \to X$ be the defined by $g(m) = f(m)$ if $m \neq n + 1, a$; $g(n + 1) = n + 1$; $g(b) = a$. By construction, the restriction $g : I_n \to X - \{n + 1\}$ is a bijection and obviously $X - \{n + 1\} \subseteq I_n$, hence $X - \{n + 1\} = I_n$ and it follows that $X = I_{n+1}$. □

Corollary 35. *(Number of elements is well-defined) If there is a bijection $f : I_n \to I_m$ then $m = n$. Therefore, if $f : I_n \to X$ and $g : I_m \to X$ are bijections then $n = m$.*

Proof. The first part follows directly from the theorem. For the second part, consider the composition $(f^{-1} \circ g) : I_m \to I_n$. □

Corollary 36. *There is no bijection $f : X \to Y$ between a finite set X and a proper subset $Y \subseteq X$.*

Proof. By definition there is a bijection $\varphi : I_n \to X$ for some $n \in \mathbb{N}$. Since Y is proper, $A := \varphi^{-1}(Y)$ is also proper in I_n. Let $\varphi_A : A \to Y$ be the restriction of φ from I_n to A. Suppose there is a bijection $f : X \to Y$, then the composite function $\varphi_A^{-1} \circ f \circ \varphi : I_n \to A$ defines a bijection, a contradiction. □

Theorem 37. *Let X be a finite set and $Y \subseteq X$, then Y is finite and $|Y| \leq |X|$, the equality occurs only if $X = Y$.*

Proof. It's enough to prove the result for $X = I_n$. If $n = 1$ the result is obvious. Suppose the result is valid for I_n and consider $Y \subseteq I_{n+1}$. If $Y \subseteq I_n$, the induction hypothesis gives the result, so assume $n+1 \in Y$. Then $Y - \{n+1\} \subseteq I_n$ and by induction, there is a bijection $f : I_p \to Y - \{n+1\}$, where $p \leq n$. Let $g : I_{p+1} \to Y$ be a bijection defined by $g(n) = f(n)$ if $n \in I_n$, and $g(p+1) = n+1$. This proves that Y is finite, moreover since $p \leq n \Rightarrow p+1 \leq n+1$, $|Y| \leq n$. The last statement says that if $Y \subseteq I_n$ and $|Y| = n$ then $Y = I_n$, but this is a direct consequence of theorem 34. □

The following Corollary is immediate:

Corollary 38. *Let Y be finite and $f : X \to Y$ be an injective function. Then X is also finite and $|X| \leq |Y|$.*

Corollary 39. *Let X be finite and $f : X \to Y$ be an surjective function. Then Y is also finite and $|Y| \leq |X|$.*

Proof. Since f is surjective, by the proof of proposition 22, f has an injective right-inverse $g : Y \to X$. The result follows by the corollary above. □

A set X that is not finite is said to be **infinite**. More, precisely X is infinite when it's not empty and there is no bijection $f : I_n \to X$ for any $n \in \mathbb{N}$.

Example 40. *The natural numbers $\mathbb{N}$ is an infinite set since there is no surjection between I_n and $\mathbb{N}$, because given any function $f : I_n \to \mathbb{N}$, the number $f(1) + f(2) + \ldots + f(n)$ is not in the range.*

Example 41. *$\mathbb{Z}$ and $\mathbb{Q}$ are also infinite sets since they contain $\mathbb{N}$, which is infinite.*

A set $X \subseteq \mathbb{N}$ is **bounded**, if there is a number $M \in \mathbb{N}$ such that $n \leq M$ for all $n \in X$.

Theorem 42. *Let $X \subseteq \mathbb{N}$ be nonempty. The following are equivalent:*

a. X is finite;

b. X is bounded;

c. X has a greatest element.

Proof. The proof is based on the implications a $\Rightarrow$ b, b $\Rightarrow$ c, c $\Rightarrow$ a.

(a $\Rightarrow$ b) Let $X = \{x_1, x_2, \ldots, x_n\}$. Then $M = x_1 + \ldots + x_n$ satisfies $n \leq M$ for all $n \in X$.

(b $\Rightarrow$ c) Consider the set $A = \{ n \in \mathbb{N} \,|\, n \geq x, \forall x \in X \}$. Since X is bounded, $A \neq \emptyset$. By the principle of well ordering, A has a minimum element, say $m \in A$. If $m \in X$ then m is the greatest element, so suppose $m \notin X$. By definition, $m > n$ for all $n \in X$, and since $X \neq \emptyset$, $m > 1$, that is $m = p + 1$, for some $p \in \mathbb{N}$. If $p \geq x$ for all $x \in X$ then $p \in A$, a contradiction since $p < m$ and m is minimal. If there is a $x \in X$ such that $x > p$, then $x \geq m$ a contradiction unless $x = m$, but $m \notin X$ by assumption. It follows that $m \in X$ and m is the greatest element.

(c $\Rightarrow$ a) If X has a greatest element, say M, then $X \subseteq I_M$ and it follows that X is finite.

□

The Theorem below follows directly from the definitions, the proof will be omitted.

Theorem 43. *Let X and Y be two sets such that $|X| = m, |Y| = n$ and $X \cap Y = \emptyset$. Then $X \cup Y$ is finite and $|X \cup Y| = m + n$.*

The following corollary is immediate:

Corollary 44. *Let $X_1, X_2, \ldots, X_n$, be a finite collection of sets such that each X_i is finite and $X_i \cap X_j = \emptyset$ if $i \neq j$. Then $\bigcup_{i=1}^{n} X_i$ is finite and*

$$|\bigcup_{i=1}^{n} X_i| = \sum_{i=1}^{n} |X_i|$$

.

Corollary 45. *Let* $X_1, X_2, \ldots, X_n$*, be a finite collection of sets such that each* X_i *is finite. Then* $\bigcup\limits_{i=1}^{n} X_i$ *is finite and*

$$|\bigcup_{i=1}^{n} X_i| \leq \sum_{i=1}^{n} |X_i|$$

.

Proof. For each $i = 1, \ldots, n$, set $Y_i = X_i \times \{i\}$. Then the projection

$$\pi_i : \bigcup_{i=1}^{n} Y_i \to \bigcup_{i=1}^{n} X_i$$

in the first coordinate is surjective, by Corollaries 39 and 44, the proof is complete. □

Corollary 46. *Let* $X_1, X_2, \ldots, X_n$*, be a finite collection of sets such that each* X_i *is finite. Then* $X_1 \times \ldots \times X_n$ *is finite and*

$$|X_1 \times \ldots \times X_n| = \prod_{i=1}^{n} |X_i|$$

.

Proof. It's enough to prove for $n = 2$, since the general case follows from this one. Let $X_2 = \{y_1, \ldots, y_m\}$, notice that $X_1 \times X_2 = X_1 \times \{y_1\} \cup \ldots \cup X_2 \times \{y_m\}$, the result follows by Corollary 44. □

7 Countable Sets

A set X is **countable** if it is finite or there is a bijection $f : \mathbb{N} \to X$. In the latter case, it is necessarily an infinite set, since as $\mathbb{N}$ is infinite, and we use the term **countably infinite**.

Example 47. *The set* $X = \{\, 2n \in \mathbb{N} \,|\, n \in \mathbb{N} \,\}$ *of all even numbers is countable. The function* $f(x) = 2x$ *defines a bijection between* X *and* $\mathbb{N}$*.*

Theorem 48. *Let* X *be an infinite set. Then* X *has a countably infinite subset.*

Proof. It's enough to find an injective function $f : \mathbb{N} \to X$, since every injective function is a bijection over its image. Choose an element $a_1 \in X$, set $X_1 = X - \{a_1\}$ and $f(1) = a_1$. Since X is infinite, X_1 is also infinite, choose an element a_2 in X_1, and set $f(2) = a_2$. Proceeding by induction, we have $f(n) = a_n$, $a_n \in X_{n-1}$, where $X_{n-1} = X - \{a_1, a_2, \dots, a_{n-1}\}$.

Suppose $f(n) = f(m)$, with $n, m \in \mathbb{N}$, then $a_n = a_m$, which is possible only if $n = m$. Therefore, f is injective. □

Corollary 49. *A set X is infinite $\iff$ there is a bijection $f : X \to Y$, where $Y \subsetneq X$ is a proper subset.*

Proof. (⇒) Suppose X infinite, by theorem 48, X has a countably infinite subset, say $Z = \{a_1, a_2, a_3, \dots\}$. Set $Y = (X - Z) \cup \{a_2, a_4, a_6, \dots\}$ and define $f(x) = x$ if $x \in X - Z$, and $f(a_n) = a_{2n}$ otherwise. The function $f(x)$, defined this way, is clearly a bijection.

(⇐) Follows from Corollary 36.

□

A function $f : X \to Y$ is called *increasing* if $x < y \Rightarrow f(x) < f(y)$.

Theorem 50. *Every subset X of $\mathbb{N}$ is countable.*

Proof. The proof is very similar to the one in theorem 48. If X is finite then is countable, so assume X infinite. We define an increasing bijection $f : \mathbb{N} \to X$ by induction. Let $X_1 = X$, $a_1 = \min X$ (which exists by Theorem 29), and set $f(1) = a_1$. Now, define $X_2 = X - \{a_1\}$ and $f(2) = a_2 = \min X_2$. By induction, we define $f(n) = a_n = \min X_n$, where $X_n = X - \{a_1, a_2, \dots, a_{n-1}\}$. The function $f(n)$ is injective by construction, suppose $f(n)$ not surjective. There is $x \in X$ such that $x \notin f(\mathbb{N})$. So $x \in X_n$ for every n, which implies that $x > f(n)$ for every n, and x is a bound for the infinite set $f(\mathbb{N})$, a contradiction by Theorem 42. □

Corollary 51. *Let X be a countable set. Then for any $Y \subseteq X$, Y is countable.*

Corollary 52. *The set of all prime numbers is countable.*

Corollary 53. *Let Y be a countable set and $f : X \to Y$ an injective function. Then X is countable.*

Corollary 54. *The set $\mathbb{Z}$ of integers is countable.*

Proof. The function $f : \mathbb{Z} \to \mathbb{N}$ defined by $f(0) = 1, f(m) = 2m, \text{if } m > 0$ and $f(m) = -2m + 1, \text{if } m < 0$, is bijective. □

Corollary 55. *Let X be a countable set and $f : X \to Y$ a surjective function. Then Y is countable.*

Proposition 56. *The set $\mathbb{N} \times \mathbb{N}$ is countable.*

Proof. The function defined by $f(m, n) = 2^m 3^n$ is a bijection $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. □

Corollary 57. *Let $X_1, X_2, \ldots$ be a countable collection of countable sets. Set $X = \bigcup_{i=1}^{\infty} X_i$, then X is countable.*

Proof. Let $f_i : \mathbb{N} \to X_i$ be a counting function for each $i \in \mathbb{N}$. Then $f(i, m) := f_i(m)$ defines a surjection $f : \mathbb{N} \times \mathbb{N} \to X$. By Corollary 55, X is countable. □

Corollary 58. *If X, Y are countable sets then $X \times Y$ is countable.*

Proof. Let $f_1 : \mathbb{N} \to X, f_2 : \mathbb{N} \to Y$ be counting functions. Then $f(m, n) := (f_1(m), f_2(n))$ defines a bijection, Proposition 56 concludes the proof. □

Corollary 59. *The set $\mathbb{Q}$ of rational numbers is countable.*

Proof. Let $\mathbb{Z}^*$ denote the set of nonzero integers. Define the surjective function $f : \mathbb{Z} \times \mathbb{Z}^* \to \mathbb{Q}$ given by $f(m, n) = \frac{m}{n}$. By Corollary 55, $\mathbb{Q}$ is countable. □

8 Uncountable sets

A set X is **uncountable** if it's not countable. Given two sets X and Y, if there is a bijection $f : X \to Y$, we say X and Y have the same **cardinality**, in symbols:

$$\text{card}(X) = \text{card}(Y).$$

If we assume f injective only and there is no surjective function $g : X \to Y$, then we say

$$\text{card}(X) < \text{card}(Y).$$

The cardinality of the Natural numbers $\mathbb{N}$ is denoted by

$$\text{card}(\mathbb{N}) = \aleph_0.$$

If the set X is finite with n elements, we say $\text{card}(X) = n$. By definition, for any infinite set X:

$$\aleph_0 \leq \text{card}(X).$$

Recall that given two sets X and Y, the set $\mathcal{F}(X, Y)$ denotes the set of all functions betwenn X and Y.

Theorem 60. *(Cantor) Let X and Y be sets such that Y has at least two elements. There is no surjective function $\phi : X \to \mathcal{F}(X, Y)$.*

Proof. Suppose a function $\phi : X \to \mathcal{F}(X, Y)$ is given and let $\phi_x = \phi(x) : X \to Y$ be the image of $x \in X$, which itself is a function. We claim that there is a $f : X \to Y$ that is not ϕ_x for any X. Indeed, for each $x \in X$ let $f(x)$ be an element different than $\phi_x(x)$ (this is possible sice $|Y| \geq 2$), then $f \neq \phi_x$ for every $x \in X$ and hence, ϕ is not surjective. □

Corollary 61. *Let $X_1, X_2, \ldots$ be a countable collection of countably infinite sets. Then the infinite cartesian product $X = \prod_{i=1}^{\infty} X_i$ is uncountable.*

Proof. It's enough to prove the result for $X_i = \mathbb{N}$. In this case, $X = \mathcal{F}(\mathbb{N}, \mathbb{N})$ and the result follows from Theorem 60. □

Example 62. *The set $X = \{(a_1, a_2, a_3, a_4, \ldots\}$ of all sequence of natural numbers is uncountable.*

Example 63. *The set of all real numbers $\mathbb{R}$ is uncountable. This will be proved in the next sections.*

II The real numbers $\mathbb{R}$

1 Fields

A **field** K is a set K together with two operations:

$$+ : K \times K \to K \text{ and } \cdot : K \times K \to K$$

satisfying the following properties (also called *field axioms*):

Given $x, y, z \in K$, we have:

1. $(x + y) + z = x + (y + z)$;
2. $x + y = y + x$;
3. There is an element $0 \in K$ such that $\forall x \in K$, $x + 0 = x$;
4. For any $x \in K$ there is an element $y \in K$ such that $x + y = 0$. We define $-x := y$, and write $z - x$ instead of $z + (-x)$;
5. $(x \cdot y) \cdot z = x \cdot (y \cdot z)$;
6. $x \cdot y = y \cdot x$;
7. There is an element $1 \in K$ such that $1 \neq 0$ and $\forall x \in K$, $x \cdot 1 = x$;
8. For any $x \neq 0$ there is an element $y \in K$ such that $x \cdot y = 1$. We define $x^{-1} := y$, and write $\frac{z}{x}$ instead of $z \cdot x^{-1}$;
9. $x \cdot (y + z) = x \cdot y + x \cdot z$.

Given two fields K and L, we say a function $f : K \to L$ is a *homomorphism*, if $f(x+y) = f(x)+f(y)$ and $f(c\cdot x) = c\cdot f(x)$. We say f is an *isomorphism* if, in addition, f is bijective and f^{-1} is also a homomorphism. An *automorphism* $f : K \to K$ is an isomorphism between K and itself.

Example 1. *The set rational numbers $\mathbb{Q}$ together with the operations*

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{db} \text{ and } \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$$

is a field. In this case, $0 = \frac{0}{1}$, $1 = \frac{1}{1}$ and $(\frac{a}{b})^{-1} = \frac{b}{a}$.

Example 2. *If p is prime, the set of integers mod p, $\mathbb{Z}_p = \{\bar{0}, \dots, \overline{p-1}\}$, with operations $\bar{a} + \bar{b} = \overline{a+b}$ and $\bar{a} \cdot \bar{b} = \overline{a \cdot b}$, is a field. It easy to see that $0 = \bar{0}, 1 = \bar{1}$ in this case. Moreover, by Fermat's little theorem $\bar{a} \cdot \bar{a}^{p-2} = \bar{1}$, hence $\bar{a}^{-1} = \bar{a}^{p-2}$.*

Example 3. *The set of rational functions, $\mathbb{Q}(t) = \{ \frac{p(t)}{q(t)} \,; p(t), q(t) \in \mathbb{Q}[t], q(t) \not\equiv 0 \}$, where $\mathbb{Q}[t]$ is the set of polynomials with rational coefficients, with the usual operations of fractions is a field.*

Proposition 4. *Let K be a field and $x, y \in K$, then*

a. $x \cdot 0 = 0$*;*

b. $x \cdot z = y \cdot z$ *and* $z \neq 0$ *then* $x = y$*;*

c. $x \cdot y = 0 \Rightarrow x = 0$ *or* $y = 0$*;*

d. $x^2 = y^2 \Rightarrow x = \pm y$*.*

Proof. a. Indeed, $x \cdot 0 + x = x \cdot (0 + 1) = x$, hence $x \cdot 0 = 0$.

b. We have $x = x \cdot z \cdot z^{-1} = y \cdot z \cdot z^{-1} = y$.

c. If $x \neq 0$ then $x \cdot y = 0 \cdot x \Rightarrow y = 0$.

d. Notice that $x^2 = y^2 \Rightarrow x^2 - y^2 = 0 \Rightarrow (x - y)(x + y) = 0$.

□

2 Ordered Fields

An ordered field is a field K together with a subset $P \subseteq K$, called the set of *positive elements*, such that for any $x, y \in P$ the following properties hold:

(I) *(Close under addition/multiplication)* $x + y \in P, x \cdot y \in P$;

(II) *(Trichotomy)* For any $x \in K$, only one of the following occurs: $x = 0$, $x \in P$,$-x \in P$.

If we denote $-P = \{ -p \,; p \in P \}$, then K can be written as a disjoint union

$$K = P \cup -P \cup \{0\}$$

Notice that in an ordered field if $x \neq 0$ then $x^2 \in P$. In particular $1 \in P$ in an ordered field.

Example 5. *The field of rational numbers* $\mathbb{Q}$ *together with the set*

$$P = \left\{ \frac{a}{b} \in \mathbb{Q}\,;\, a \cdot b \in \mathbb{N} \right\}$$

is an ordered field.

Example 6. *The field* $\mathbb{Z}_p$ *can't be ordered, since if we add* $\bar{1}$*,* p *times, the result is* $\bar{0}$*, i.e.* $\bar{1} + \cdot + \bar{1} = \bar{0}$*, but in an ordered field the sum of positive elements has to be positive, in particular nonzero.*

Example 7. *The field* $\mathbb{Q}(t)$ *of example 3 together with the set*

$$P = \left\{ \frac{p(t)}{q(t)}\,;\, \textit{the leading coefficient of } p(t) \cdot q(t) \textit{ is positive} \right\}$$

is an ordered field.

In an ordered field K, if $x - y \in P$ we write $x > y$ (or $y < x$). In particular, $x > 0$ implies $x \in P$ and $x < 0$ implies $x \in -P$. Notice that if $x \in P$ and $y \in -P$ then $x > y$.

We use the notation $x \leq y$ to indicate $x < y$ or $x = y$, in a similar way we can define $x \geq y$ as well.

Proposition 8. *Let* K *be an ordered field and* $x, y, z \in K$*, then*

(I) (Transitivity) $x < y$ *and* $y < z \Rightarrow x < z$*;*

(II) (Trichotomy) Only one of the following occurs: $x = y$*,* $x > y$*,*$x < y$*;*

(III) (Sum monotoneity) $x < y \Rightarrow x + z < y + z$*;*

(IV) (Multiplication monotoneity)If $z > 0$*, then* $x < y \Rightarrow x \cdot z < y \cdot z$ *and if* $z < 0$*, then* $x < y \Rightarrow x \cdot z > y \cdot z$*.*

Since in an ordered field K, 1 is always positive we have $1 + 1 > 1 > 0$ and $1 + 1 + 1 > 1 + 1$, so we can easily define an increasing injection

$$f : \mathbb{N} \to K$$

by $f(n) = \overbrace{1 + 1 + \cdots + 1}^{n}$, or more precisely, $f(1) = 1$ and $f(n+1) = f(n)+1$. Therefore, it makes sense to identify $\mathbb{N}$ with $f(\mathbb{N}) \subseteq K$, so henceforward we will simply write

$$\mathbb{N} \subseteq K$$

whenever K is an ordered field.

Notice in particular that $f(n)$ is never zero in this case, hence every ordered field is infinite. Whenever $f(n)$ is never zero, for f defined above, we say K has **characteristic zero**; if $f(p) = 0$, then we say K has **characteristic p**.

Example 9. *The field $\mathbb{Q}$ clearly has characteristic zero. The field $\mathbb{Z}_p$ has characteristic p.*

Proceeding as before, we can extend the bijection above to $f : \mathbb{Z} \to K$ and view $\mathbb{Z} \subseteq K$ as well. Hence, we have $\mathbb{N} \subseteq \mathbb{Z} \subseteq K$.

Finally, we can use $f : \mathbb{Z} \to K$ to define a bijection $g : \mathbb{Q} \to K$ by $g(\frac{a}{b}) = f(a) \cdot f(b)^{-1}$. So we may identify $\mathbb{Q}$ with $g(\mathbb{Q}) \subseteq K$ and write

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq K$$

whenever K is an ordered field.

Example 10. *If $K = \mathbb{Q}$ in the above discussion, then $g : \mathbb{Q} \to \mathbb{Q}$ is the identity automorphism. i.e. $g(\frac{a}{b}) = \frac{a}{b}$.*

Proposition 11. *(Bernoulli's inequality) Let K be an ordered field and $x \in K$. If $x \geq -1$ and $n \in \mathbb{N}$, then*

$$(1 + x)^n \geq 1 + n \cdot x$$

Proof. We use induction on $n \in \mathbb{N}$. The case $n = 1$ is clear, suppose the result valid for n. Then $(1 + x)^{n+1} = (1 + x)^n(1 + x) \geq (1 + n \cdot x)(1 + x) = 1 + x + n \cdot x + x^2 \geq 1 + x + n \cdot x$, as expected. *(Notice that we used the fact that $x \geq -1$ in the first inequality and proposition 8(IV).)* □

3 Intervals

Let K be an ordered field and $a < b$ be elements of K. We call any subset of the following form an interval:

$[a, b] = \{x \in K; a \leq x \leq b\}$ *(closed interval)*

$(a, b) = \{x \in K; a < x < b\}$ *(open interval)*

$[a, b) = \{x \in K; a \leq x < b\}$ and $(a, b] = \{x \in K; a < x \leq b\}$

$$(-\infty, b) = \{x \in K; x < b\} \text{ and } (-\infty, b] = \{x \in K; x \leq b\}$$

$$(a, \infty) = \{x \in K; a < x\} \text{ and } [a, \infty) = \{x \in K; a \leq x\}$$

$$(-\infty, \infty) = K$$

If $a = b$, then $[a, a] = a$ and $(a, a) = \emptyset$. We say the interval $[a, a]$ is degenerate.

Let K be an ordered field and $x \in K$. We define the absolute value of x, denoted by $|x|$, by

$$|x| := \max\{x, -x\},$$

which is to say, $|x|$ is the greater of the two numbers x or $-x$. Geometrically, if the elements of K are put in a straight line, $|x|$ measures the distance between x and 0, hence $|x - a|$ is the distance between x and a.

Theorem 12. *Let x, y be elements of an ordered field K. The following are equivalent:*

(i) $-y \leq x \leq y$

(ii) $x \leq y$ *and* $-x \leq y$

(iii) $|x| \leq y$

Corollary 13. *Let $x, a, \epsilon \in K$ then*

$$|x - a| \leq \epsilon \iff a - \epsilon \leq x \leq a + \epsilon.$$

Remark 2. *The theorem and corollary remains valid if we exchange $\leq$ by $<$.*

Theorem 14. *Let x, y, z be elements of an ordered field K.*

(i) $|x + y| \leq |x| + |y|$;

(ii) $|x \cdot y| = |x| \cdot |y|$;

(iii) $|x| - |y| \leq ||x| - |y|| \leq |x - y|$;

(iv) $|x - z| \leq |x - y| + |y - z|$.

Let K be an ordered field and $X \subseteq K$. An **upper bound** of X is an element $M \in K$ such that $x \leq M$ for every $x \in X$. Similarly, a **lower bound** is an element $m \in K$ such that $m \leq x$ for every $x \in X$. We say X is *bounded from above* if it has an upper bound, *bounded from below* if it has a lower bound, and *bounded* if it has upper and lower bounds, i.e. $X \subseteq [m, M]$.

Example 15. *The principle of well-ordering guarantees that $\mathbb{N}$ is bounded from below when viewed as a set inside the ordered field $\mathbb{Q}$. $\mathbb{N}$ is obviously not bounded from above in $\mathbb{Q}$, since given any n, $n+1 > n$.*

Example 16. *Oddly enough, $\mathbb{N}$ is bounded from above in the ordered field $\mathbb{Q}(t)$ from example 7. Since given any $n \in \mathbb{N}$, the rational function $r(t) = t$ satisfies $r(t) - n > 0$. Therefore, $r(t) \in \mathbb{Q}(t)$ is an upper bound for $\mathbb{N}$ and the latter is bounded from above, hence bounded, in $\mathbb{Q}(t)$.*

Theorem 17. *Let K be an ordered field. The following are equivalent:*

1. *$\mathbb{N}$ is not bounded from above;*
2. *Given $a, b \in K$, with $a > 0$, $\exists n \in \mathbb{N}$ such that $n \cdot a > b$;*
3. *Given $a > 0$ in K, $\exists n \in \mathbb{N}$ such that $0 < \frac{1}{n} < a$.*

*A field K satisfying the above conditions is called **Archimedean field**.*

Proof. The proof is based on the implications $1 \Rightarrow 2$, $2 \Rightarrow 3$, $3 \Rightarrow 1$.

$(1 \Rightarrow 2)$ Since $\mathbb{N}$ is unbounded, $\frac{b}{a} < n$ for some $n \in \mathbb{N}$, hence $n \cdot a > b$.

$(2 \Rightarrow 3)$ Take $b = 1$ in 2.

$(3 \Rightarrow 1)$ For any $a > 0$, consider $\frac{1}{a}$, by 3., $\exists n \in \mathbb{N}$ such that $\frac{1}{n} < \frac{1}{a} \iff n > a$. Therefore, no positive element is an upper bound. Similarly, no negative element can be an upper bound since if x is negative $-x$ is positive and we can apply the same argument.

□

Example 18. *Examples 15 and 16 say that $\mathbb{Q}$ is Archimedean but $\mathbb{Q}(t)$ isn't.*

4 The real numbers $\mathbb{R}$

Let K be an ordered field and $X \subseteq K$ be a bounded from above subset. The **supremum** of X, denoted $\sup X$ is the least upper bound of X, in other words, among all upper bounds $M \in K$ of X, i.e. $x \leq M$ for every $x \in X$, $\sup X \in K$ is the least of them. Therefore, $\sup X \in K$ has the following properties:

(i) (upper bound) For every $x \in X$, $x \leq \sup X$.

(ii) (least upper bound) Given any $a \in K$ such that $x \leq a$ for every $x \in X$, then $\sup X \leq a$. In other words, if $a < \sup X$ then $\exists b \in X$ such that $a < b$.

Lemma 19. *If the supremum of a set X exists, it is unique.*

Proof. Suppos $a = \sup X$ and $b = \sup X$. By (ii) above, $a \leq b$ since a is the least upper bound, but for the same reason we also have $b \leq a$, hence $a = b$. □

Lemma 20. *If a set X has a maximum element, then $\max X = \sup X$.*

Proof. Indeed, $\max X$ is obviously an upper bound and any other upper bound is greater than or equal to the maximum. □

Example 21. *Consider the set $I_n = \{1, 2, \ldots, n\} \subseteq \mathbb{Q}$. Then $\sup I_n = \max I_n = n$.*

Example 22. *Consider the set $X = \{-\frac{1}{n}; n \in \mathbb{N}\} \subseteq \mathbb{Q}$. Then $\sup X = 0$. Indeed, 0 is an upper bound and given any number $a < 0$ we can find $-\frac{1}{n}$ such that $a < -\frac{1}{n}$ since $\mathbb{Q}$ is an Archimedean field.*

Similar to the idea of supremum, the **infimum** of a bounded from below set $X \subseteq K$, denoted $\inf X$, is the greatest lower bound. The element $\inf X \in K$ has the following properties:

(i) (lower bound) For every $x \in X$, $x \geq \inf X$.

(ii) (greatest lower bound) Given any $a \in K$ such that $x \geq a$ for every $x \in X$, then $\inf X \geq a$.

The lemmas 19 and 20 extend naturally to the notion of infimum, namely, if $X \subseteq K$ has a minimum element m then $m = \inf X$. Additionally, the infimum is unique. More generally, we easily conclude that:

Proposition 23. *Let $X \subseteq K$ be a bounded subset of an ordered field K. Then,* $\inf X \in X \iff \inf X = \min X$ *and* $\sup X \in X \iff \sup X = \max X$*. In particular, every finite set has a supremum and infimum.*

Example 24. *Consider the set $X = (a, b)$, an open interval in a ordered field K. Then* $\inf X = a$ *and* $\sup X = b$*. Indeed, a is a lower bound, by definition of interval, suppose $c > a$, we claim c can't be a lower bound. For instance, consider $d = \frac{a+c}{2} \in (a, b)$. We have $d < c$ if $c < b$, hence the conclusion.*

Example 25. *Let $X = \{\frac{1}{2^n}; n \in \mathbb{N}\} \subseteq \mathbb{Q}$. Then* $\inf X = 0$ *and* $\sup X = \frac{1}{2}$*. Notice that* $\max X = \frac{1}{2}$*, by lemma 20* $\sup X = \frac{1}{2}$*. Now, 0 is obviously a lower bound. Suppose $c > 0$, since $\mathbb{Q}$ is Archimedean we can find $n \in \mathbb{N}$ such that $n + 1 > \frac{1}{c}$. By Bernoulli's inequality (Proposition 11), we have $2^n = (1 + 1)^n \geq 1 + n > \frac{1}{c}$, hence $c > \frac{1}{2^n}$ and c can't be a lower bound, so* $\inf X = 0$.

Lemma 26. *(Pythagoras) There is no $x \in \mathbb{Q}$ satisfying $x^2 = 2$.*

Proof. Suppose not, then $x = \frac{p}{q}$ satisfies $\left(\frac{p}{q}\right)^2 = 2$, or $p^2 = 2q^2$, where $p, q \in \mathbb{Z}$ and $q \neq 0$. If we decompose p^2 in prime factors, it will have an even number of factors equal to two, the same occurs for q^2. Since $2q^2$ has an odd number of factors two, we can't have $p^2 = 2q^2$. □

Proposition 27. *Consider the sets of rational numbers $X = \{x \in \mathbb{Q}; x \geq 0$ and $x^2 < 2\}$ and $Y = \{y \in \mathbb{Q}; y > 0$ and $y^2 > 2\}$. There are no rational numbers $a, b \in \mathbb{Q}$ such that $a = \sup X$ and $b = \inf Y$.*

Proof. We prove the result concerning the supremum, the result about infimum can be proven similarly. We first claim X doesn't have a maximum element. Given $x \in X$, take $r < 1$ satisfying $0 < r < \frac{2-x^2}{2x+1}$, then $x + r \in X$, so $x \in X$ can't be the maximum. Indeed, since $r < 1 \Rightarrow r^2 < r$, and we have

$$(x + r)^2 = x^2 + 2xr + r^2 < x^2 + 2xr + r = x^2 + r(2x + 1) < x^2 + 2 - x^2 = 2.$$

By a similar reasoning, given $y \in Y$, it's possible to find $r > 0$ such that $y - r \in Y$, so Y doesn't have a minimum element. Finally, notice that if $x \in X$, $y \in Y$ then $x < y$, since $x^2 < 2 < y^2 \Rightarrow 0 < (x - y)(x + y) \Rightarrow 0 < (x - y)$.

Suppose there is a number $a \in \mathbb{Q}$ such that $a = \sup X$. Then $a \notin X$, otherwise it would be its maximum. If $a \in Y$, since Y doesn't have a minimum, there would be a $b \in Y$ such that $b < a$, then $x < b < a$, a contradiction since a is the supremum. We conclude that $a \notin X$ and $a \notin Y$, so a has to satisfy $a^2 = 2$, a contradiction by lemma 26. □

Since every ordered field contains $\mathbb{Q}$, in the proposition above, if there is an ordered field K such that every nonempty bounded from above set has a supremum, then $a = \sup X$ is an element of K satisfying $a^2 = 2$.

Example 28. *(A bounded set with no supremum) Let K be a non-Archimedean field. Then, by definition, $\mathbb{N} \subseteq K$ is bounded from above. Let $M \in K$ be an upper bound for $\mathbb{N}$. So $n + 1 \leq M$ for all $n \in \mathbb{N}$, but then $n \leq M - 1$ and $M - 1$ is also an upper bound. We conclude that if M is an upper bound, $M - 1$ is one as well, hence* $\sup \mathbb{N}$ *doesn't exists in K.*

We say that an ordered field K is **complete**, if every nonempty bounded from above subset $X \subseteq K$ has a supremum in K. This motivates the following axiom (also called **the fundamental axiom of mathematical analysis**):

Axiom. There is a complete ordered field, represented by $\mathbb{R}$, called the field of real numbers.

Remark 3. *Notice that in a complete ordered field K, if $X \subseteq K$ is bounded from below then X has an infimum.*

Remark 4. *From example 28 we conclude that every complete ordered field is Archimedean.*

Proposition 29. *If K, L are complete ordered fields, then there is an isomorphism $f : K \to L$.*

The proposition above says that, in some suitable sense, $\mathbb{R}$ is the only complete ordered field.

Until the end of the semester, **every** topic we discuss will involve the complete ordered field $\mathbb{R}$ and its properties.

The discussion above leads to the conclusion that despite there is no number $x \in \mathbb{Q}$ satisfying $x^2 = 2$, there is a positive number $x \in \mathbb{R}$ such that $x^2 = 2$. We denote that number by $\sqrt{2}$. There is nothing special about 2, so we can generalize the proof above to any $n \in \mathbb{N}$ that is not a perfect square and conclude that we can find a positive number, denoted by $\sqrt{n}$, such that $(\sqrt{n})^2 = n$.

We can generalize even further and talk about the n^{th}-*root* of $m \in \mathbb{N}$, denote by $\sqrt[n]{m}$. Namely, a positive number $x \in \mathbb{R}$ such that $x^n = m$.

We call the elements of the set $\mathbb{R} - \mathbb{Q}$, **irrational numbers**. As we've just seen, there are many of them, namely, numbers of the form $\sqrt[n]{2}$, for

$n \geq 2$, are all irrational. In fact, we shall see next that irrational numbers are everywhere, in a precise sense, as a subset of the real numbers.

A subset $X \subseteq \mathbb{R}$ is said to be **dense** in $\mathbb{R}$ if for every $a, b \in \mathbb{R}$, with $a < b$, we can find $x \in X$ such that $a < x < b$. In other words, X is dense in $\mathbb{R}$ if every open non-degenerate interval (a, b) contains a point $x \in X$.

Example 30. *Let $X = \mathbb{R} - \mathbb{Z}$. Then X is dense in $\mathbb{R}$. Indeed, every open interval (a, b) is an infinite set (since $\mathbb{R}$ is ordered). On the other hand, $\mathbb{Z} \cap (a, b)$ is finite, hence we can always find a number $x \notin \mathbb{Z}$ with $x \in (a, b)$.*

Theorem 31. *The set of rational numbers, $\mathbb{Q}$, and the set of irrational numbers, $\mathbb{R} - \mathbb{Q}$, are both dense in $\mathbb{R}$.*

Proof. Let $(a, b) \in \mathbb{R}$ be a non-degenerate open interval. The idea of the proof is that since $b - a > 0$, there is a natural number $n \in \mathbb{N}$ such that $\frac{1}{n} < b - a$, hence a multiple of this number, say $\frac{m}{n}$ eventually will be in (a, b). More formally, let $X = \{m \in \mathbb{Z}; \frac{m}{n} \geq b\}$. Since $\mathbb{R}$ is Archimedean, $X \neq \emptyset$. Notice that X is bounded from below by $nb \in \mathbb{R}$. By the well ordering principle, X has a smallest element, say $m_0 \in X$. By the smallness of m_0, the number $m_0 - 1 \notin X$, so $\frac{m_0 - 1}{n} < b$. We claim $a < \frac{m_0 - 1}{n}$. Suppose not, then $\frac{m_0 - 1}{n} \leq a < b < \frac{m_0}{n}$, which implies that $b - a \leq \frac{m_0}{n} - \frac{m_0 - 1}{n} = \frac{1}{n}$, a contradiction. Therefore, the rational number $\frac{m_0 - 1}{n}$ satisfies $a < \frac{m_0 - 1}{n} < b$ and $\mathbb{Q}$ is dense in $\mathbb{R}$. We can apply the same argument *mutatis mutandis* to conclude that $\mathbb{R} - \mathbb{Q}$ is dense. Namely, instead of using $\frac{1}{n}$ in our argument, we use an irrational number, say $\frac{\sqrt{2}}{n}$. $\square$

Theorem 32. *(The nested intervals principle) Let $I_1 \supseteq I_2 \supseteq \dots I_n \supseteq \dots$ be a decreasing sequence of closed intervals of the form $I_n = [a_n, b_n]$. Then $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$, or more precisely,*

$$\bigcap_{n=1}^{\infty} I_n = [a, b],$$

where $a = \sup a_n = \sup\{a_n; n \in \mathbb{N}\}$ and $b = \inf b_n = \inf\{b_n; n \in \mathbb{N}\}$

Proof. By hypothesis, $I_n \supseteq I_{n+1}, \forall n \in \mathbb{N}$, which implies:

$$a_1 \leq a_2 \leq \dots a_n \leq \dots \leq b_n \leq \dots \leq b_2 \leq b_1.$$

Notice that a_n is bounded from above by b_1, hence the supremum of a_n, $a \in \mathbb{R}$, is well defined. Similarly, the infimum of b_n, $b \in \mathbb{R}$, is well defined. Since b_n is an upper bound for a_n, we have $a \leq b_n, \forall n \in \mathbb{N}$. On the other hand, a is also an upper bound and we conclude that

$$a_n \leq a \leq b_n, \forall n \in \mathbb{N}.$$

A similar reasoning can be applied to b, hence

$$[a, b] \subseteq I_n, \forall n \in \mathbb{N}.$$

If $x < a$, we can find a_{n_0} such that $x < a_{n_0}$, so $x \notin I_{n_0} \Rightarrow x \notin \bigcap_{n=1}^{\infty} I_n$. Similarly, if $x > b$, then we can find n_1 such that $b_{n_1} < x$, so $x \notin I_{n_1} \Rightarrow x \notin \bigcap_{n=1}^{\infty} I_n$. We conclude that $\bigcap_{n=1}^{\infty} I_n = [a, b]$. □

Theorem 33. $\mathbb{R}$ *is uncountable.*

Proof. Let $X = \{x_1, x_2, \dots\} \subseteq \mathbb{R}$ be a countable subset of $\mathbb{R}$, which we know exists by theorem 48. We claim there is always an $x \in \mathbb{R}$ such that $x \notin X$. Pick a closed interval I_1 not containing x_1, this is possible since $\mathbb{R}$ is infinite. Proceed by induction, after setting I_n not containing x_n, we select $I_{n+1} \subseteq I_n$ as a closed interval which doesn't contain x_{n+1}. Proceeding this way, we construct a nested sequence of closed intervals $I_1 \supseteq I_2 \supseteq \dots I_n \supseteq \dots$. Therefore, by theorem 32, there is at least one $x \in \mathbb{R}$ that is not in X. □

Corollary 34. *Any non-degenerate interval* $(a, b) \subseteq \mathbb{R}$ *is uncountable.*

Proof. The function $f : (0, 1) \to (a, b)$ defined by $f(x) = (b-a)x+a$ is bijective, so it suffices to prove the result for $(0, 1)$. Suppose $(0, 1)$ is countable, then $(0, 1]$ is also countable and reasoning as before, $(n, n + 1]$ is countable for every $n \in \mathbb{N}$. Then $\mathbb{R} = \bigcup_{n \in \mathbb{Z}} (n, n + 1]$ is countable, a contradiction. □

Corollary 35. *The set of irrational numbers* $\mathbb{R} - \mathbb{Q}$ *is uncountable.*

Proof. Suppose not, then $\mathbb{R} = \mathbb{Q} \cup (\mathbb{R} - \mathbb{Q})$ is countable, a contradiction. □

III Sequences and series

1 Sequences

A **sequence of real numbers**, denoted by $x_n := x(n)$, is a function $x : \mathbb{N} \to \mathbb{R}$ that associates to each natural number $n \in \mathbb{N}$, a real number $x(n) \in \mathbb{R}$. There is no universally defined notation for a sequence x_n, but here are examples of common notation found in the literature:

$$\{x_n\}_{n\in\mathbb{N}},\ \mathbf{x}_n,\ \{x_1, x_2, \ldots\},\ (x_n)$$

We say that a sequence x_n is *bounded* if there are $a, b \in \mathbb{R}$ such that

$$a \leq x_n \leq b,$$

this is equivalent of saying that $x(\mathbb{N}) \subseteq [a, b]$, i.e. $x(n)$ is bounded as a function. A sequence is *unbounded* when is not bounded.

A sequence x_n is *bounded from above* when there is $b \in \mathbb{R}$ such that $x_n \leq b$, and *bounded from below* if there is an $a \in \mathbb{R}$ such that $a \leq x_n$. Notice that a sequence is bounded if and only if is bounded from above and below.

Let $K \subseteq \mathbb{N}$ be an infinite subset. Then K is countably infinite, let $b : \mathbb{N} \to K$, given by $k \mapsto n_k$ be a bijection. Given any sequence $x : \mathbb{N} \to \mathbb{R}$, the composition $x_{n_k} := x \circ b : K \to \mathbb{R}$ is also a sequence, called a **subsequence** of x_n.

Example 1. *Let $K = \{n; n \text{ is even}\} \subseteq \mathbb{N}$ and $b(k) = 2k$. In this case, given a sequence x_n, the sequence $x_{n_k} := x_{2n}$ is a subsequence of x_n. For example, if $x_n = (-1)^n$,i.e. $\{-1, 1, -1, \ldots\}$, then x_{2n} is the constant subsequence $x_{2n} = \{1, 1, 1, \ldots\}$.*

Notice that every subsequence x_{n_k} of a bounded sequence x_n is itself bounded by definition. We say a sequence x_n is *nondecreasing* if $x_n \leq x_{n+1}, \forall n \in \mathbb{N}$, and if the inequality is strict, i.e. $x_n < x_{n+1}$, we call x_n an *increasing* sequence. We define *nonincreasing* and *decreasing* sequences in a similar way by placing $\geq$ ($>$) instead of $\leq$ ($<$).

A sequence that is either nondecreasing, nonincreasing, increasing, or decreasing will be called **monotone**.

Lemma 2. *A monotone sequence x_n is bounded $\iff$ it has a bounded subsequence.*

Proof. Only the converse is not obvious. Suppose x_{n_k} is a bounded monotone subsequence, say $x_{n_1} \leq x_{n_2} \leq \ldots \leq b$. Given any $n \in \mathbb{N}$, we can find $n_k > n$, hence $x_n \leq x_{n_k} \leq b$. □

Example 3. *$x_n = 1$, i.e. $\{1, 1, 1, \ldots\}$, is a constant, bounded, nonincreasing and nondecreasing sequence.*

Example 4. *$x_n = n$, i.e. $\{1, 2, 3, \ldots\}$, is an unbounded increasing sequence.*

Example 5. *$x_n = \frac{1}{n}$, i.e. $\{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$, is a bounded decreasing sequence, since $0 < x_n \leq 1$.*

Example 6. *$x_n = 1 + (-1)^n$, i.e. $\{0, 2, 0, 2, \ldots\}$, is a bounded sequence that is not monotone.*

Example 7. *$x_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}$ is increasing, and bounded, since $0 < x_n < 1 + 1 + \frac{1}{2} + \frac{1}{4} + \ldots + \frac{1}{2^{n-1}} < 3$. The sequence $y_n = (1 + \frac{1}{n})^n$ is related to this sequence, since by the binomial theorem $y_n \leq x_n$, therefore y_n is also bounded, $0 < y_n < 3$.*

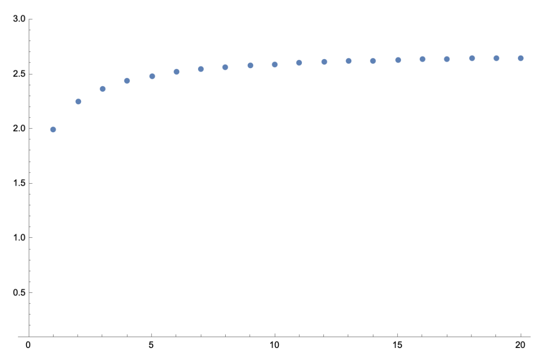


Figure 1: $y_n = (1 + \frac{1}{n})^n$

Example 8. *Let $x_1 = 0$ and $x_2 = 1$, and consider, by induction, $x_{n+2} = x_{n+1} + x_n$. It's easy to see that $0 \leq x_n \leq 1$, and moreover a quick computation shows that $x_{2n} = 1 - \left(\frac{1}{4} + \frac{1}{16} + \ldots + \frac{1}{4^{n-1}}\right)$ and $x_{2n+1} = \frac{1}{2}\left(1 + \frac{1}{4} + \frac{1}{16} + \ldots + \frac{1}{4^{n-1}}\right)$. So x_n is a bounded sequence that is not monotone.*

Example 9. *Let $a \in \mathbb{R}$ such that $0 < a < 1$. The sequence $x_n = 1 + a + a^2 + \ldots + a^n = \frac{1 - a^{n+1}}{1 - a}$ is increasing, since $a > 0$, and bounded since $0 < x_n \leq \frac{1}{1-a}$.*

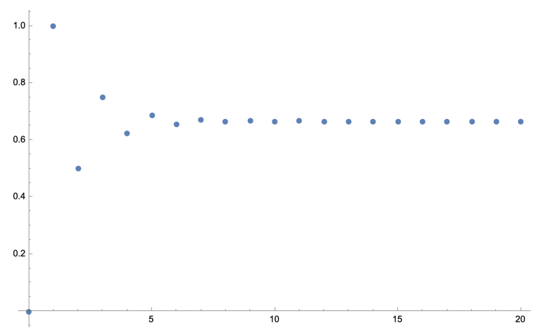


Figure 2: $x_{n+2} = x_{n+1} + x_n$

Example 10. *The sequence* $\{1, \sqrt{2}, \sqrt[3]{3}, \sqrt[4]{4}, \ldots\}$ *given by* $x_n = \sqrt[n]{n}$*, increases for* $n = 1, 2$*. We claim that starting at the third term, this sequence is decreasing. Indeed,* $x_{n+1} < x_n$ *is equivalent to* $(n+1)^n < n^{n+1}$*, which is equivalent to* $(1 + \frac{1}{n})^n < n$*, which is true for* $n \geq 3$ *by Example 7. Hence,* x_n *is bounded.*

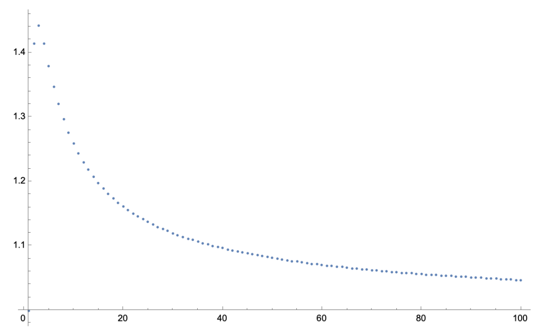


Figure 3: $x_n = \sqrt[n]{n}$

2 The limit of a sequence

Informally, to say $a \in \mathbb{R}$ is the limit of the sequence x_n is to say that the terms of the sequence are very close to a, when n is large. More precisely, we quantify this using the following definition:

$$\lim_{n\to\infty} x_n = a := \forall \epsilon > 0 \, \exists n_0 \in \mathbb{N}; n > n_0 \Rightarrow |x_n - a| < \epsilon$$

In other words: *"The limit of sequence* x_n *is* a*, if for every positive number* ϵ*, no matter how small it is, it's always possible to find an index* n_0 *such that* *the distance between x_n and a is less then ϵ, for $n > n_0$."*

Additionally, the above is the same of saying that any open interval

$$(a - \epsilon, a + \epsilon)$$

centered at a and with length 2ϵ, contains all the points of the sequence x_n except possibly a finite amount of them.

Remark 5. *It's a common practice to omit "$n \to \infty$ "and write* $\lim x_n$ *only.*

When $\lim x_n = a$, we say x_n *converges* to a, also denoted by $x_n \to a$, and call x_n *convergent*. If x_n is not convergent, we call it *divergent*, i.e. there is no $a \in \mathbb{R}$ such that $\lim x_n = a$.

Theorem 11. *(Uniqueness of the limit) If* $\lim x_n = a$ *and* $\lim x_n = b$*, then* $a = b$.

Proof. Let $\lim x_n = a$ and $b \neq a$, it's enough to prove that $\lim x_n \neq b$. Take $\epsilon = \frac{|b-a|}{2}$, then since $\lim x_n = a$, we can find n_0 such that $n > n_0 \Rightarrow |x_n - a| < \epsilon$. Therefore, $x_n \notin (b - \epsilon, b + \epsilon)$ if $n > n_0$ and we can't have $\lim x_n = b$. □

Theorem 12. *If* $\lim x_n = a$*, then for every subsequence* x_{n_k} *of* x_n*, we also have* $\lim x_{n_k} = a$.

Proof. Indeed, since given $\epsilon > 0$ it's possible to find n_0 such that $n > n_0 \Rightarrow |x_n - a| < \epsilon$, the same n_0 works for x_{n_k} as well, namely, $n_k > n_0 \Rightarrow |x_{n_k} - a| < \epsilon$. □

Corollary 13. *Let* $k \in \mathbb{N}$*. If* $\lim x_n = a$ *then* $\lim x_{n+k} = a$*, since* x_{n+k} *is a subsequence of* x_n.

In other words, Corollary 13 says that the limit of a sequence doesn't change if we omit the first k terms.

Theorem 14. *Every convergent sequence* x_n *is bounded.*

Proof. Suppose $\lim x_n = a$. Then it's possible to find n_0 such that $x_n \in (a - 1, a + 1)$ for $n > n_0$. Let $M = \max\{|x_1|, \ldots, |x_{n_0}|, |a - 1|, |a + 1|\}$, then $x_n \in (-M, M)$. □

Example 15. *The sequence* $\{0, 1, 0, 1, 0, 1, \ldots\}$ *can't be convergent by theorem 12, since it has two subsequences converging to different values, namely,* $x_{2n} = 1$ *and* $x_{2n-1} = 0$*. Also, this sequence is an example of a bounded sequence which is not convergent, illustrating the fact that the converse of theorem 14 is false.*

Theorem 16. *Every bounded monotone sequence is convergent.*

Proof. Suppose $x_n \leq x_{n+1}$, the other cases are proved similarly. Since x_n is bounded, $\sup x_n$ is well defined, say $a = \sup x_n$. Let $\epsilon > 0$ be given, then $\exists n_0 \in \mathbb{N}$ such that $a - \epsilon < x_{n_0}$, but since $x_n \leq x_{n+1}$, we must have have $a - \epsilon < x_n$, $\forall n \geq n_0$. We obviously have $x_n \leq a$, hence $a - \epsilon < x_n < a + \epsilon$ for $n > n_0$ and $\lim x_n = a$. □

Corollary 17. *If a monotone sequence x_n has a convergent subsequence then x_n is convergent.*

Example 18. *Every constant sequence $x_n = k \in \mathbb{R}$ is convergent and $\lim x_n = k$.*

Example 19. *The sequence $\{1, 2, 3, 4, \ldots\}$ is divergent because it's unbounded.*

Example 20. *The sequence $\{1, -1, 1, -1, \ldots\}$ is divergent because it has two subsequences converging to different values.*

Example 21. *The sequence $x_n = \frac{1}{n}$ is convergent and $\lim x_n = 0$, since $\mathbb{R}$ is Archimedian and given any $\epsilon > 0$ it's possible to find $n_0 \in \mathbb{N}$ such that $0 < \frac{1}{n_0} < \epsilon$. Hence, $n > n_0 \Rightarrow \frac{1}{n} < \epsilon$.*

Example 22. *Let $0 < a < 1$. The sequence $x_n = a^n$ is monotone and bounded, hence convergent. Notice that $\lim x_n = 0$ in this case.*

3 Properties of limits

Theorem 23. *Let $\lim x_n = 0$ and y_n a bounded sequence. Then*

$$\lim x_n \cdot y_n = 0.$$

Proof. Let $c > 0$ be such that $|y_n| < c$. Let $\epsilon > 0$ be given, and $n_0 \in \mathbb{N}$ a number such that $n > n_0 \Rightarrow |x_n| < \frac{\epsilon}{c}$. Then, $n > n_0 \Rightarrow |x_n y_n| < \frac{\epsilon}{c} \cdot c = \epsilon$. □

Example 24. *Using the theorem above we have $\lim\limits_{n\to\infty} \frac{\sin n}{n} = 0$*

Theorem 25. *Let $\lim x_n = a$ and $\lim y_n = b$. Then*

1. $\lim x_n + y_n = a + b$, $\lim x_n - y_n = a - b$;
2. $\lim x_n \cdot y_n = ab$;

3. If $b \neq 0$ *then* $\lim \frac{x_n}{y_n} = \frac{a}{b}$

Example 26. *Let* $a \in \mathbb{R}$ *be a positive number. The sequence* $x_n = \sqrt[n]{a}$ *is bounded and monotone, hence converges. We claim*

$$\lim \sqrt[n]{a} = 1.$$

Indeed, let $L := \lim \sqrt[n]{a}$ *and consider the subsequence* $y_n = x_{n(n+1)}$ *then*

$$L = \lim y_n = \lim a^{\frac{1}{n(n+1)}} = \lim a^{\frac{1}{n} - \frac{1}{n+1}} = \frac{\lim a^{\frac{1}{n}}}{\lim a^{\frac{1}{n+1}}} = 1$$

Example 27. *Similar to the example above is the sequence* $x_n = \sqrt[n]{n}$*. It is bounded and monotone (starting from the third term), hence converges. We claim*

$$\lim \sqrt[n]{n} = 1.$$

Let $L := \lim \sqrt[n]{n}$ *and consider the subsequence* $y_n = x_{2n}$ *then*

$$L^2 = \lim y_n \cdot y_n = \lim \sqrt[n]{2n} = \lim \sqrt[n]{2}\sqrt[n]{n} = 1 \cdot L = L$$

Hence, $L = 0$ *or* $L = 1$*, but* $L \neq 0$ *since* $x_n \geq 1$*.*

Theorem 28. *If* $\lim x_n = a$ *and* $a > 0$*, then* $\exists n_0$ *such that* $x_n > 0$ *for* $n > n_0$*. An equivalent statement is valid if* $a < 0$*, namely, up to a finite amount of indexes,* $x_n < 0$*.*

Proof. It's possible to find n_0 such that $n > n_0 \Rightarrow |x_n - a| < \frac{a}{2}$, in particular, $x > \frac{a}{2} > 0$ if $n > n_0$. The case $a < 0$ is proved similarly. □

Corollary 29. *If* x_n, y_n *are convergent sequences and* $x_n \leq y_n$ *then* $\lim x_n \leq \lim y_n$*.*

Corollary 30. *If* x_n *is convergent and* $x_n \geq a \in \mathbb{R}$ *then* $\lim x_n \geq a$*.*

Theorem 31. *(Squeeze theorem) If* $x_n \leq y_n \leq z_n$ *and* $\lim x_n = \lim z_n$*, then* $\lim y_n = \lim x_n = \lim z_n$*.*

4 $\liminf x_n$ and $\limsup x_n$

A number $a \in \mathbb{R}$ is an accumulation point of the sequence x_n, if there is a subsequence x_{n_k} such that $\lim_{k\to\infty} x_{n_k} = a$.

Theorem 32. *$a \in \mathbb{R}$ is an accumulation point of the sequence x_n if and only if $\forall \epsilon > 0$, there are infinitely many values of $n \in \mathbb{N}$ such that $x_n \in (a-\epsilon, a+\epsilon)$.*

Proof. The implication is clear, we prove the converse only. Take $\epsilon = 1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{k}, \ldots$, then it's possible to find x_{n_k} such that $|x_{n_k} - a| < \frac{1}{k}$ for every $k \in \mathbb{N}$ and moreover $n_k < n_{k+1}$, in particular, $\lim_{k\to\infty} x_{n_k} = a$. $\square$

Example 33. *If $\lim x_n = a$ then x_n has only one accumulation point, namely $a \in \mathbb{R}$. This follows directly from theorem 12.*

Example 34. *The sequence $\{0, 1, 0, 2, 0, 3, \ldots\}$ is divergent. However, it has 0 as an accumulation point, due to the constant subsequence $x_{2n-1} = 0$. Similarly, the divergent sequence $\{1, -1, 1, -1, 1, -1, \ldots\}$ has only two accumulation points: 0 and 1. The divergent sequence $\{1, 2, 3, 4, 5, 6, \ldots\}$ doesn't have an accumulation point.*

Example 35. *By theorem 31, every real number $r \in \mathbb{R}$ is an accumulation point of a sequence of rational numbers.*

We shall see below that every bounded sequence has at least two accumulation points, and the sequence converges if and only if they coincide.

Let x_n be a bounded sequence, say $m \leq x_n \leq M$, with $m, M \in \mathbb{R}$. Set

$$X_n = \{x_n, x_{n+1}, \ldots\}.$$

Then $X_n \subseteq [m, M]$ and $X_{n+1} \subseteq X_n$. Define $a_n := \inf X_n$ and $b_n := \sup X_n$, then

$$m \leq a_1 \leq a_2 \leq \ldots \leq a_n \leq \ldots \leq b_n \leq \ldots \leq b_2 \leq b_1 \leq M,$$

and the following limits are well defined $a = \lim a_n = \sup a_n$ and $b = \lim b_n = \inf b_n$. We define the *limit inferior* of x_n as

$$\liminf x_n := a$$

and the *limit superior* of x_n as

$$\limsup x_n := b.$$

We obviously have

$$\liminf x_n \leq \limsup x_n.$$

Example 36. *Consider the sequence* $x_n = \{0, 1, 0, 1, 0, 1, \ldots\}$*. Using the notation above,* $a_n \equiv 0$ *and* $b_n \equiv 1$*. Therefore,* $\liminf x_n = 0$ *and* $\limsup x_n = 1$*. More generally, we have:*

Theorem 37. *Let* x_n *be a bounded sequence. Then* $\liminf x_n$ *is the smallest accumulation point and* $\limsup x_n$ *is the greatest one.*

Proof. We prove the limit inferior claim, the other part can be proved analogously. First, we claim that $a = \liminf x_n$ is an accumulation point. Indeed, using the notation above, $a = \lim a_n$, hence given any $\epsilon > 0$, for $n > n_0$, we have $a - \epsilon < a_n < a + \epsilon$. In particular, choose $n_1 > n_0$, then $a - \epsilon < a_{n_1} < a + \epsilon$. Therefore, for $n > n_1$ we have $a_{n_1} \leq x_n < a + \epsilon$. We conclude that $a - \epsilon < x_n < a + \epsilon$, by theorem 32, a is an accumulation point. To prove the minimality, let $c < a$. We claim c is not an accumulation point. Since $c < a \Rightarrow c < a_{n_0}$, for some $n_0 \in \mathbb{N}$. Hence, $c < a_{n_0} \leq x_n$ for $n \geq n_0$. Finally, setting $\epsilon = a_{n_0} - c$, we conclude that the interval $(c - \epsilon, c + \epsilon)$ doesn't contain any x_n for $n > n_0$, by theorem 32 this concludes the proof. □

Corollary 38. *(Bolzano–Weierstrass theorem) Every bounded sequence* x_n *has a convergent subsequence.*

Proof. Since x_n is bounded, $a = \liminf x_n$ is well defined and is an accumulation point. In particular, there's a subsequence of x_n converging to a. □

Corollary 39. *A sequence* x_n *is convergent if and only if* $\liminf x_n = \limsup x_n$ *(*x_n *has a unique accumulation point)*

Proof. If x_n is convergent, all subsequences converge to the same limit, in particular $\liminf x_n = \limsup x_n = \lim x_n$. Conversely, suppose $a = \liminf x_n = \limsup x_n$. Then, using the notation above, we can find n_0 such that $a - \epsilon < a_{n_0} \leq a \leq b_{n_0} < a + \epsilon$ and $n > n_0$ implies $a_{n_0} \leq x_n \leq b_{n_0}$. We conclude that $a - \epsilon < x_n < a + \epsilon$. □

Corollary 40. *If* $c < \liminf x_n$ *then* $\exists n_0 \in \mathbb{N}$ *such that* $n > n_0 \Rightarrow c < x_n$*. Similarly, if* $c > \limsup x_n$ *then* $\exists n_1 \in \mathbb{N}$ *such that* $n > n_1 \Rightarrow c > x_n$*.*

5 Cauchy Sequences

A sequence x_n is called a **Cauchy sequence** if given $\epsilon > 0$ we can find $n_0 \in \mathbb{N}$ such that for $n, m > n_0$ we have

$$|x_n - x_m| < \epsilon$$

In other words, a Cauchy sequence is a sequence such that its terms x_n are infinitely close for sufficiently large n. It's reasonable to expect that a sequence with this property converges, and that is indeed true as the theorem below shows *(for sequences in $\mathbb{R}$, we will see in a few weeks when we talk about topology, that it's possible to construct a topological space where no Cauchy sequence converges.)*

Theorem 41. *Every Cauchy sequence is convergent.*

The proof is a direct consequence of the two lemmas below.

Lemma 42. *Every Cauchy sequence is bounded.*

Proof. By definition, we can find $n_0 \in \mathbb{N}$ such that $m, n > n_0 \Rightarrow |x_n - x_m| < 1$. Fix x_m and set $M := \max\{|x_1|, |x_2|, \ldots, |x_{n_0}|, |x_m - 1|, |x_m + 1|\}$, then $x_n \in [-M, M]$. □

Lemma 43. *If a Cauchy sequence x_n has a convergent subsequence x_{n_k} with $\lim_{k\to\infty} x_{n_k} = a$ then it converges and $\lim x_n = a$.*

Proof. Given $\epsilon > 0$, it's possible to find n_0 such that $m, n > n_0 \Rightarrow |x_n - x_m| < \frac{\epsilon}{2}$. Additionally, it's possible to find m_0 such that $n_k > m_0 \Rightarrow |x_{n_k} - a| < \frac{\epsilon}{2}$, take one $n_k > n_0$ such that this is true. Then $n > n_0 \Rightarrow |x_n - a| < |x_n - x_{n_k}| + |x_{n_k} - a| < \epsilon$. □

Now we prove the converse of the theorem above.

Theorem 44. *Every convergent sequence is a Cauchy sequence.*

Proof. Suppose $a := \lim x_n$. Then it's possible to find n_0 and n_1 such that $n > n_0 \Rightarrow |x_n - a| < \frac{\epsilon}{2}$ and $m > n_1 \Rightarrow |x_m - a| < \frac{\epsilon}{2}$. We conclude that

$$|x_n - x_m| < |x_n - a| + |x_m - a| < \epsilon,$$

for $m, n > \max\{n_0, n_1\}$. □

We conclude that

Corollary 45. *A sequence x_n of real numbers is a Cauchy sequence if and only if it converges.*

6 Infinite limits

A divergent sequence x_n *converges to infinity*, denoted by $\lim x_n = +\infty$, if for any number $M > 0$, there is $n_0 > 0$ such that $n > n_0 \Rightarrow x_n > M$. Similarly, A sequence x_n *converges to negative infinity*, denoted by $\lim x_n = -\infty$, if for any number $M > 0$, there is $n_0 > 0$ such that $n > n_0 \Rightarrow x_n < -M$.

Example 46. *The sequence $x_n = n$ converges to infinity, since given any $M > 0$, take any natural number $n_0 > M$, then $x_n = n > M$ if $n > n_0$. On the other hand, the sequence $x_n = (-1)^n n$ is divergent but doesn't converge to ∞, nor to $-\infty$, since it is unbounded from above and below, and as a consequence of the definition a sequence converges, say to $+\infty$, then it's bounded from below, and similarly, converges to $-\infty$, then it's bounded from above.*

The following theorem, similar to theorem 25 gives some properties of infinite limits. The proof will be omitted.

Theorem 47. *(Arithmetic operations with infinite limits)*

1. *If $\lim x_n = +\infty$ and y_n is bounded from below, then $\lim(x_n + y_n) = +\infty$ and $\lim(x_n \cdot y_n) = +\infty$;*
2. *If $x_n > 0$ then $\lim x_n = 0$ if and only if $\lim \frac{1}{x_n} = +\infty$;*
3. *Let $x_n, y_n > 0$ be positive sequences. Then:*
   (a) *If x_n is bounded from below and $\lim y_n = 0$ then $\lim \frac{x_n}{y_n} = +\infty$;*
   (b) *If x_n is bounded and $\lim y_n = +\infty$ then $\lim \frac{x_n}{y_n} = 0$.*

Example 48. *Let $x_n = \sqrt{n+1}$ and $y_n = -\sqrt{n}$. Then $\lim x_n = \infty$,$\lim y_n = -\infty$. We have:*

$$\lim(x_n + y_n) = \lim \sqrt{n+1} - \sqrt{n} = \lim \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} = \lim \frac{1}{\sqrt{n+1} + \sqrt{n}},$$

which gives $\lim(x_n + y_n) = 0$. However, it's ***not true in general*** *that $\lim(x_n + y_n) = \lim x_n + \lim y_n$ if both sequences have infinite limit. For example, $x_n = n^2$ and $y_n = -n$ give a counter-example, since $\lim x_n = +\infty$, $\lim y_n = -\infty$, but $\lim(x_n + y_n) = +\infty$.*

Example 49. *Let* $x_n = [2 + (-1)^n]n$ *and* $y_n = n$. *Then* $\lim x_n = \lim y_n = +\infty$, *but* $\lim \frac{x_n}{y_n} = \lim[2 + (-1)^n]$ *doesn't exists. So it's not true in general that* $\lim \frac{x_n}{y_n} = 1$ *if* $\lim x_n = \lim y_n = +\infty$.

Example 50. *Let* $a > 1$. *Then* $\lim \frac{a^n}{n} = +\infty$. *Indeed,* $a = 1 + s$ *with* $s > 0$, *so* $a^n = (1 + s)^n \geq 1 + ns + \frac{n(n-1)}{2}s^2$ *for* $n \geq 2$, *but* $\lim \frac{1+ns+\frac{n(n-1)}{2}s^2}{n} = +\infty$, *hence* $\lim \frac{a^n}{n} = +\infty$. *Arguing by induction, it's easy to show that for any* $m \in \mathbb{N}$, $\lim \frac{a^n}{n^m} = +\infty$.

Example 51. *Let* $a > 0$. *Then* $\lim \frac{n!}{a^n} = +\infty$. *Indeed, pick* $n_0 \in \mathbb{N}$ *such that* $\frac{n_0}{a} > 2$. *Then*

$$\frac{n!}{a^n} = \frac{n(n-1)\dots(n_0+1)n_0!}{a^{n_0}\underbrace{a\dots a}_{n-n_0}} > \frac{n_0!}{a^{n_0}}2^{n-n_0},$$

and it follows that $\lim \frac{n!}{a^n} = +\infty$.

7 Series

Given a sequence of real numbers x_n, the purpose of this section if to give meaning to expressions of the form, $x_1 + x_2 + x_3 + \dots$, that is, the formal sum of all the elements of the sequence x_n.

A natural way of doing this is to set $s_n := x_1 + \dots + x_n$, called *partial sums*, and define

$$\sum_{n=1}^{\infty} x_n := \lim s_n$$

It's a common practice to write $\sum x_n$ instead of $\sum\limits_{n=1}^{\infty} x_n$, and to call x_n the general term of the series. In these notes we shall adopt these conventions.

Since we define $\sum x_n$ as a limit, it may or may not exist. In case $\sum x_n = L \in \mathbb{R}$ we say that the series $\sum x_n$ converges, otherwise we say $\sum x_n$ diverges.

Theorem 52. *If the series* $\sum x_n$ *converges then* $\lim x_n = 0$.

Proof. Indeed, we have $x_n = s_n - s_{n-1}$. Therefore, $\lim x_n = \lim(s_n - s_{n-1}) = \lim s_n - \lim s_{n-1} = 0$. □

The converse of the theorem above is not true. Here's a counterexample:

Example 53. *(harmonic series) Consider the series* $\sum \frac{1}{n}$. *We obviously have* $\lim \frac{1}{n} = 0$, *however, we claim* $\sum \frac{1}{n}$ *diverges. Indeed, in order to prove that* $\lim s_n$ *diverges, it's enough to find a divergent subsequence. Take for example* s_{2^n}:

$$\begin{aligned} s_{2^n} &= 1 + \frac{1}{2} + \ldots + \frac{1}{2^n} \\ &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \ldots \\ &> 1 + \frac{1}{2} + \frac{2}{4} + \frac{4}{8} + \frac{8}{16} + \ldots + \frac{2^{n-1}}{2^n} \\ &= 1 + n \cdot \frac{1}{2} \end{aligned}$$

Hence, $s_{2^n} > 1 + n \cdot \frac{1}{2}$ *and* $\lim s_{2^n} = +\infty$.

Example 54. *(geometric series) The series* $\sum a^n$, *with* $a \in \mathbb{R}$, *diverges if* $|a| \geq 1$, *since the general term* $x_n = a^n$ *doesn't satisfy* $\lim x_n = 0$. *If* $|a| < 1$, *then* $\sum a^n$ *converges. Indeed, we can show by induction that*

$$s_n = \frac{1 - a^{n+1}}{1 - a},$$

and hence $\sum a^n = \lim s_n = \frac{1}{1-a}$, *if* $|a| < 1$.

Theorem 55. *Given series* $\sum a_n, \sum b_n$, *we have:*

1. *If* $\sum a_n$ *and* $\sum b_n$ *converge, then* $\sum(a_n + b_n)$ *converges and* $\sum(a_n + b_n) = \sum a_n + \sum b_n$.
2. *Let* $c \in \mathbb{R}$. *If* $\sum a_n$ *converges, then* $\sum c\, a_n$ *also converges, and* $\sum c\, a_n = c \sum a_n$.
3. *Suppose* $\sum a_n$ *and* $\sum b_n$ *converge, set* $c_n := \sum\limits_{i=1}^{n} a_i b_n + \sum\limits_{j=1}^{n-1} a_n b_j$. *Then* $\sum c_n$ *converges and* $\sum c_n = (\sum a_n) \cdot (\sum b_n)$.

Example 56. *(telescoping series) The series* $\sum \frac{1}{n(n+1)}$ *is convergent. Since* $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, *we easily see that* $s_n = 1 - \frac{1}{n+1}$, *so* $\sum \frac{1}{n(n+1)} = 1$.

Example 57. *The series* $\sum(-1)^n$ *is divergent since the sequence* $(-1)^n$ *has two distinct accumulation points, so it's impossible to have* $\lim(-1)^n = 0$.

Theorem 58. *Let $a_n \geq 0$ be a nonnegative sequence of real numbers. Then $\sum a_n$ converges if and only if the partial sum s_n is a bounded sequence for every $n \in \mathbb{N}$.*

Proof. The implication is clear. The converse follows from the fact that every bounded monotone sequence converges. □

Corollary 59. *(Comparison principle) Suppose $\sum a_n$ and $\sum b_n$ are series of nonnegative real numbers, i.e. $a_n, b_n \geq 0$. If there are $c \in \mathbb{R}$ and $n_0 \in \mathbb{N}$ such that $a_n \leq c\, b_n$ for $n > n_0$, then if $\sum b_n$ converges, $\sum a_n$ converges. Moreover, if $\sum a_n$ diverges then $\sum b_n$ diverges.*

Example 60. *If $r > 1$, the series $\sum \frac{1}{n^r}$ converges. Indeed, the general term of this series is positive, so the partial sums s_n are increasing, hence it's enough to prove that a subsequence of s_n is bounded. We claim s_{2^n-1} is bounded. We have:*

$$\begin{aligned} s_{2^n-1} &= 1 + \frac{1}{2^r} + \ldots + \frac{1}{(2^n-1)^r} \\ &= 1 + \left(\frac{1}{2^r} + \frac{1}{3^r}\right) + \left(\frac{1}{4^r} + \frac{1}{5^r} + \frac{1}{6^r} + \frac{1}{7^r}\right) + \ldots + \frac{1}{(2^n-1)^r} \\ &< 1 + \frac{2}{2^r} + \frac{4}{4^r} + \frac{8}{8^r} + \ldots + \frac{2^{n-1}}{2^{(n-1)r}} \\ &= \sum_{j=0}^{n-1} \left(\frac{2}{2^r}\right)^j \end{aligned}$$

On the other hand, the geometric series $\sum_{j=0}^{\infty} \left(\frac{2}{2^r}\right)^j$ converges since $\frac{2}{2^r} < 1$. We conclude that s_{2^n-1} is bounded and the claim follows.

Corollary 61. *(Cauchy's criteria) The series $\sum a_n$ is convergent if and only if given $\epsilon > 0$, there is $n_0 \in \mathbb{N}$ such that $|a_{n+1} + \ldots + a_{n+p}| < \epsilon$ for $n > n_0$.*

Proof. Notice that s_n converges if and only if it is a Cauchy sequence (see Corollary 45). □

A series $\sum a_n$ is **absolutely convergent** if $\sum |a_n|$ is convergent. A series with all of its terms positive (or negative) is convergent if and only if is absolutely convergent. Hence, in this case the two notion coincide. Here's a classical counterexample that shows that they don't coincide in general:

Example 62. *Consider the series* $\sum \frac{(-1)^n}{n}$. *We already know that* $\sum \frac{1}{n}$ *diverges, however we claim that* $\sum \frac{(-1)^n}{n}$ *converges. Indeed, notice that the subsequence* s_{2n} *satisfies*

$$s_2 < s_4 < s_6 < \ldots < s_{2n},$$

and is a Cauchy sequence, hence convergent. Whereas s_{2n-1} *satisfies*

$$s_1 > s_3 > s_5 > \ldots > s_{2n-1},$$

so it's bounded and monotone, hence convergent as well. Set $a := \lim s_{2n}$, $b := \lim s_{2n-1}$, *then since* $s_{2n} - s_{2n-1} = \frac{1}{2n} \to 0$, *we necessarily have* $a = b$. *We conclude that* s_n *has only one accumulation point, hence converges. (We will see later that* $a = b = \log 2$)

A series $\sum a_n$ is **conditionally convergent** if $\sum a_n$ is convergent but $\sum |a_n|$ is divergent. The example above shows that $\sum \frac{(-1)^n}{n}$ is conditionally convergent.

Theorem 63. *Every absolutely convergent series* $\sum a_n$ *is convergent.*

Proof. By hypothesis, $\sum a_n$ is Cauchy, so we can find $n_0 \in \mathbb{N}$ such that $n > n_0, \forall p \in \mathbb{N} \Rightarrow |a_{n+1}| + \ldots + |a_{n+p}| < \epsilon$. In particular, $|a_{n+1} + \ldots + a_{n+p}| < |a_{n+1}| + \ldots + |a_{n+p}| < \epsilon$, the conclusion follows from Cauchy's criteria (Corollary 61). ☐

Corollary 64. *Let* $\sum b_n$ *a convergent series with* $b_n \geq 0$. *If there are* $n_0 \in \mathbb{N}$ *and* $c \in \mathbb{R}$ *such that* $n > n_0 \Rightarrow |a_n| \leq c\, b_n$ *then the series* $\sum a_n$ *is absolutely convergent.*

Corollary 65. *(The root test) If there are* $n_0 \in \mathbb{N}$ *and* $c \in \mathbb{R}$ *such that* $n > n_0 \Rightarrow \sqrt[n]{|a_n|} \leq c < 1$, *then the series* $\sum a_n$ *is absolutely convergent. In other words, if* $\limsup \sqrt[n]{|a_n|} < 1$ *then* $\sum a_n$ *is absolutely convergent. On the other hand, if* $\limsup \sqrt[n]{|a_n|} > 1$, *then* $\sum a_n$ *diverges.*

Proof. In this case, we can compare $\sum |a_n|$ with $\sum c^n$, the latter (absolutely) converges since it's a geometric series with $0 < c < 1$. If $\sqrt[n]{|a_n|} > 1$ for n sufficiently large, then $\lim a_n \neq 0$. ☐

Corollary 66. *(The root test – second version) If* $\lim \sqrt[n]{|a_n|} < 1$, *then the series* $\sum a_n$ *is absolutely convergent. If* $\lim \sqrt[n]{|a_n|} > 1$, *then the series* $\sum a_n$ *is divergent.*

Example 67. *Let $a \in \mathbb{R}$ and consider the series $\sum na^n$. Notice that* $\lim \sqrt[n]{n\,|a|^n} = \lim \sqrt[n]{n} \lim |a| = |a|$. *Hence, if $|a| < 1$ the series $\sum na^n$ is absolutely convergent and if $|a| > 1$ it diverges. If $|a| = 1$ the series also diverges, since* $\lim na^n \neq 0$ *in this case.*

Theorem 68. *(The ratio test) Let $\sum a_n$ and $\sum b_n$ be series of real numbers such that $a_n \neq 0, b_n > 0, \forall n \in \mathbb{N}$ and $\sum b_n$ convergent. If there is $n_0 \in \mathbb{N}$ such that $n > n_0 \Rightarrow \left|\frac{a_{n+1}}{a_n}\right| \leq \left|\frac{b_{n+1}}{b_n}\right|$, then $\sum a_n$ is absolutely convergent.*

Proof. Consider the inequalities:

$$\left|\frac{a_{n_0+2}}{a_{n_0+1}}\right| \leq \left|\frac{b_{n_0+2}}{b_{n_0+1}}\right|$$
$$\left|\frac{a_{n_0+3}}{a_{n_0+2}}\right| \leq \left|\frac{b_{n_0+3}}{b_{n_0+2}}\right|$$
$$\cdots$$
$$\left|\frac{a_n}{a_{n-1}}\right| \leq \left|\frac{b_n}{b_{n-1}}\right|$$

Multiplying them together, we have:

$$\left|\frac{a_n}{a_{n_0+1}}\right| \leq \left|\frac{b_n}{b_{n_0+1}}\right|$$

Hence, $|a_n| \leq c\, b_n$ and the result follows by the comparison principle. □

Corollary 69. *(The ratio test – second version) If* $\limsup \left|\frac{a_{n+1}}{a_n}\right| < 1$, *then the series $\sum a_n$ is absolutely convergent. If* $\limsup \left|\frac{a_{n+1}}{a_n}\right| > 1$, *then the series $\sum a_n$ is divergent.*

Proof. For the convergence, take $b_n = (\limsup \left|\frac{a_{n+1}}{a_n}\right|)^n$ in theorem 68. If $\limsup \left|\frac{a_{n+1}}{a_n}\right| > 1$ then $\lim a_n \neq 0$. □

Corollary 70. *(The ratio test – third version) If* $\lim \left|\frac{a_{n+1}}{a_n}\right| < 1$ *then $\sum a_n$ is absolutely convergent, if* $\lim \left|\frac{a_{n+1}}{a_n}\right| > 1$ *then $\sum a_n$ diverges.*

Example 71. *Fix $x \in \mathbb{R}$ and consider the series $\sum \frac{x^n}{n!}$, then $\left|\frac{a_{n+1}}{a_n}\right| = \frac{|x|}{n+1} \to 0$ regardless of x, and the series is absolutely convergent. We will see later that this series coincides with e^x.*

Theorem 72. *(Root test is stronger than the ratio test) For any bounded sequence a_n of positive numbers we have*

$$\liminf \frac{a_{n+1}}{a_n} \leq \liminf \sqrt[n]{a_n} \leq \limsup \sqrt[n]{a_n} \leq \limsup \frac{a_{n+1}}{a_n},$$

In particular, if $\lim \frac{a_{n+1}}{a_n} = c$ then $\lim \sqrt[n]{a_n} = c$.

Proof. It's enough to prove that $\limsup \sqrt[n]{a_n} \leq \limsup \frac{a_{n+1}}{a_n}$, the first inequality can be proven *mutatis mutandis*. We argue by contradiction, suppose there is a $k \in \mathbb{R}$ such that

$$\limsup \sqrt[n]{a_n} > k > \limsup \frac{a_{n+1}}{a_n}$$

Proceeding as in the proof of theorem 68, we can find $n_0 \in \mathbb{N}$ such that $n > n_0 \Rightarrow a_n < c\, k^n$, which implies that $\sqrt[n]{a_n} < c^{\frac{1}{n}}\, k$ and hence:

$$\limsup \sqrt[n]{a_n} \leq k$$

a contradiction. ☐

Example 73. *A nice application of the theorem above is the computation of $\lim \frac{n}{\sqrt[n]{n!}}$. Set $x_n = \frac{n}{\sqrt[n]{n!}}$ and $y_n = \frac{n^n}{n!}$, then $x_n = \sqrt[n]{y_n}$. On the other hand, $\frac{y_{n+1}}{y_n} = (1 + \frac{1}{n})^n$, hence $\lim \frac{y_{n+1}}{y_n} = e$, and it follows that $\lim \frac{n}{\sqrt[n]{n!}} = e$.*

Example 74. *Given two distinct numbers $a, b \in \mathbb{R}$, consider the sequence $x_n = \{a, ab, a^2b, a^2b^2, a^3b^2, \ldots\}$, then the ratio $\frac{x_{n+1}}{x_n} = b$ if n is odd, and $\frac{x_{n+1}}{x_n} = a$ if n is even, hence the sequence $\frac{x_{n+1}}{x_n}$ doesn't converge and $\lim \frac{x_{n+1}}{x_n}$ doesn't exist. On the other hand, we have $\lim \sqrt[n]{x_n} = \sqrt{ab}$. This demonstrates that in the theorem above the inequalities can be strict.*

Theorem 75. *(Dirichlet) Let b_n be a nonincreasing sequence of positive numbers with $\lim b_n = 0$, and $\sum a_n$ be a series such that the partial sum s_n is a bounded sequence. Then the series $\sum a_n b_n$ converges.*

Proof. Notice that

$$\begin{aligned} a_1b_1 + a_2b_2 + \ldots + a_nb_n &= a_1(b_1 - b_2) + (a_1 + a_2)(b_2 - b_3) + \\ &+ (a_1 + a_2 + a_3)(b_3 - b_4) + \ldots + (a_1 + \ldots + a_n)b_n \\ &= \sum_{i=2}^{n} s_{i-1}(b_{i-1} - b_i) + s_nb_n \end{aligned}$$

Since s_n is bounded, say $|s_n| \leq k$ and $b_n \to 0$, we have $\lim s_nb_n = 0$. Moreover, $|\sum_{i=2}^{n} s_{i-1}(b_{i-1} - b_i)| \leq k|\sum_{i=2}^{n}(b_{i-1} - b_i)| = k(b_1 - b_n)$. So $\sum_{i=2}^{n} s_{i-1}(b_{i-1} - b_i)$ converges, and therefore, by comparison, $\sum a_nb_n$ converges as well. □

We can weaken the hypothesis $\lim b_n = 0$ if we take $\sum a_n$ convergent. Indeed, if $\lim b_n = c$ just take $b_n^* := b_n - c$ and use this new sequence instead. We conclude:

Corollary 76. *(Abel) If $\sum a_n$ is convergent and b_n is a nonincreasing sequence of positive numbers then $\sum a_nb_n$ converges.*

Corollary 77. *(Leibniz) Let b_n be a nonincreasing sequence of positive numbers with $\lim b_n = 0$. Then the series $\sum(-1)^nb_n$ converges.*

Proof. In this case, $a_n = (-1)^n$ has bounded partial sum, namely $|s_n| \leq 1$, and the result follows directly from theorem 75. □

Example 78. *Some periodic real valued functions can be written as a linear combination of $\sum \cos(nx)$ and $\sum \sin(nx)$. The properties of such functions and generalizations are addressed in area of mathematics called* ***Fourier Analysis****. E. Stein's book on the subject is a wonderful first-read of the topic.*

Take the example of $f(x) = \sum \frac{\cos(nx)}{n}$, we claim that if $x \neq 2\pi k$, $k \in \mathbb{Z}$ then $f(x)$ is well-defined, i.e. $\sum \frac{\cos(nx)}{n}$ converges. Indeed, let $a_n = \cos(nx)$ and $b_n = \frac{1}{n}$, then b_n is decreasing, so by theorem 75, it's enough to prove that the partial sums s_n of $\sum a_n$ are bounded. In other words, we need to show that

$$s_n = \cos(x) + \cos(2x) + \cos(3x) + \ldots + \cos(nx)$$

is bounded. Recall, that $e^{ix} = \cos(x) + i\sin(x)$*. Therefore:*

$$
\begin{aligned}
1 + s_n &= Re[1 + e^{ix} + e^{2ix} + e^{3ix} + \ldots + e^{nix}] \\
1 + s_n &= Re[\frac{1 - e^{(n+1)ix}}{1 - e^{ix}}] \\
1 + s_n &\leq \frac{2}{|1 - e^{ix}|}
\end{aligned}
$$

It follows that s_n *is bounded and we conclude that* $\sum \frac{\cos(nx)}{n}$ *converges if* $x \neq 2\pi k$*.*

Given a series $\sum a_n$, we define the *positive part* of $\sum a_n$ as the series $\sum p_n$, where $p_n = a_n$ if $a_n > 0$, and $p_n = 0$ if $a_n \leq 0$. Similarly, the *negative part* of $\sum a_n$ as the series $\sum q_n$, where $q_n = -a_n$ if $a_n < 0$, and $q_n = 0$ if $a_n \geq 0$. It follows immediately from the definition that $p_n, q_n \geq 0$ and $a_n = p_n - q_n, |a_n| = p_n + q_n \, \forall n \in \mathbb{N}$.

Proposition 79. *The series* $\sum a_n$ *is absolutely convergent if and only if* $\sum p_n$ *and* $\sum q_n$ *converge.*

Proof. Notice that $p_n \leq |a_n|$ and $q_n \leq |a_n|$, hence if $\sum |a_n|$ converge then by comparison $\sum p_n$ and $\sum q_n$ also converge. The converse is obvious. □

Example 80. *If* $\sum a_n$ *is not absolutely convergent, then the proposition is false. Take the example of* $\sum \frac{(-1)^n}{n}$*. In this case,* $\sum p_n = \sum \frac{1}{2n}$ *and* $\sum q_n = \sum \frac{1}{2n-1}$*, and both diverge.*

Proposition 81. *If* $\sum a_n$ *is conditionally convergent then* $\sum p_n$ *and* $\sum q_n$ *diverge.*

Proof. Suppose not, say $\sum q_n$ converge. Then $\sum |a_n| = \sum p_n + \sum q_n = \sum a_n + 2\sum q_n$ also converges, a contradiction. □

Let $f : \mathbb{N} \to \mathbb{N}$ be a bijection and $\sum a_n$ be a series of real numbers. Set $b_n = a_{f(n)}$. We say $\sum a_n$ is **commutatively convergent** if $\sum a_n = \sum b_n$ for every bijection $f : \mathbb{N} \to \mathbb{N}$. We will show below that the notion of commutative convergence coincides with absolute convergence.

Theorem 82. *A series* $\sum a_n$ *is absolutely convergent if and only if is commutatively convergent.*

Proof. Suppose $\sum a_n$ absolutely convergent, and let $b_n = a_{f(n)}$ for some bijection $f : \mathbb{N} \to \mathbb{N}$. It's enough to assume that $a_n \geq 0$, otherwise just use the fact that $a_n = p_n - q_n$, for $p_n, q_n \geq 0$, and apply the result for p_n and q_n. Now, fix $n \in \mathbb{N}$ and let $s_n = \sum_{i=1}^{n} a_i$ denote the partial sum of $\sum a_n$, and $t_n = \sum_{i=1}^{n} b_i$, the partial sum of $\sum b_n$. If we set $m := \max\{f(x); 1 \leq x \leq n\}$, it follows that $t_n = \sum_{i=1}^{n} a_{f(i)} \leq \sum_{i=1}^{m} a_i = s_m$. We conclude that for each $n \in \mathbb{N}$ it's possible to find $m \in \mathbb{N}$ such that $t_n \leq s_m$, and similarly using $f^{-1}(y)$ instead of $f(x)$, given $m \in \mathbb{N}$ it's possible to find $n \in \mathbb{N}$, such that $s_m \leq t_n$, which implies $\lim s_n = \lim t_n$, hence $\sum a_n = \sum b_n$.

Conversely, we want to show that if $\sum a_n$ is commutatively convergent then it is absolutely convergent. We prove the contra-positive, that is, suppose $\sum a_n$ is not absolutely convergent then $\sum a_n$ is not commutatively convergent. Indeed, if $\sum a_n$ is divergent, just take $b_n = a_n$. Otherwise, $\sum a_n$ is conditionally convergent, say $\sum a_n = S \in \mathbb{R}$, and by proposition 81, both $\sum p_n$ and $\sum q_n$ diverge. Moreover, since $\lim a_n = 0$, we have $\lim p_n = \lim q_n = 0$. Take any number $c \neq S$, we will show that we can reorder a_n into b_n in such a way that $\sum b_n = c$, hence $\sum a_n$ can't be commutatively convergent. Let n_1 be the smallest natural such that

$$p_1 + p_2 + \ldots + p_{n_1} > c,$$

and $n_2 > n_1$, be smallest number such that

$$p_1 + \ldots + p_{n_1} - q_1 - q_2 - \ldots - q_{n_2} < c.$$

Proceeding by induction, we obtain a new series $\sum b_n$, such that the partial sums t_n approach c. Indeed, for odd i we have $t_{n_i} - c \leq p_{n_i}$, be definition of n_i, and similarly, $c - t_{n_{i+1}} \leq q_{n_{i+1}}$. Since $\lim p_n = \lim q_n = 0$, we have $\lim t_{n_i} = c$. A similar argument holds for i even. $\square$

IV Topology of $\mathbb{R}$

1 Open sets

Let $X \subseteq \mathbb{R}$. A point $p \in X$ is called *an interior point* if there is an open interval (a, b), also called a *neighborhood*, such that $p \in (a, b) \subseteq X$. In other words, p is an interior point if all points sufficiently close to p remain in X.

It's easy to see that $p \in X$ is an interior point if and only if $\exists \epsilon > 0$ such that $(p-\epsilon, p+\epsilon) \subseteq X$. Equivalently, p is an interior point if and only if$\exists \epsilon > 0$ such that $|x - p| < \epsilon \Rightarrow x \in X$.

The set of all interior points of X, denoted by $\text{int}(X)$ (also by X°), is called *the interior of* X. Notice that by definition, we necessarily have $\text{int}(X) \subseteq X$.

A set $X \subseteq \mathbb{R}$ is **open** if $X = \text{int}(X)$. That is to say, every point of X is an interior point.

Example 1. *By definition if X has an interior point then it contains an open interval, in particular it is an infinite set. Hence, if $X = \{x_1, \ldots, x_n\}$ is finite then it has no interior points. Moreover, if int$(X) \neq \emptyset$ then X is uncountable since it contains an interval. Therefore,*

$$\textit{int}(\mathbb{N}) = \textit{int}(\mathbb{Z}) = \textit{int}(\mathbb{Q}) = \emptyset,$$

and they can't be open sets. Similarly, since $\mathbb{Q}$ is dense, any open interval containing an irrational point also contains a rational point, hence

$$\textit{int}(\mathbb{R} - \mathbb{Q}) = \emptyset,$$

and it's not open as well.

Example 2. *The open interval (a, b) is open. Indeed, any $x \in (a, b)$ is an interior point because (a, b) itself contains x. On the other hand, the closed interval $[a, b]$ is not open because int$([a, b]) = (a, b) \neq [a, b]$. Indeed, any open interval containing the endpoints necessarily contain points outside $[a, b]$, so the endpoints can't be interior points. Similarly, if $X = [a, b)$ or $X = (a, b]$ then int$(X) = (a, b)$*

Example 3. *The empty set $\emptyset$ is open since its interior is also empty, i.e. int$(\emptyset) = \emptyset$.*

Example 4. *The union of two open intervals* $X = (a, b) \cup (c, d)$ *is open. Indeed, any interior point of* X *has to be an interior point of* (a, b) *or* (c, d).

Theorem 5. *a) If* $A, B \subseteq \mathbb{R}$ *are open then* $A \cap B$ *is open*

b) Given an arbitrary set L*. If* $\{A_i\}_{i\in L}$ *is a family of open sets, then* $\bigcup_{i\in L} A_i$ *is open.*

Proof. a) Let $x \in A \cap B$, then we can find $a, b, c, d \in \mathbb{R}$ such that $x \in (a, b) \subseteq A$ and $x \in (c, d) \subseteq B$. Let $m := \max\{a, c\}$ and $M := \min\{b, d\}$, then $x \in (m, M) \subseteq A \cap B$.

b) Let $x \in \bigcup_{i\in L} A_i$, then there is at least one $i_0 \in L$ such that $x \in A_{i_0}$. Since A_{i_0} is open by definition, we can find a neighborhood $(a, b) \ni x$ such that $(a, b) \subseteq A_{i_0} \subseteq \bigcup_{i\in L} A_i$. We conclude that every point is an interior point.

□

Corollary 6. *Every open set* $X \subseteq \mathbb{R}$ *is a union of open intervals.*

Proof. For each $x \in X$, take an open interval $I_x \ni x$ such that $I_x \subseteq X$. Then $X = \bigcup_{x\in X} I_x$. □

Corollary 7. *If* $A_1, A_2, \ldots, A_n$ *are open sets then* $A_1 \cap A_2 \cap \ldots \cap A_n$ *is an open set.*

The corollary above is false for countably infinite intersections, take for example the open intervals $A_n = (-\frac{1}{n}, \frac{1}{n})$. Then $\bigcap_{i=1}^{\infty} A_i = \{0\}$, which is not open (since it's finite).

Example 8. *Let* $a \in \mathbb{R}$*, then the set* $X = \mathbb{R} - \{a\}$ *is open. Indeed, set* $A = (-\infty, a)$ *and* $B = (a, +\infty)$*. Then both* A *and* B *are open and* $X = A \cup B$*, hence* X *is open. More generally, we can use induction to show that* $\mathbb{R} - \{a_1, \ldots, a_n\}$ *is open.*

Before proving the next theorem, we need the following lemma:

Lemma 9. *Let* $\{I_j\}_{j\in L}$ *be a family of open intervals containing a point* $x \in \mathbb{R}$*. Then* $I = \bigcup_{j\in L} I_j$ *is itself an open interval.*

Proof. Suppose $I_j = (a_j, b_j)$. By hypothesis,

$$a_j < x < b_j, \forall j \in L.$$

Set $a := \inf a_j$ and $b := \sup b_j$ (Notice that it's possible that $a = -\infty, b = +\infty$.) We claim that $I = (a, b)$. The inclusion $I \subseteq (a, b)$ is clear. Conversely, let $y \in (a, b)$. Then by definition of supremum and infimum, we can find a_j and b_k such that $a_j < y < b_k$, if $y < b_j$ then $y \in I_j$. Otherwise, $y \geq b_j$, and $a_j < b_j \leq y$, which implies that $a_k < y < b_k$, and $y \in I_k$. In conclusion, $(a, b) \subseteq I$, hence $I = (a, b)$. □

Theorem 10. *(Structure of open sets) Every open set $X \subseteq \mathbb{R}$ can be written uniquely as a countable union of pairwise disjoints open intervals, called the interval components of X.*

Proof. Given $x \in X$, let I_x be the union of all open intervals I_j contained in X such that $I_j \ni x$. By lemma 9, I_x is an open interval. We claim that either $I_x \cap I_y = \emptyset$ or $I_x = I_y$. Indeed, if $I_x \cap I_y \neq \emptyset$ then $I_x \cap I_y$ itself is an interval containing, say x, hence $I_x \cap I_y \subseteq I_x$, and $I_y \subseteq I_x$. Similarly, $I_x \cap I_y \subseteq I_y \Rightarrow I_x \subseteq I_y$ and it follows that $I_x = I_y$.

Define $L = \{\overline{x} \in X; x \sim y \text{ if } I_x = I_y\}$, that is, L is constructed by identifying elements of X who have the same component. Then X is the union $X = \bigcup_{\overline{x} \in L} I_x$ of pairwise disjoints open intervals. In order to prove that this union is countable we define a function that associates to each $\overline{x} \in L$ a random rational number $r(\overline{x}) \in \mathbb{Q}$ contained in I_x. Since $I_x \neq I_y \Rightarrow I_x \cap I_y = \emptyset \Rightarrow r(\overline{x}) \neq r(\overline{y})$, hence the function $r : L \to \mathbb{Q}$ is injective and corollary 53 implies that L is countable.

We are left to prove uniqueness. Suppose $X = \bigcup_{i=k}^{\infty} J_k$, where J_k are open intervals, say $J_k = (a_k, b_k)$, pairwise disjoints. We claim the endpoints of J_k are not in X. Indeed, if $a_k \in X$ then $\exists J_l$ such that $a_k \in (a_l, b_l)$, but then if we set $b := \min\{b_k, b_l\}$, we have $(a_k, b) \subseteq J_k \cap J_l$, a contradiction since $J_k \cap J_l = \emptyset$. Therefore, for each $x \in J_k$, J_k is the largest open interval containing x inside X, and we must have $J_k = I_x$. □

Corollary 11. *(Connectedness of intervals) Let $I \subseteq \mathbb{R}$ be an open interval. If $I = A \cup B$, where A and B are open and $A \cap B = \emptyset$, then either $A = I$ or $B = I$ ($B = \emptyset$ or $A = \emptyset$.)*

2 Closed sets

We say a point $a \in \mathbb{R}$ is *adherent (or closure point)* of the set $X \subseteq \mathbb{R}$ if it is limit of a sequence of points in X. Every point of X is adherent to itself, since any point $x \in X$ is the limit of the constant sequence $x_n = x$.

Example 12. *Consider $X = (0, +\infty)$. Then $0 \notin X$ but 0 is an adherent point, since $0 = \lim x_n$, where $x_n = \frac{1}{n} \in X$.*

Theorem 13. *A point $a \in \mathbb{R}$ is adherent of the set $X \subseteq \mathbb{R}$ if and only if for every $\epsilon > 0$, $(a - \epsilon, a + \epsilon) \cap X \neq \emptyset$.*

Proof. Suppose a is an adherent point, say $\lim x_n = a$, where $x_n \in X$. Given any $\epsilon > 0$, we can find $n_0 \in \mathbb{N}$ such that $n > n_0 \Rightarrow x_n \in (a - \epsilon, a + \epsilon)$, in particular, $(a - \epsilon, a + \epsilon) \cap X \neq \emptyset$. Conversely, suppose $(a - \epsilon, a + \epsilon) \cap X \neq \emptyset$ for every $\epsilon > 0$. By choosing $\epsilon = \frac{1}{n}$ for each $n \in \mathbb{N}$, we are able to construct a sequence $x_n \in X$ such that $x_n \in (a - \frac{1}{n}, a + \frac{1}{n})$, and hence $\lim x_n = a$. □

Corollary 14. *A point $a \in \mathbb{R}$ is adherent of the set $X \subseteq \mathbb{R}$ if and only if every open interval $I \ni a$ we have $I \cap X \neq \emptyset$.*

Corollary 15. *Suppose $X \subseteq \mathbb{R}$ is bounded, then $\sup X$ and $\inf X$ are adherent points.*

The set of all adherent points of X, denoted by $\overline{X}$ is called the *closure* of X. A set $X \subseteq \mathbb{R}$ is **closed** if $X = \overline{X}$. In other words, a set X is closed if and only if it contains all of its adherent points.

Notice that a set $X \subseteq \mathbb{R}$ is dense in $\mathbb{R}$ if and only if $\overline{X} = \mathbb{R}$.

Example 16. *The closed interval $[a, b]$ is a closed set. Indeed, for any sequence $x_n \in [a, b]$, we must have $a \leq \lim x_n \leq b$, hence $\overline{[a, b]} = [a, b]$. Similarly, $\overline{(a, b)} = [a, b]$, since in this case the endpoints aren't in (a, b); but still, we have $a = \lim(a + \frac{1}{n})$ and $b = \lim(b - \frac{1}{n})$.*

Example 17. *Using the density of the rationals in $\mathbb{R}$ we have $\overline{\mathbb{Q}} = \mathbb{R}$ and $\overline{\mathbb{R} - \mathbb{Q}} = \mathbb{R}$.*

Theorem 18. *A set $X \subseteq \mathbb{R}$ is closed if and only if X^c is open.*

Proof. X is closed if and only if X^c doesn't contain any adherent points, which is the case if and only if $\forall x \in X^c, \exists \epsilon > 0$ such that $(x - \epsilon, x + \epsilon) \subseteq X^c$, that is to say, X^c is open. □

Corollary 19. $\mathbb{R}$ *itself and* $\emptyset$ *are closed sets.*

Corollary 20. *If* A *and* B *are closed sets then* $A \cup B$ *is closed.*

Proof. Notice that $(A \cup B)^c = A^c \cap B^c$ is open. $\square$

Corollary 21. *Let* $\{A_j\}_{j \in L}$ *be a family of closed sets. Then* $\bigcap_{j \in L} A_j$ *is closed.*

Example 22. *Arbitrary union of closed sets need not to be closed. For example, for each* $x \in (0,1)$*, the set* $\{x\}$ *is closed since it's finite, but* $\bigcup_{x \in (0,1)} \{x\} = (0,1)$ *is open.*

Theorem 23. *Let* $X \subseteq \mathbb{R}$ *be an arbitrary set. Then* $\overline{X}$ *is closed. (i.e.* $\overline{\overline{X}} = \overline{X}$*)*

Proof. Take $x \in \overline{X}^c$, then we can find an open interval $I \ni x$ such that $I \cap \overline{X} = \emptyset$, hence x in an interior point of $\overline{X}^c$. $\square$

Example 24. $\mathbb{R}$ *itself is closed, and so is* $\emptyset$*. Every finite set* $\{x_1, \ldots, x_n\} \subseteq \mathbb{R}$ *is closed, since its complement is open. Similarly,* $\mathbb{Z}$ *is closed.*

Example 25. *The sets* $\mathbb{Q}$*,* $\mathbb{R} - \mathbb{Q}$*,* $(a,b]$*,* $[a,b)$ *are not open nor closed.*

Theorem 26. *Every set* $X \subseteq \mathbb{R}$ *has a countable dense subset* D*, i.e.* $\overline{D} = X$*.*

Proof. Notice that, if we fix $n \in \mathbb{N}$, we can write $\mathbb{R} = \bigcup_{p \in \mathbb{Z}} \left[\frac{p}{n}, \frac{p+1}{n}\right)$. For each $n \in \mathbb{N}$ and $p \in \mathbb{Z}$ if $X \cap \left[\frac{p}{n}, \frac{p+1}{n}\right) \neq \emptyset$, choose a number $x_{np} \in X \cap \left[\frac{p}{n}, \frac{p+1}{n}\right)$, and let D be the set of all such x_{np}. By construction, D is countable. We claim $\overline{D} = X$. Indeed, let I be an open interval of length $\epsilon > 0$ containing a point $x \in X$. For n sufficiently large such that $\frac{1}{n} < \epsilon$, we can find a $p \in \mathbb{Z}$ such that $\left[\frac{p}{n}, \frac{p+1}{n}\right) \subseteq I$, and hence $x_{np} \in I$. $\square$

A point $a \in \mathbb{R}$ is an *accumulation point* of the set $X \subseteq \mathbb{R}$ if $a = \lim x_n$, for $x_n \in X$ and x_n is sequence with pairwise disjoint elements. Alternatively, every open interval containing a contains points of X other than a itself.

The set of all accumulation points of X is called the *derived set* of X, denoted by X'.

We easily see that if $X' \neq \emptyset$ then X is infinite.

Example 27. *Let* $X = \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$*. Then* $X' = \{0\}$*.*

Example 28. $(a,b)' = [a,b]$. *Also,* $\mathbb{Q}' = (\mathbb{R} - \mathbb{Q})' = \mathbb{R}' = \mathbb{R}$, *whereas* $\mathbb{Z}' = \emptyset$.

Given a point $a \in \mathbb{R}$ and a set $X \subseteq \mathbb{R}$. We say a is an *isolated point* of X if a is not an accumulation point. In other words, a is isolated if we can find an open interval $I \ni a$ such that $I \cap X = \{a\}$.

Example 29. *Every natural number* $n \in \mathbb{N}$ *is isolated. More generally, every* $n \in \mathbb{Z}$ *is isolated.*

Theorem 30. *For every* $X \subseteq \mathbb{R}$, *we have*

$$\overline{X} = X \cup X'.$$

Proof. Since $X \subseteq \overline{X}$ and $X' \subseteq \overline{X}$, we have $X \cup X' \subseteq \overline{X}$. Conversely, let $a \in \overline{X}$. Then every open interval I containing a also contains points of X, either a itself or a point different from a, hence $a \in X \cup X'$. □

Corollary 31. *A set* X *is closed if and only if* $X' \subseteq X$.

Corollary 32. *If all the points of* X *are isolated then* X *is countable.*

Proof. Let D be a countable dense subset of X, i.e. $\overline{D} = X$, and $x \in X$. By definition, any interval containing x contains points of D, since x is isolated, that can only happen if $x \in D$. Hence $X = D$. □

We need the following lemma to prove the next theorem.

Lemma 33. *Let* $X \subseteq \mathbb{R}$ *be a closed nonempty set with no isolated points. Then* $\forall x \in \mathbb{R}$, $\exists I_x \subseteq X$, *a closed bounded nonempty subset with no isolated points, such that* $x \notin I_x$.

Proof. Since X is infinite, we can find a point $y \in X$, with $y \neq x$. Take a interval $(a,b) \subseteq \mathbb{R}$ such that $x \notin [a,b]$ and $y \in (a,b)$. Set $A = (a,b) \cap X$, then $A \subseteq X$ is bounded and nonempty. The set $I_x = \overline{A}$ satisfies the desired properties. □

Theorem 34. *Let* $X \subseteq \mathbb{R}$ *be a nonempty closed set such that* $X' = X$ *(*X *has no isolated points). Then* X *is uncountable.*

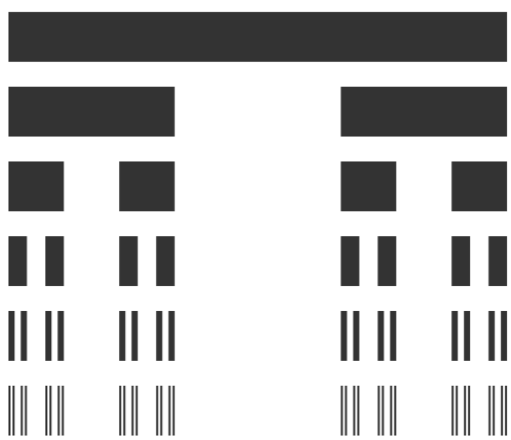
Proof. The proof is based on lemma 33 applied inductively in the following way: Let $\{x_1, x_2, \ldots\}$ be any countable subset of X. We use the lemma to find $I_1 \subseteq X$ such that $x_1 \notin I_1$, and proceed inductively by finding $I_n \subseteq I_{n-1}$

such that $x_n \notin I_n$. Choose $y_n \in I_n$ for each n. Then the sequence y_n is bounded, by Bolzano-Weierstrass theorem, it has a converging subsequence, say $y_{n_k} \to y$. For n sufficiently large we have $y \in I_n$, hence $y \in I_n$ for every $n \in \mathbb{N}$, since the I_n are nested, and moreover $y \neq x_n$ by construction. We conclude that it's impossible for X to be $\{x_1, x_2, \ldots\}$, a countable set. □

Corollary 35. *(The contrapositive version) If X is a closed countable nonempty set then X has an isolated point.*

3 The Cantor set

The Cantor set is a bounded set $K \subseteq [0,1]$ defined in the following way: Start with the interval $[0,1]$ and remove the middle third open interval $(\frac{1}{3}, \frac{2}{3})$. We are left with $[0, \frac{1}{3}]$ and $[\frac{2}{3}, 1]$. Proceed inductively, removing the middle third of each interval obtained in the previous interation, what is left is the Cantor set K.



For example, the numbers $\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \ldots$ which are endpoints of removed intervals in each iteration are elements of the Cantor set K. So K has a countable subset. Interesting enough, those are not the only points of K, as a matter of fact most points of K are not endpoints of removed intervals, and it turns out the K is actually uncountable as we shall see.

Since in each iteration we remove a finite amount of intervals, the number of intervals removed is countable. If we denote each open interval removed by I_j, then

$$K = [0,1] - \bigcup_{j=1}^{\infty} I_j = [0,1] \cap \left(\mathbb{R} - \bigcup_{j=1}^{\infty} I_j \right).$$

Since K is the union of two closed sets, it is closed.

Lemma 36. *K doesn't have interior points, i.e. int$(K) = \emptyset$.*

Proof. K doesn't have any open intervals, because after each interaction the remaining intervals shrink, so it's impossible to exists an interval $I \subseteq K$ of length l, for any $l \in \mathbb{R}$. Hence, K doesn't have interior points. □

Lemma 37. *Let R be the set of endpoints of removed intervals in each iteration. Then R is dense in K, i.e. $\overline{R} = K$.*

Proof. We have to show that given any $x \in K$, for every $\epsilon > 0$, we must have $(x-\epsilon, x+\epsilon) \cap R \neq \emptyset$. If $\epsilon > \frac{1}{2}$, the result is immediate, so let's assume $\epsilon \leq \frac{1}{2}$. At least one of intervals, $(x-\epsilon, x]$ or $[x, x+\epsilon)$, is entirely contained in $[0,1]$, say $(x-\epsilon, x]$. After the n-th iteration, only intervals of length $\frac{1}{3^n}$ are left, hence when $\frac{1}{3^n} < \epsilon$, part of $(x-\epsilon, x]$ will be removed (or was removed already previously), and it can't be the whole $(x-\epsilon, x]$ because $x \in K$. Hence, the endpoint of the removed interval is the point of R we are looking for. □

Corollary 38. *K is uncountable.*

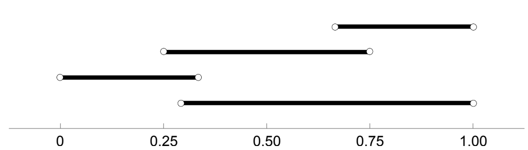
Proof. It follows directly from lemma 37 and theorem 34. □

4 Compact Sets

A open *cover* of a set $X \subseteq \mathbb{R}$ is a collection $\mathcal{C} = \{U_j\}_{j \in L}$ (not necessarily countable) of open sets $U_j \subseteq \mathbb{R}$, such that $X \subseteq \bigcup_{j \in L} U_j$. A *subcover* $\mathcal{C}'$ of $\mathcal{C}$ is a collection formed by sub-indexes $L' \subseteq L$, that is, $\mathcal{C}' = \{U_j\}_{j \in L'}$, such that $X \subseteq \bigcup_{j \in L'} U_j$.

A set $X \subseteq \mathbb{R}$ is called **compact**, if every open cover has a finite subcover, that is to say, we can take L' a finite set in the definition above.

Example 39. *Let $X = (\frac{7}{24}, 1)$. The sets $U_1 = (0, \frac{1}{3}), U_2 = (\frac{1}{4}, \frac{3}{4}), U_3 = (\frac{2}{3}, 1)$ form a (finite) open cover of X, since $X \subseteq U_1 \cup U_2 \cup U_3$. Also, $U_2 = (\frac{1}{4}, \frac{3}{4})$ and $U_3 = (\frac{2}{3}, 1)$ form a subcover, since it is still true that $X \subseteq U_2 \cup U_3$*



Example 40. *Consider the set $X = \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$, which has all of its points isolated, so it's possible to find an open interval I_n around each point $\frac{1}{n} \in X$, such that $I_n \cap \{\frac{1}{n}\} = \{\frac{1}{n}\}$. Therefore, $\mathcal{C} = \{I_n\}_{n\in\mathbb{N}}$ forms an open cover of X, and moreover, $\mathcal{C}$ doesn't have any open subcover, since if we remove at least one I_n of $\mathcal{C}$, it ceases to be a cover in the first place.*

Theorem 41. *(Borel-Lebesgue Theorem – simple version) Any closed interval $[a, b] \subseteq \mathbb{R}$ is compact.*

Proof. We need to prove that any open cover $\mathcal{C} = \{I_j\}_{j\in L}$ of $[a, b]$ has a finite subcover. We may assume that I_j are open intervals, since each I_j is open, so it has to contain an interval around each point.

Let X be the set of all points $x \in [a, b]$ such that $[a, x]$ can be cover be finitely many I_j. Notice that $X \neq \emptyset$, since $a \in X$. Set $c = \sup X$, we claim $c = b$. First, we prove $c \in X$. Indeed, $c \leq b$, so we can find $I_{j_0} = (a_0, b_0)$ covering c. Since $c > a_0$, we can find $a_0 < x \leq c$ such that $[a, x] \subseteq I_1 \cup \ldots \cup I_n$, but then $[a, c] \subseteq I_1 \cup \ldots \cup I_n \cup I_{j_0}$, hence $c \in X$. If $c < b$, then we can find $c' \in I_{j_0}$ such that $c < c' < b$. But then $[a, c']$ would still be covered by $I_1 \cup \ldots \cup I_n \cup I_{j_0}$, and c isn't an upper bound, a contradiction. □

Corollary 42. *(Borel-Lebesgue Theorem – classical version) Any bounded and closed set $X \subseteq \mathbb{R}$ is compact.*

Proof. Since X is closed, its complement $X^c = \mathbb{R} - X$ is open. Moreover, we can find $[a, b] \supseteq X$, because X is also bounded. Let $\mathcal{C} = \{I_j\}_{j\in L}$ be a open cover of X, then $\mathcal{C} \cup X^c$ is an open cover of $[a, b]$, by the theorem above we can extract $I_{j_1} \cup \ldots \cup I_{j_n} \cup X^c$, a finite subcover of $[a, b]$. Thus $I_{j_1} \cup \ldots \cup I_{j_n}$ is a finite subcover of X. □

Example 43. *The real line $\mathbb{R}$ is not compact. Indeed, consider the cover $\mathbb{R} = \bigcup_{n=1}^{\infty} (-n, n)$. Any finite subcover would be equal to the largest interval since they are nested, and hence can't cover the whole line. Similarly, $(0, 1]$ is not compact either, if we consider the nested cover $\bigcup_{n=1}^{\infty} (\frac{1}{n}, 2)$, we can argue like before.*

Theorem 44. *(Heine–Borel theorem) Let $K \subseteq \mathbb{R}$. The following are equivalent:*

1. *K is closed and bounded;*
2. *K is compact;*
3. *Every infinite subset of K has an accumulation point in K;*
4. *(Sequential compactness) Every sequence $x_n \in K$ has a convergent subsequence with limit in K.*

Proof. We already know that $1 \Rightarrow 2$. We first prove $2 \Rightarrow 3$. It's easy to show the contrapositive of 3, namely, if $X \subseteq K$ doesn't have accumulation points in K then X is finite. Indeed, we can find for each $x \in K$ an interval I_x such that $I_x \cap X = \emptyset$ if $x \notin X$, and $I_x \cap X = \{x\}$ if $x \in X$. Then $\bigcup I_x$ is a cover of K, by compactness, we extract a finite subcover, say $I_{x_1} \cup \ldots I_{x_n}$, but this would force $X = \{x_1, \ldots, x_n\}$,i.e. X is finite.

We now show $3 \Rightarrow 4$. Consider the set $X = \{x_1, x_2, \ldots\}$ formed by elements of the sequence $x_n \in K$. If X is finite then at least one member of the sequence repeat itself infinitely many times, hence forms a constant (convergent) subsequence. Otherwise, by hypothesis we have some $a \in X'$ that is also in K. Equivalently, every neighborhood of $a \in K$ contains point of the sequence x_n, hence a subsequence of x_n converges to a.

Finally, we show $4 \Rightarrow 1$. The proof is by contradiction, namely, suppose K is not bounded or not closed. If K is not closed, at least one sequence x_n converges to a point outside K, so any subsequence of this sequence would also converge to point not in K, a contradiction. If K is not bounded we can easily construct an unbounded sequence, say K is unbounded from above, then construct a sequence satisfying $x_n + 1 < x_{n+1}$, and any subsequence would also be increasing and unbounded, hence can't converge. □

Corollary 45. *(Bolzano-Weierstrass alternative version) Every infinite bounded set $X \subseteq \mathbb{R}$ has an accumulation point.*

Proof. Apply theorem 44 to $\overline{X}$. □

Corollary 46. *Let $K_1 \supseteq K_2 \supseteq \ldots$ be a nested sequence of nonempty compact sets. Then $\bigcap_{j=1}^{\infty} K_j$ is compact and nonempty.*

Example 47. *The Cantor set K is compact since it's closed and bounded. Every finite set is compact. $\mathbb{Z}$ is not compact because it's unbounded, nor is $\mathbb{R}$ itself. $\mathbb{Q} \cap [0, 1]$ is bounded but it's not compact because it's not closed.*

V Limits

1 The limit of a function

Let $f : X \subseteq \mathbb{R} \to \mathbb{R}$ be a function of a real variable, and $a \in X'$. We say the number $L \in \mathbb{R}$ is the limit of $f(x)$ as x approaches a, denoted by

$$\lim_{x \to a} f(x) = L,$$

if given $\epsilon > 0$, we can find $\delta > 0$, such that for every $x \in X$:

$$0 < |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon.$$

In other words, $f(x)$ can be made arbitrarily close to L by choosing $x \neq a$ in a sufficiently small neighborhood $(a - \delta, a + \delta)$ of a.

Notice that $a \in X'$ is an accumulation point, so the definition makes sense even if $a \notin X$. In fact, most interesting cases are when $a \notin X$. If a is not an accumulation point, i.e. an isolated point, then the same definition would imply that every number $L \in \mathbb{R}$ is a limit! Hence, the definition only makes sense if $a \in X'$.

Theorem 1. *(Uniqueness of limits) Let $X \subseteq \mathbb{R}$, $f : X \to \mathbb{R}$ and $a \in X'$. If $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} f(x) = M$, then $L = M$.*

Proof. Given any $\epsilon > 0$, we can find δ, γ such that

$$|x - a| < \delta \Rightarrow |f(x) - L| < \frac{\epsilon}{2}, \text{ and } |x - a| < \gamma \Rightarrow |f(x) - M| < \frac{\epsilon}{2}$$

Let $\alpha = \min\{\delta, \gamma\}$ then

$$|x - a| < \alpha \Rightarrow |L - M| \leq |L - f(x)| + |f(x) - M| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

This is only possible if $L - M = 0 \Rightarrow L = M$. □

Theorem 2. *(Restriction of limits) Let $Y \subseteq X \subseteq \mathbb{R}$, $f : X \to \mathbb{R}$, $a \in X' \cap Y'$. Consider the restriction $g : Y \to \mathbb{R}$ given by $g(x) = f(x)$ (Also written as $f_{|Y}(x)$). If $\lim_{x \to a} f(x) = L$ then $\lim_{x \to a} g(x) = L$.*

Proof. Self-evident. □

Theorem 3. *(Local boundedness) If* $\lim_{x\to a} f(x) = L$*, then* $\exists M > 0, \delta > 0$ *such that* $0 < |x - a| < \delta \Rightarrow |f(x)| < M$.

Proof. Take $\epsilon = 1$ in the definition. Then we can find $\delta > 0$ such that $0 < |x - a| < \delta \Rightarrow |f(x) - L| < 1 \Rightarrow |f(x)| < |L| + 1 =: M$. □

Theorem 4. *(Squeeze-theorem) Let* $X \subseteq \mathbb{R}$*,* $f, g, h : X \to \mathbb{R}$ *and* $a \in X'$*. If for every* $x \neq a$*:*

$$f(x) \leq g(x) \leq h(x),$$

then

$$\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = L \Rightarrow \lim_{x\to a} g(x) = L$$

Proof. We can find $\delta, g > 0$ such that $0 < |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon \Rightarrow L - \epsilon < f(x)$, and $0 < |x - a| < \gamma \Rightarrow |h(x) - L| < \epsilon \Rightarrow h(x) < L + \epsilon$.

Hence, if we set $\alpha = \min\{\delta, \gamma\}$ then $0 < |x - a| < \alpha \Rightarrow L - \epsilon < f(x) \leq g(x) \leq h(x) < L + \epsilon \Rightarrow |g(x) - a| < \epsilon$. □

Theorem 5. *(Monotonicity preservation) Let* $X \subseteq \mathbb{R}$*,* $f, g : X \to \mathbb{R}$ *and* $a \in X'$*. If* $\lim_{x\to a} f(x) = L$ *and* $\lim_{x\to a} g(x) = M$ *and* $L < M$ *then there exists* $\delta > 0$*, such that* $0 < |x - a| < \delta \Rightarrow f(x) < g(x)$.

Proof. Set $\epsilon := \frac{M-L}{2}$. There exists $\delta > 0$ such that $0 < |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon$ and $|g(x) - M| < \epsilon$. It follows that, $f(x) < \epsilon + L < g(x)$. □

Corollary 6. *If* $\lim_{x\to a} f(x) > 0$*, then there exists* $\delta > 0$ *such that* $0 < |x-a| < \delta \Rightarrow f(x) > 0$.

Corollary 7. *If* $f(x) \leq g(x)$ *for every* x*, then* $\lim_{x\to a} f(x) \leq \lim_{x\to a} g(x)$.

Theorem 8. *(Equivalent definition of limit) Let* $X \subseteq \mathbb{R}$*,* $f : X \to \mathbb{R}$ *and* $a \in X'$*. Then* $\lim_{x\to a} f(x) = L$ *if and only if for every sequence* $x_n \in X - \{a\}$*, with* $x_n \to a$*, we have* $\lim_{x\to a} f(x_n) = L$.

Proof. Suppose $\lim_{x\to a} f(x) = L$ and $x_n \to a$. Given $\epsilon > 0$, there exists $\delta > 0$ and $n_0 \in \mathbb{N}$ such that $0 < |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon$ and $n > n_0 \Rightarrow 0 < |x_n - a| < \delta$. Therefore, $n > n_0 \Rightarrow |f(x_n) - L| < \epsilon$.

Conversely, suppose $f(x_n) \to L$ for every $x_n \to a$ but $\lim_{x\to a} f(x) \neq L$. There exists $\epsilon > 0$, such that we can find a sequence $x_n \in X - \{a\}$ satisfying $0 < |x_n - a| < \frac{1}{n} \Rightarrow |f(x_n) - L| \geq \epsilon$, but then this sequence converges to a, yet it's not true that $f(x_n) \to L$, a contradiction. □

Corollary 9. *(Properties of limits) Let $X \subseteq \mathbb{R}$, $f, g : X \to \mathbb{R}$ and $a \in X'$.*

1. $\lim\limits_{x\to a}[f(x) \pm g(x)] = \lim\limits_{x\to a} f(x) \pm \lim\limits_{x\to a} g(x)$

2. $\lim\limits_{x\to a}[f(x) \cdot g(x)] = \lim\limits_{x\to a} f(x) \cdot \lim\limits_{x\to a} g(x)$

3. *Suppose* $\lim\limits_{x\to a} g(x) \neq 0$ *then* $\lim\limits_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x\to a} f(x)}{\lim\limits_{x\to a} g(x)}$

4. *Suppose* $\lim\limits_{x\to a} f(x) = 0$ *and* $|g(x)| \leq M$ *then* $\lim\limits_{x\to a}[f(x) \cdot g(x)] = 0$.

Proof. We proved the equivalent result for sequences, the result then follows by theorem 8. □

Example 10. *It follows from the definition of limit that $\lim\limits_{x\to a} x = a$. Similarly, using the properties of limits (Corollary 9), we obtain $\lim\limits_{x\to a} x^2 = a^2$. Proceeding by induction, we conclude that $\lim\limits_{x\to a} x^n = a^n$, and hence for every polynomial $p(x) \in \mathbb{R}[x]$, $\lim\limits_{x\to a} p(x) = p(a)$. Similarly, for any rational function $r(x) = \frac{p(x)}{q(x)}$, if $q(a) \neq 0$ then $\lim\limits_{x\to a} \frac{p(x)}{q(x)} = \frac{p(a)}{q(a)}$.*

Example 11. *Consider the function:*

$$f(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q} \\ 0, & \text{if } x \in \mathbb{R} - \mathbb{Q} \end{cases}$$

Then for any $a \in \mathbb{R}$, the limit $\lim\limits_{x\to a} f(x)$ doesn't exist. Indeed, given any real number a we can construct two sequences $x_n \in \mathbb{Q}$ and $y_n \in \mathbb{R} - \mathbb{Q}$, with $x_n \to a$ and $y_n \to a$. Therefore, $f(x_n) \to 1$ but $f(y_n) \to 0$, so $\lim\limits_{x\to a} f(x)$ doesn't exist.

Example 12. *Consider the function $f : \mathbb{R} - \{0\} \to \mathbb{R}$ given by $f(x) = \sin(\frac{1}{x})$. We claim $\lim\limits_{x\to 0} f(x)$ doesn't exist. It's enough to find two sequences $x_n \to 0$ and $y_n \to 0$ such that $f(x_n)$ and $f(y_n)$ converge to different limits. Take $x_n = \frac{1}{n\pi}$ and $y_n = (\frac{\pi}{2} + 2n\pi)^{-1}$, then $f(x_n) \to 0$ but $f(y_n) \to 1$.*

2 One sided and infinite limits

Let $X \subseteq \mathbb{R}$ and $a \in \mathbb{R}$. We say a is *accumulation point to the right* (or one-sided right accumulation point) if for every $\epsilon > 0$, $(a, a+\epsilon) \cap X \neq \emptyset$. Similarly, a is *accumulation point to the left* if for every $\epsilon > 0$, $(a-\epsilon, a) \cap X \neq \emptyset$.

We denote $X'_+ (X'_-)$, the set of all accumulation points to the right (left) of X. The definition of limit can be extended in this scenario as well. For example, let $X \subseteq \mathbb{R}$, $f : X \to \mathbb{R}$ and $a \in X'_+$, then we write

$$\lim_{x \to a^+} f(x) = L$$

If $\forall \epsilon > 0, \exists \delta > 0,\ 0 < x - a < \delta \Rightarrow |f(x) - L| < \epsilon$. We define $\lim_{x \to a^-} f(x) = L$ analogously.

Theorem 13. *Let $X \subseteq \mathbb{R}$, $f : X \to \mathbb{R}$ and $a \in X'$. Then $\lim_{x \to a} f(x) = L$ if and only if $\lim_{x \to a^+} f(x) = \lim_{x \to a^-} f(x) = L$.*

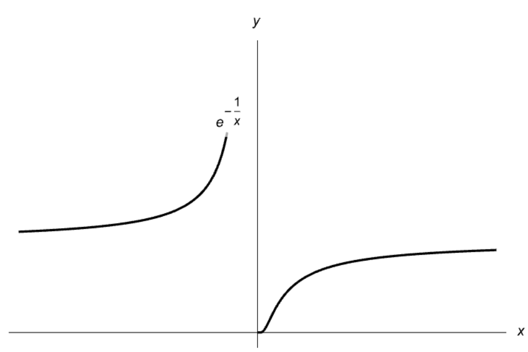
Proof. The conditional implication is trivial, we prove the converse. Suppose $\lim_{x \to a^+} f(x) = \lim_{x \to a^-} f(x) = L$. Then we can find $\delta, \gamma > 0$ such that given $\epsilon > 0$, $0 < x - a < \delta \Rightarrow |f(x) - L| < \epsilon$ and $0 < a - x < \gamma \Rightarrow |f(x) - L| < \epsilon$. If we set $\alpha = \min\{\delta, \gamma\}$, then $0 < |x - a| < \alpha \Rightarrow |f(x) - L| < \epsilon$. □

Example 14. *Consider the function sign* $: \mathbb{R} - \{0\} \to \mathbb{R}$ *given by*

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

Then $\lim_{x \to 0^-} sign(x) = -1$ but $\lim_{x \to 0^+} sign(x) = 1$, so $\lim_{x \to 0} sign(x)$ doesn't exist.

Example 15. *Consider the function $f(x) : \mathbb{R} \to \mathbb{R}$ given by $f(x) = e^{-\frac{1}{x}}$.*



Then $\lim_{x\to 0^+} f(x) = 0$ *but* $\lim_{x\to 0^-} f(x)$ *doesn't exist.*

Recall that a function is *increasing* if $x < y \Rightarrow f(x) < f(y)$, *nondecreasing* if $x \le y \Rightarrow f(x) \le f(y)$. We define *decreasing, nonincreasing* in a similar way. Finally we say a function is *monotone* if satisfies any of the above conditions.

Theorem 16. *Let* $X \subseteq \mathbb{R}$ *and* $f : X \to \mathbb{R}$ *a bounded monotone function. Given* $a \in X'_+, b \in X'_-$, *the one sided limits* $\lim_{x\to a^+} f(x)$ *and* $\lim_{x\to b^-} f(x)$ *exist.*

Proof. Without loss of generality, suppose $f(x)$ increasing. We prove $\lim_{x\to a^+} f(x)$ exist, the other limit is analogous. Set $L := \inf\{f(x); x > a\}$. We claim $\lim_{x\to a^+} f(x) = L$. Indeed, given $\epsilon > 0$ the number $\epsilon + L$ is not a lower bound, hence we can find $\delta > 0$ such that $L \le f(a+\delta) < L + \epsilon$. Since $f(x)$ is increasing, it follows that $a < x < a + \delta \Rightarrow L \le f(x) < L + \epsilon$, as required. □

Let $X \subseteq \mathbb{R}$ be a set unbounded from above. Given $f : X \to \mathbb{R}$ we write

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

if there is a number $L \in \mathbb{R}$ such that

$$\forall \epsilon > 0, \exists M > 0,\ M < x \Rightarrow |f(x) - L| < \epsilon.$$

The limit $\lim_{x\to -\infty} f(x)$ is defined analogously. Notice that both infinite limits are, in a way, one sided limits. In particular, the limit of a sequence x_n is an infinite limit when we consider the sequence as a function $x : \mathbb{N} \to \mathbb{R}$, i.e. $\lim x_n = \lim_{n\to +\infty} x(n)$.

Example 17. *We have* $\lim_{x\to -\infty} \frac{1}{n} = \lim_{x\to +\infty} \frac{1}{n} = 0$. *Also,* $\lim_{x\to -\infty} e^x = 0$ *but* $\lim_{x\to +\infty} e^x$ *doesn't exist.*

Let $X \subseteq \mathbb{R}$, $f : X \to \mathbb{R}$ and $a \in X'$. We write

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

if $\forall M > 0, \exists \delta > 0,\ 0 < |x - a| < \epsilon \Rightarrow f(x) > M$.

The definition of $\lim_{x\to a} f(x) = -\infty$, $\lim_{x\to \pm\infty} f(x) = \pm\infty$, and $\lim_{x\to a^\pm} f(x) = \pm\infty$ can be given *mutatis mutandis.*

Example 18. *With the definitions above we have, for example,* $\lim_{x\to+\infty} e^x = +\infty$, $\lim_{x\to-\infty} x^2 = +\infty$, $\lim_{x\to 2^-} \left(\frac{1}{x-2}\right) = -\infty$, $\lim_{x\to 2^+} \left(\frac{1}{x-2}\right) = +\infty$.

The theorem below can be proven using the same arguments we used to prove their finite counterpart, so the proof will be ommitted.

Theorem 19. *(Properties of infinite limits) Let* $X \subseteq \mathbb{R}$, $f : X \to \mathbb{R}$ *and* $a \in X'$.

- *(Uniqueness) If* $\lim_{x\to a} f(x) = +\infty$ *then it's impossible to have* $\lim_{x\to a} f(x) = L$ *for* $L \in \mathbb{R}$ *or* $L = -\infty$.

- *(Restriction) If* $\lim_{x\to a} f(x) = +\infty$, *then for every* $Y \subseteq X$, *if we set* $g(x) = f_{|_Y}(x)$, *we still have* $\lim_{x\to a} g(x) = +\infty$.

- *(Unboundedness) If* $\lim_{x\to a} f(x) = +\infty$, *then* $f(x)$ *is not bounded in any neighborhood of* $a \in X$.

- *(Monotonicity) If* $f(x) \leq h(x)$ *and* $\lim_{x\to a} f(x) = +\infty$, *then* $\lim_{x\to a} h(x) = +\infty$.

- *(Preservation of the sign) If* $\lim_{x\to a} f(x) = L$ *and* $\lim_{x\to a} h(x) = +\infty$, *then* $\exists \delta > 0$ *such that* $0 < |x - a| < \delta \Rightarrow f(x) < h(x)$.

- *(Equivalent definition)* $\lim_{x\to a} f(x) = +\infty$ *if and only if for every sequence* $x_n \in X - \{a\}$ *with* $\lim x_n = a$, *we have* $\lim_{n\to\infty} f(x_n) = +\infty$.

3 Limit superior and inferior of functions

Let $X \subseteq \mathbb{R}$, $f : X \to \mathbb{R}$ and $a \in X'$. We say f is *bounded in a neighborhood* of a, if there is $k, \delta > 0$ such that

$$0 < |x - a| < \delta \Rightarrow |f(x)| \leq k$$

A number $c \in \mathbb{R}$ is an *adherent value* of f at a if there exists a sequence $x_n \in X$ such that $\lim x_n = a$ and $\lim f(x_n) = c$. In particular, if a function has a limit $\lim_{x\to a} f(x) = L$, then L is the only adherent value.

Given $a \in X'$ and $\delta > 0$, we denote by I_δ the $\delta-$neighborhood around a given by $I_\delta = X - \{a\} \cap (a - \delta, a + \delta)$.

Theorem 20. *A number $c \in \mathbb{R}$ is an adherent value of f at a if and only if for every $\delta > 0$ we have $c \in \overline{f(I_\delta)}$.*

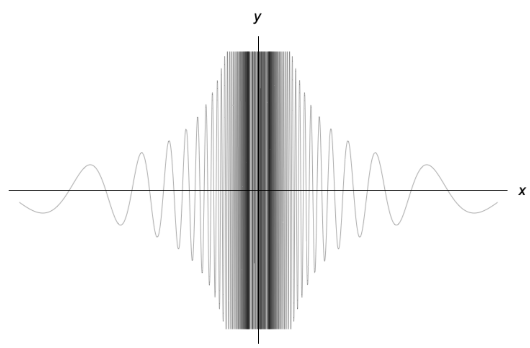
Proof. Suppose $c \in \mathbb{R}$ is an adherent value. Then $a = \lim x_n$ and $c = \lim f(x_n)$. Since $I_\delta \ni a$, $x_n \in I_\delta$ for n sufficiently large, so $f(x_n) \in \overline{f(I_\delta)}$. Conversely, suppose $c \in \overline{f(I_\delta)}$ for every $\delta > 0$. We can take δ of the form $\delta = \frac{1}{n}$, for $n \in \mathbb{N}$, to obtain a sequence $x_n \in I_{\frac{1}{n}}$, such that $|f(x_n) - c| < \frac{1}{n}$. We conclude that $\lim x_n = a$ and $\lim f(x_n) = c$. $\square$

Let's denote the set of all adherent values at a of a function f by $AV(f, a)$.

Corollary 21. $AV(f,a) = \bigcap_{\delta > 0} \overline{f(I_\delta)}$

Corollary 22. *$AV(f,a)$ is a closed set. If f is bounded in a neighborhood of a, then $AV(f,a)$ is compact and nonempty.*

Example 23. *Let $f(x) = \frac{\sin\left(\frac{1}{x}\right)}{x}$, whose graph is shown below.*



Every $c \in \mathbb{R}$ is an adherent value of f at 0, that is, $AV(f, 0) = \mathbb{R}$. Indeed, given any $c \in \mathbb{R}$ and an open intervals $(c - \epsilon, c + \epsilon) \ni c$ and $I_\delta := (-\delta, \delta) \ni 0$, we claim $(c - \epsilon, c + \epsilon) \cap f(I_\delta) \neq \emptyset$, or equivalently, $c - \epsilon < \frac{\sin\left(\frac{1}{a}\right)}{a} < c + \epsilon$ for some $a \in (-\delta, \delta)$, which is easily true by the periodicity of $\sin(x)$ and the behavior of $\frac{1}{x}$.

Example 24. *Let $f(x) = \frac{1}{x}$, then $AV(f, 0) = \emptyset$.*

According to corollary 22, if f is bounded in a neighborhood of a, the set $AV(f, a) \neq \emptyset$ is compact, hence has a maximum and minimum value.

We call the maximum value of $AV(f,a)$ the *limit superior* of f at a and denote it by

$$\limsup_{x\to a} f(x).$$

Similarly, the minimum value of $AV(f,a)$ is called the *limit inferior* of f at a and denote it by

$$\liminf_{x\to a} f(x).$$

We use the convention that when f is not bounded around a, we write $\limsup_{x\to a} f(x) = +\infty$ and $\liminf_{x\to a} f(x) = -\infty$.

Example 25. *Let $f(x) = \sin\left(\frac{1}{x}\right)$ then $AV(f,0) = [-1,1]$. Indeed, for a fixed $a \in [-1,1]$ consider $x_n = (a+2\pi n)^{-1}$, then $f(x_n) = a$. Therefore, $\liminf_{x\to a} f(x) = -1$ and $\limsup_{x\to a} f(x) = 1$.*

Theorem 26. *Let f be a bounded function in a neighborhood of a. Then given $\epsilon > 0$, there exists $\delta > 0$ such that*

$$0 < |x-a| < \delta \Rightarrow \liminf_{x\to a} f(x) - \epsilon < f(x) < \limsup_{x\to a} f(x) + \epsilon.$$

Corollary 27. $\lim_{x\to a} f(x) = L$ *if and only if f has only one adherent value at a, namely L itself.*

4 Continuity

Intuitively, a continuous function is a function whose graph has no gaps or holes. More precisely, let $f : X \to \mathbb{R}$ be a real valued function and $a \in X$. We say f is *continuous* at a if

$$\forall \epsilon > 0, \exists \delta > 0;\ |x-a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon$$

If f is continuous for every $a \in X$ we simply say f is continuous.

Notice that if $a \in X$ is an isolated point then any function $f : X \to \mathbb{R}$ is continuous at a. In particular, if $X' = \emptyset$ then any function $f : X \to \mathbb{R}$ is continuous.

Example 28. *Any function $f : \mathbb{Z} \to \mathbb{R}$ is continuous, since $\mathbb{Z}' = \emptyset$.*

Theorem 29. *If $a \in X'$, then f is continuous at a if and only if $\lim_{x\to a} f(x) = f(a)$.*

Proof. Self-evident. □

By using the already proven properties of limits we conclude:

Theorem 30. *If $f : X \to \mathbb{R}$ is continuous then for any $Y \subseteq X$ the restriction $f_{|Y}$ is also continuous. Conversely, if $Y = I \cap X$ for some open interval I containing a point $a \in X$, then if $f_{|Y}$ is continuous at a, f is also continuous at a.*

In other words, theorem 30 says that continuity is a *local property*. More precisely, if f coincides with a continuous function in a neighborhood of $a \in X$, then f itself is continuous at a.

Corollary 31. *If f is continuous at $a \in X$, then f is bounded in a neighborhood of a.*

Corollary 32. *If f, g are continuous at $a \in X$ and $f(a) < g(a)$, then $f(x) < g(x)$ in a neighborhood of a.*

Corollary 33. *If f is continuous at $a \in X$ and $f(a) < k$ ($f(a) > k$), for some $k \in \mathbb{R}$, then $f(x) < k$ ($f(x) > k$) in a neighborhood of a.*

Using the alternate definition of limit we can prove:

Theorem 34. *f is continuous at $a \in X$ if and only if for every sequence $x_n \to a$, we have $f(x_n) \to f(a)$.*

Theorem 35. *f, g are continuous at $a \in X$, them $f + g$,$f - g$, and $f \cdot g$ are also continuous at a. If $g(a) \neq 0$ then f/g is also continuous at a. Moreover, the composition of continuous function is also continuous.*

Example 36. *The function $f(x) = x$ is clearly continuous, hence its self-product x^n is also continuous, and so is any polynomial $p(x) = a_n x^n + \ldots + a_1 x + a_0$. A rational function $p(x)/q(x)$ is continuous at points where $q(x) \neq 0$.*

Example 37. *The function $f(x) = |x|$ is continuous on the open interval $(0, +\infty)$ since it is constant there, for the same reason it's also continuous in $(-\infty, 0)$. Finally, it's continuous at 0, since $\lim\limits_{x \to 0^-} |x| = \lim\limits_{x \to 0^+} |x| = 0$. On the other hand, the function defined by $g(x) = \frac{x}{|x|}$, if $x \neq 0$, and $g(0) = 1$, is not continuous at the origin since $\lim\limits_{x \to 0^-} g(x) = -1 \neq \lim\limits_{x \to 0^+} g(x) = 1$.*

Theorem 38. *Suppose $X \subseteq A \cup B$, where $A, B \subseteq \mathbb{R}$ are closed sets. If the function $f : X \to \mathbb{R}$ satisfies $f_{|X\cap A}$ is continuous and $f_{|X\cap B}$ is continuous, then f itself is continuous.*

Proof. Let $a \in X$ and $\epsilon > 0$ be given. Suppose first $a \in A \cap B$. Then there are $\delta, \gamma > 0$ such that $\forall x \in X \cap A, |x - a| < \delta \Rightarrow |f(x) - f(a)| < \epsilon$ and $\forall x \in X \cap B, |x - a| < \gamma \Rightarrow |f(x) - f(a)| < \epsilon$. Set $\alpha = \min\{\delta, \gamma\}$, then $\forall x \in X, |x - a| < \alpha \Rightarrow |f(x) - f(a)| < \epsilon$, which implies f is continuous at a.

Now suppose $a \in A$ but $a \notin B$. There exists $\delta > 0$, such that $\forall x \in X \cap A, |x-a| < \delta \Rightarrow |f(x)-f(a)| < \epsilon$. Since B is closed, $\overline{B} = B$, and we can find $\gamma > 0$ such that $|x-a| < \gamma \Rightarrow x \notin B$. As before, if we set $\alpha = \min\{\delta, \gamma\}$, then $\forall x \in X, |x - a| < \alpha \Rightarrow |f(x) - f(a)| < \epsilon$, as desired. The case $a \notin A$ but $a \in B$ can be proven analogously. □

Corollary 39. *Suppose $X = A \cup B$, where $A, B \subseteq \mathbb{R}$ are closed sets. If the restrictions $f_{|A}, f_{|B}$ of a function $f : X \to \mathbb{R}$ are continuous, then f itself is continuous.*

We can generalize the result above if we take the cover $A \cup B$ to be open. In fact, a stronger result is valid. (The proof follows directly from theorem 30 and will be omitted.)

Theorem 40. *(Sheaf property) Let $X \subseteq \bigcup_{\lambda\in L} A_\lambda$ be an open cover of X. If the restrictions $f_{|X\cap A_\lambda}$ of a function $f : X \to \mathbb{R}$ are continuous, then f itself is continuous*

Corollary 41. *Suppose $X = \bigcup_{\lambda\in L} A_\lambda$, where each A_λ is open. If the restrictions $f_{|A_\lambda}$ of a function $f : X \to \mathbb{R}$ are continuous, then f itself is continuous*

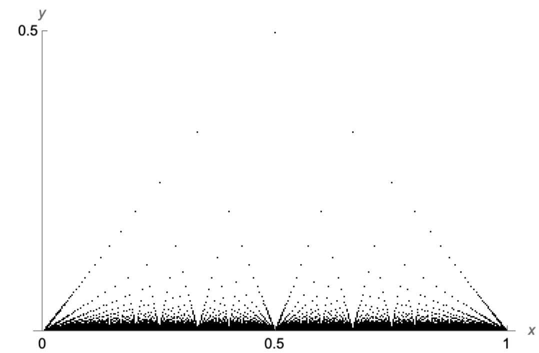
Example 42. *Consider again $f(x) = \frac{x}{|x|}$ but this time with domain $X = (-\infty, 0) \cup (0, +\infty)$. Then f is continuous by the corollary above.*

Let $f : X \to \mathbb{R}$ be a real valued function and $a \in X$. If f is not continuous at a, we say it is *discontinuous* at a.

Example 43. *(Thomae's function)The function $f : \mathbb{R} \to \mathbb{R}$ given by:*

$$f(x) = \begin{cases} \frac{1}{q}, & \text{if } x \in \mathbb{Q} \text{ and } x = \frac{p}{q}, p \in \mathbb{Z}, q \in \mathbb{N}, \gcd(p, q) = 1 \\ 0, & \text{if } x \in \mathbb{R} - \mathbb{Q} \end{cases}$$

The graph of $f(x)$ on the interval $(0, 1)$ is shown below.



Notice that $f(x)$ is periodic, since $f(x+1) = f(x)$. We claim that f is discontinuous at any $a \in \mathbb{Q}$. Indeed, we can find a sequence, say $x_n = a + \frac{\sqrt{2}}{n}$, of irrational numbers, with $x_n \to a$ but $f(x_n) \to 0$, since $f(a) \neq 0$ in this case, f can't be continuous at a.

Surprisingly enough, f is continuous at every $a \notin \mathbb{Q}$. Equivalently, we must have $\lim\limits_{x \to a} f(x) = 0$. Since f is periodic, it's enough to prove the continuity for $a \in (0,1) \cap (\mathbb{R} - \mathbb{Q})$.

Suppose $\epsilon > 0$ is given. Using the Archimedean property of $\mathbb{R}$, there is $n \in \mathbb{N}$ such that $\frac{1}{n} < \epsilon$. Decompose $(0,1)$ into k subintervals of length $\frac{1}{k}$, for $k = 1, 2, \dots, n$. Then 'a' will be in one of these intervals, for each k, say $a \in (\frac{m_k}{k}, \frac{m_k+1}{k})$. Let $\delta_k = \min\left\{|a - \frac{m_k}{k}|, |a - \frac{m_k+1}{k}|\right\}$, the minimum distance between a and the endpoints of $(\frac{m_k}{k}, \frac{m_k+1}{k})$, and define $\delta := \min\limits_{1 \leq k \leq n} \delta_k$.

Given $x \in (a - \delta, a + \delta)$ if $x \notin \mathbb{Q}$ then $f(x) = 0 < \epsilon$. Otherwise, $x = \frac{p}{q}$ and by minimality of δ, we must have $q > n$, hence $f(x) = \frac{1}{q} < \frac{1}{n} < \epsilon$ and we conclude that $\lim\limits_{x \to a} f(x) = f(a) = 0$.

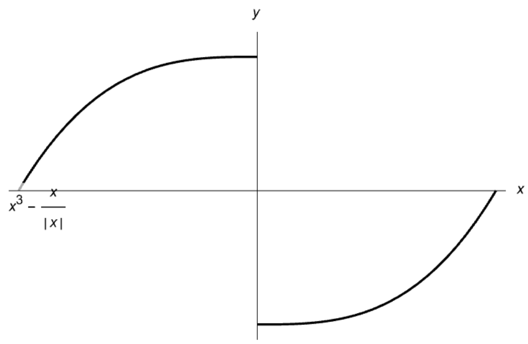
It's impossible to have a function which is discontinuous at every irrational number, see the exercises.

Example 44. *If $f : \mathbb{R} \to \mathbb{R}$ is given by:*

$$f(x) = \begin{cases} 1, & \text{if } x \in \mathbb{Q} \\ 0, & \text{if } x \in \mathbb{R} - \mathbb{Q} \end{cases}$$

Then f is discontinuous at every $a \in \mathbb{R}$, since the limit $\lim\limits_{x \to a} f(x)$ doesn't exist.

Example 45. *Consider $f : \mathbb{R} \to \mathbb{R}$ given by $f(0) = 1$ and $f(x) = x^3 - \frac{x}{|x|}$ if $x \neq 0$. Then f is discontinuous at 0 only.*



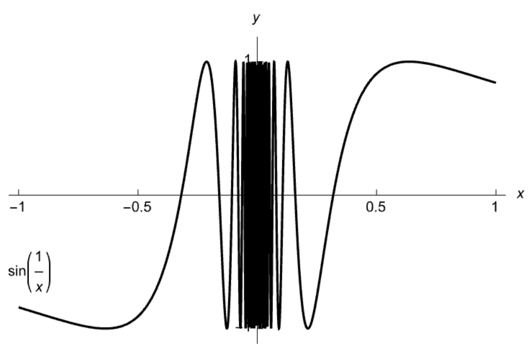
Example 46. *Let* K *be the Cantor set. Consider the function* $f : [0, 1] \to \mathbb{R}$ *given by*

$$f(x) = \begin{cases} 0, & \text{if } x \in K \\ 1, & \text{if } x \notin K \end{cases}$$

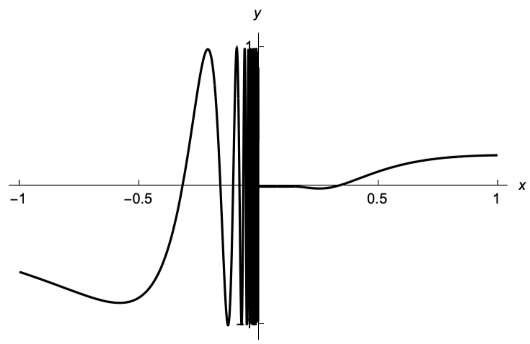
Then f *is discontinuous at every point* $a \in K$ *and continuous at the open set* K^c. *Indeed,* f *is constant, hence continuous, at every* $a \in K^c$.

Suppose now $a \in K$. *Since every point of* K *is an accumulation point, it's possible to find a sequence* $x_n \notin K$ *such that* $x_n \to a$, *hence* $f(x_n) \to 1 \neq 0$, *so* f *is discontinuous at* a.

Example 47. *The function* $f(0) = a$ *and* $f(x) = \sin \frac{1}{x}$ *if* $x \neq 0$ *is discontinuous at* 0, *regardless of* $a \in \mathbb{R}$, *since* $\lim_{x \to 0} f(x)$ *doesn't exist.*



Example 48. *The function* $f(0) = 0$ *and* $f(x) = \frac{\sin \frac{1}{x}}{1 + e^{\frac{1}{x}}}$ *if* $x \neq 0$ *is discontinuous at* 0, *since* $\lim_{x \to 0^-} f(x)$ *doesn't exist. In this case,* $\lim_{x \to 0^+} f(x) = 0$ *however.*



Example 49. *The function* $f(0) = 0$ *and* $f(x) = \frac{1}{1+e^{\frac{1}{x}}}$ *if* $x \neq 0$ *is discontinuous at* 0*, since* $\lim_{x \to 0^-} f(x) = 1$ *but* $\lim_{x \to 0^+} f(x) = 0$*.*

Let $f : X \to \mathbb{R}$, $a \in X$ and suppose f is discontinuous at a. Then we say $a \in X$ is a *jump discontinuity*, if both one sided limits $\lim_{x \to a^+} f(x)$ and $\lim_{x \to a^-} f(x)$ exists but are different. If at least one of the one sided limits doesn't exist, then we say $a \in X$ is an *essential discontinuity.*

Theorem 50. *A monotone function* $f : X \to \mathbb{R}$ *can't have essential discontinuities.*

Proof. Suppose f nondecreasing and $a \in X$. If $x + \delta \in X$ then f is bounded in $[x, x + \delta] \cap X$. The result then follows from theorem 16. $\square$

Theorem 51. *Let* $f : X \to \mathbb{R}$ *be a function having only jump discontinuities. Then the set of discontinuities of* f *is countable.*

Proof. Define the jump function $j(x) : X \to \mathbb{R}$ of f by:

$$j(a) = \begin{cases} 0, \text{ if } a \text{ is isolated.} \\ |f(a) - \lim_{x \to a^+} f(x)|, \text{ if } a \in X'_+ \text{ only.} \\ |f(a) - \lim_{x \to a^-} f(x)|, \text{ if } a \in X'_- \text{ only.} \\ \max\{|f(a) - \lim_{x \to a^+} f(x)|, |f(a) - \lim_{x \to a^-} f(x)|\}, \text{ if } a \in X'_+ \cap X'_-. \end{cases}$$

Intuitively, $j(x)$ measures the length of the 'jump' of $f(x)$. Consider the set

$$C_n := \{x \in X; j(x) \geq \frac{1}{n}\}.$$

The set of discontinuities of $f(x)$ is the set $\bigcup_{n=1} C_n$, hence if we can prove that each C_n is countable then we're done. We claim that for each $n \in \mathbb{N}$, the set C_n has only isolated points, hence it's countable (see corollary 32).

Let $a \in C_n$ and suppose $a \in X'_+$. By using the definition of one sided limit, if we set $L := \lim_{x \to a^+} f(x)$ we can find $\delta > 0$ such that $0 < x - a < \delta \Rightarrow |f(x) - L| < \frac{1}{4n} \Rightarrow L - \frac{1}{4n} < f(x) < L + \frac{1}{4n}$, hence if $x \in (a, a + \delta)$ then $j(x) \leq \frac{1}{2n}$, which is to say $(a, a+\delta) \cap C_n = \emptyset$. If $a \notin X'_+$, we can just choose $\delta > 0$ such that $(a, a+\delta) \cap X = \emptyset$. In any case, we can find $\delta > 0$ such that $(a, a+\delta) \cap C_n = \emptyset$. A similar argument implies we can find $\gamma > 0$ such that $(a - \gamma, a) \cap C_n = \emptyset$. We conclude that $a \in C_n$ is isolated. □

Corollary 52. *The set of discontinuities of a monotone function f is countable.*

5 Continuous functions defined on intervals

The next result highlights the fact that continuous functions can't have gaps, in other words, if two numbers $a \neq b$ are in the range, then $[a, b]$ is also in the range.

Theorem 53. *(Intermediate Value Theorem) Let $f : [a, b] \to \mathbb{R}$ be a continuous function and $d \in \mathbb{R}$ be a number such that $f(a) < d < f(b)$. Then there is $c \in [a, b]$ such that $d = f(c)$.*

Proof. Define $X = \{x \in [a, b]; f(x) < d\}$. This set is nonempty because $f(a) < f(d)$, and due to the continuity of $f(x)$, X doesn't have a maximum element. Set $c = \sup X$, then $c \notin X$. However, since c is an adherent value, there is a sequence $x_n \to c$, which implies $f(c) \leq d$. We conclude that $f(c) = d$. □

Corollary 54. *Let $f : I \to \mathbb{R}$ be a continuous function, where I is an interval (not necessarily bounded). If $a, b \in I$ and $f(a) < d < f(b)$, then there exists $c \in I$ such that $f(c) = d$.*

Corollary 55. *Let $f : I \to \mathbb{R}$ be a continuous function, where I is an interval. Then $f(I)$ is an interval.*

Proof. If we set $c = \inf f(x)$ and $d = \sup f(x)$ then $f(I)$ is an interval with endpoints c and d *(not necessarily bounded, nor open/closed).* □

Example 56. *Let $f : I \to \mathbb{R}$ be a continuous function such that $f(I) \subseteq Y$, where Y has empty interior. Then f is constant. Indeed, it follows by 55 that $f(I)$ is an interval, so it must be of the form $[c, c]$, otherwise, $f(I)$ would have an interior point. In particular, every continuous function $f : I \to \mathbb{Z}$ is constant.*

Example 57. *Every polynomial $p(x) = a_{2n-1}x^{2n-1} + \ldots + a_0$ of odd degree has at least one real root. Indeed, in this case $p(x)$ is a continuous function defined on the interval $(-\infty, +\infty)$, so its image is an interval. Since $\lim_{x\to\pm\infty} p(x) = \pm\infty$, that interval has to be $(-\infty, +\infty)$, hence $p(x)$ is surjective.*

A function $f : X \to Y$ is a *homeomorphism*, if f is a continuous bijection having a continuous inverse f^{-1}.

Theorem 58. *Let $f : I \to \mathbb{R}$ be a continuous injective function defined on a interval I. Then f is monotone, and if we set $J = f(I)$, then $f : I \to J$ is a homeomorphism.*

Proof. It's enough to prove the result for $I = [a, b]$. Suppose $f(a) < f(b)$, we claim f is increasing. Suppose not, that is, we can find $c, d \in [a, b]$ such that $c < d$ but $f(c) > f(d)$. Either $f(a) < f(d)$ or $f(a) > f(d)$. If $f(a) < f(d) < f(c)$, by theorem 53, we can find $p \in (a, c)$ such that $f(p) = f(d)$, a contradiction by the injectivity of f. For the same reason we can't have $f(d) < f(a) < f(b)$. Hence, f has to be increasing.

Using corollary 55, we see that J is an interval, hence $f^{-1} : J \to I$ is an increasing function (since f is) whose image is an interval. Suppose f^{-1} is not continuous at a point $y \in J$, say $M := \lim_{x\to y^+} f^{-1}(x) \neq L := \lim_{x\to y^-} f^{-1}(x)$. Then $f^{-1}(c) \in (L, M)$ and $(L, M) \cap I = \{f^{-1}(c)\}$, which implies I has an isolated point, a contradiction. □

Theorem 59. *Let $f : X \to \mathbb{R}$ be a continuous function. If X is compact then $f(X)$ is compact.*

Proof. We claim $f(X)$ is sequentially compact, which is equivalent to compactness by theorem 44. Let $y_n = f(x_n)$ be a sequence in $f(X)$, we claim it has a converging subsequence. By the compactness of X, there is a converging subsequence $x_{n_k} \to x \in X$. If we set $y_{n_k} = f(x_{n_k})$, then $y_{n_k} \to f(x)$, since f is continuous. □

Corollary 60. *(Weierstrass Extreme Value Theorem) Let $X \subseteq \mathbb{R}$ be compact and $f : X \to \mathbb{R}$ be a continuous function. Then f achieves its maximum and minimum value, that is to say, there are $a, b \in X$ such that $f(a) \leq f(x) \leq f(b)$ for every $x \in X$.*

Theorem 61. *Let $X \subseteq \mathbb{R}$ be compact and $f : X \to \mathbb{R}$ be a continuous injective function. If we set $Y := f(X)$, then $f : X \to Y$ is a homeomorphism.*

Proof. Let $y \in Y$, we claim f^{-1} is continuous at $y = f(x)$. Suppose $y_n = f(x_n)$ is a sequence of points in Y such that $y_n \to y = f(x)$, we claim $x_n \to x$. It's enough to prove that any converging subsequence of x_n converges to x. Let x_{n_k} be a converging subsequence, say $x_{n_k} \to a \in X$. Then $y_{n_k} \to f(a)$, but since y_{n_k} is a subsequence of y_n, it also converges to $f(x)$, by the injectivity of f we deduce that $a = x$. □

We say a function $f : X \to \mathbb{R}$ is uniformly continuous if

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x, y \in X, |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$$

It follows that every uniformly continuous function is continuous. The converse is false, as the example below illustrates.

Example 62. *The function $f(x) = \frac{1}{x}$ is continuous on $(0, +\infty)$ but is not uniformly continuous. Indeed, given $\epsilon, \delta > 0$, take a point $0 < x < \min\{\delta, \frac{1}{3\epsilon}\}$ and $y = x + \frac{\delta}{2}$. Then $|x - y| < \delta$ but*

$$|f(x) - f(y)| = \left| \frac{1}{x} - \frac{1}{x + \frac{\delta}{2}} \right| = \left| \frac{\delta}{x(2x + \delta)} \right| > \left| \frac{\delta}{3\delta x} \right| > \epsilon.$$

Example 63. *Linear functions $f(x) = mx + b$ are continuous. Indeed, given $\epsilon > 0$ just take $\delta = \frac{\epsilon}{|m|}$, so that $|x - y| < \delta \Rightarrow |f(x) - f(y)| = |m(x - y)| \leq |m| \frac{\epsilon}{|m|} = \epsilon$.*

Example 64. *A function $f : X \to \mathbb{R}$ is called Lipschitz if there exists a constant $C > 0$ such that $|f(x) - f(y)| \leq C|x - y|$. Any Lipschitz function is obviously uniformly continuous. For example, linear functions $f(x) = mx + b$ are Lipschitz, and if X is bounded, $f(x) = x^n$ is Lipschitz.*

Theorem 65. *If $f : X \to \mathbb{R}$ is uniformly continuous and x_n is a Cauchy sequence then $f(x_n)$ is also Cauchy.*

Corollary 66. *If* $f : X \to \mathbb{R}$ *is uniformly continuous and* $a \in X'$ *then* $\lim_{x \to a} f(x)$ *exists.*

Example 67. *The functions* $f(x) = \sin \frac{1}{x}$ *and* $g(x) = \frac{1}{x}$ *can't be uniformly continuous because the limit when when* x *approaches* 0 *doesn't exist.*

Theorem 68. *Let* $X \subseteq \mathbb{R}$ *be compact and* $f : X \to \mathbb{R}$ *continuous then* f *is uniformly continuous.*

VI Derivatives

1 Definition and first properties

Let $X \subseteq \mathbb{R}$, $a \in X \cap X'$, and $f : X \to \mathbb{R}$ be a real valued function. We say f is *differentiable* at $a \in X$ if the following limit exists:

$$f'(a) := \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \tag{1}$$

The number $f'(a)$ is called the derivative of f at a. If f is differentiable at every $a \in X$, we simply say f is differentiable (in X).

Intuitively speaking, for $x \neq a$, the number $\frac{f(x)-f(a)}{x-a}$ is the slope of the secant line connecting the points $(x, f(x))$ and $(a, f(a))$, hence when $x \to a$, this number becomes the slope of the tangent line.

Similarly to one-sided limits, we can define *one-sided derivatives*, $f'_+(a) := \lim\limits_{x \to a^+} \frac{f(x)-f(a)}{x-a}$, if $a \in X \cap X'_+$, and $f'_-(a) := \lim\limits_{x \to a^-} \frac{f(x)-f(a)}{x-a}$ if $X \cap X'_-$. We can easily see that $f'(a)$ exists for some $a \in X \cap X'_+ \cap X'_-$ if and only if $f'_+(a)$ and $f'_-(a)$ exist and $f'_-(a) = f'_+(a)$. In particular, a function is not differentiable if its graph has sharp corners, since this implies $f'_-(a) \neq f'_+(a)$ at the corner.

If we set $h := x - a$ in equation 1, then we can see that $f'(a)$ can be equivalently defined by

$$f'(a) := \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}. \tag{2}$$

Sometimes the latter definition is more convenient for computational purposes.

If $a \in X'_+$ but $a \notin X'_-$, and $f'_+(a)$ exists, we can set $f'(a) := f'_+(a)$ and consider f to be differentiable at a. A similar convention holds for $a \in X'_-$. According to this convention, the function $f : [a, b) \to [a, b)$, given by $f(x) = x$, is differentiable.

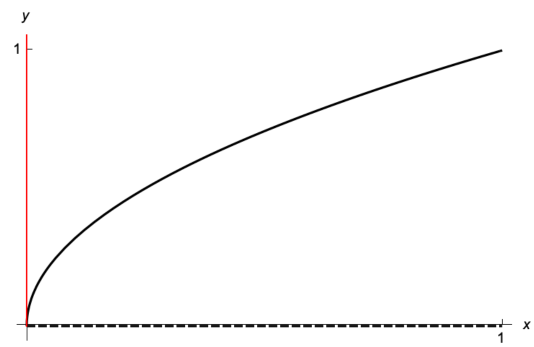
Example 1. *Let $f : \mathbb{R} \to \mathbb{R}$ be linear, $f(x) = mx + b$. Then $f'(x) = m$. In particular, if $m = 0$ and $f(x) = b$ is constant, then $f'(x) = 0$.*

Example 2. *Consider $f(x) = |x|$. Using the definition of one-sided derivatives we obtain $f'_+(0) = 1$ and $f'_-(0) = -1$. Therefore, f is not differentiable at 0. On the other hand, we easily see that $f'(x) = 1$, if $x > 0$, and $f'(x) = -1$, if $x < 0$.*

Example 3. *Let $f : [0, +\infty) \to \mathbb{R}$ be defined by $f(x) = \sqrt{x}$. Using equation 2, for $x > 0$, we obtain:*

$$f'(x) = \lim_{h\to 0} \frac{\sqrt{x+h} - \sqrt{x}}{h} = \lim_{h\to 0} \frac{h}{h(\sqrt{x+h} + \sqrt{x})} = \frac{1}{2x}$$

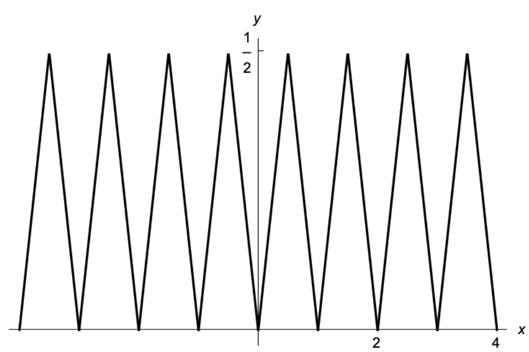
On the other hand, at $x = 0$ the quotient $\frac{\sqrt{h}}{h} = \frac{1}{\sqrt{h}} \to +\infty$ as $h \to 0^+$, hence $f'(0)$ doesn't exits. Intuitively, this is clear since the tangent line being a vertical line has 'infinite' slope.



Example 4. *(Sawtooth function)Let $f : \mathbb{R} \to \mathbb{R}$ be defined by*

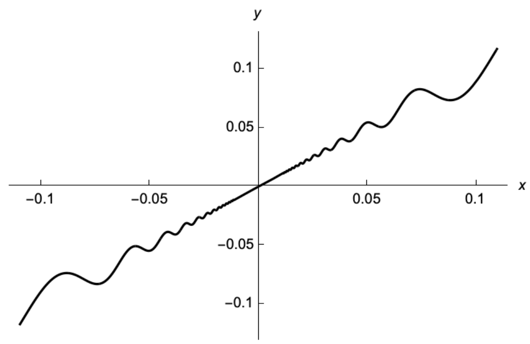
$$f(x) = \inf\{|x - n|; n \in \mathbb{Z}\}$$

.

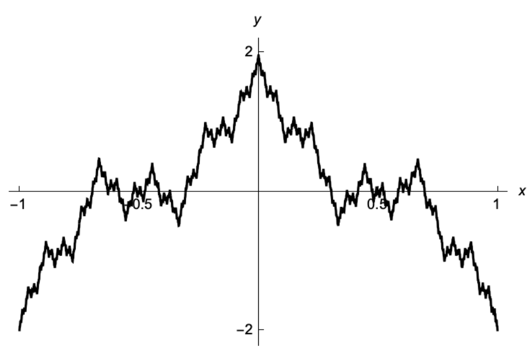


Notice that the graph of f has sharp corners at every $n, \frac{n}{2}$, for $n \in \mathbb{Z}$, hence it's not differentiable at those points. Otherwise, the function is differentiable with $f'(x) = \pm 1$, depending whether or not the fractional part of $f(x)$ is less than 0.5.

Example 5. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(0) = 0$ and $f(x) = x + 2x^2 \sin(1/x)$, if $x \neq 0$. Despite this seemly complicated definition, this function is indeed differentiable everywhere and $f'(x) = 1-2\cos(1/x)+4x\sin(1/x)$*



Example 6. *(Weierstrass function) Given $0 < a < 1$ and $b \in \mathbb{N}$, such that $ab > 1 + \frac{3}{2}\pi$. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = \sum_{n=1}^{\infty} a^n \cos(b^n \pi x)$. The figure below is the graph of $f(x)$. It is an example of a continuous function that is nowhere differentiable.*



Moreover, the graph of $f(x)$ is self-similar if we zoom in, in the sense, that if we restrict the the domain of $f(x)$ to $[-\frac{1}{n}, \frac{1}{n}]$ and take n bigger and bigger, the shape of the graph doesn't change. We will prove these claims later, when we discuss series of functions.

Theorem 7. *A real valued function $f : X \to \mathbb{R}$ is differentiable at $a \in X$ if and only if there is number $C \in \mathbb{R}$ and a real valued function $r(x)$, such that if $a + h \in X$:*

$$f(a+h) = f(a) + Ch + r(h), \tag{3}$$

and $r(x)$ satisfies $\lim_{h\to 0} \frac{r(h)}{h} = 0$. Moreover, $C = f'(a)$.

Proof. The implication is clear. We prove the converse. Suppose that there is $C \in \mathbb{R}$ satisfying (3). Then

$$f(a+h) - f(a) - r(h) = Ch \tag{4}$$

Dividing both sides by h and taking the limit when $h \to 0$ we obtain

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = C \in \mathbb{R},$$

as required. □

The theorem above says that f is differentiable at a if and only if in a neighborhood of a, f can be approximated by the linear function $p(x) = f'(a)x + f(a)$ with error $r(x)$ that goes to zero faster than $g(x) = x$. We will see soon that the more derivatives f has, the better we can make this approximation using a polynomial $p(x)$ whose degree is equal to the number of derivatives of f.

If $f : X \to \mathbb{R}$ differentiable at $a \in X \cap X'$, we define the *differential at* a, denoted by $df_a : \mathbb{R} \to \mathbb{R}$, as the linear transformation given by

$$df_a(h) = f'(a)h. \tag{5}$$

In this notation, equation 3 becomes

$$f(a+h) = f(a) + df_a(h) + r(h). \tag{6}$$

Theorem 8. *If the $f : X \to \mathbb{R}$ is differentiable at $a \in X$ then f is continuous at $a \in X$.*

Proof. Indeed, we have

$$\begin{aligned} \lim_{x \to a}[f(x) - f(a)] = \lim_{x \to a}\left[\frac{f(x) - f(a)}{x - a}(x - a)\right] &= \lim_{x \to a}\left[\frac{f(x) - f(a)}{x - a}\right] \cdot \lim_{x \to a}(x - a) \\ &= f'(a) \cdot 0 = 0. \end{aligned} \tag{7}$$

$\therefore$ f is continuous at a. □

The theorem below follows directly from the definition of derivative and the properties of limits we have already proved.

Theorem 9. *(Properties of derivatives) If $f, g : X \to \mathbb{R}$ are differentiable at $a \in X \cap X'$ then $f \pm g$, $f \cdot g$, f/g (if $g'(a) \neq 0$) are also differentiable at a. Moreover,*

$$\begin{aligned} (f \pm g)'(a) &= f'(a) \pm g'(a) \\ (f \cdot g)'(a) &= f'(a) \cdot g(a) + f(a) \cdot g'(a) \\ \left(\frac{f}{g}\right)'(a) &= \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}. \end{aligned} \tag{8}$$

Theorem 10. *(The Chain Rule) Let $f : X \to \mathbb{R}$ and $g : Y \to \mathbb{R}$ be real valued functions, such that $f(X) \subseteq Y$. If f is differentiable at $a \in X$, and g is differentiable at $b := f(a)$, then $g \circ f : X \to \mathbb{R}$ is differentiable at a, moreover $(g \circ f)'(a) = g'(b)f'(a)$.*

Proof. By hypothesis, we have

$$\begin{aligned} (g \circ f)(a + h) = g[f(a + h)] &= g[f(a) + f'(a)h + r(h)] \\ &= g[f(a)] + g'[f(a)][f'(a)h + r(h)] + s(f'(a)h + r(h)) \\ &= g(b) + g'(b)[f'(a)h] + g'(b)[r(h)] + s(f(a + h) - f(a)). \end{aligned}$$

Since

$$\lim_{h\to 0} \frac{g'(b)[r(h)] + s(f(a + h) - f(a))}{h} = g'(b) \lim_{h\to 0} \frac{r(h)}{h} + \lim_{h\to 0} \frac{s(f(a + h) - f(a))}{h} = 0$$

The proof is complete by theorem 7. □

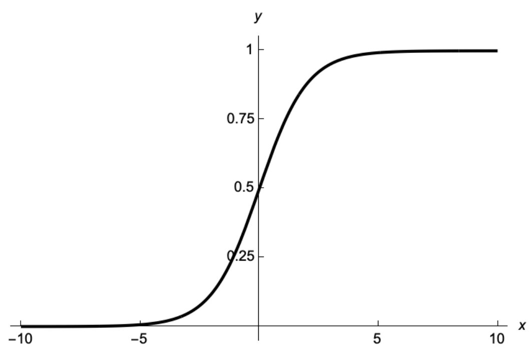
Corollary 11. *Let $f : X \to Y \subseteq \mathbb{R}$ be a bijective real valued functions. If f is differentiable at $a \in X$, and $f^{-1} : Y \to X$ is continuous at $b := f(a)$, then f^{-1} is differentiable at b if and only if $f'(a) \neq 0$, moreover, if that's the case, then $(f^{-1})'(b) = \frac{1}{f'(a)}$.*

Proof. If f^{-1} is differentiable at b, we can apply the Chain rule to $1 = (f^{-1} \circ f)'(a) = (f^{-1})'(b)f'(a)$. Conversely, suppose $f'(a) \neq 0$, set $g(y) := f^{-1}(y)$. Then

$$\lim_{y\to b} \frac{g(y) - g(b)}{y - b} = \lim_{y\to b} \frac{g(y) - a}{f[g(y)] - f(a)} = \lim_{y\to b} \left(\frac{f[g(y)] - f(a)}{g(y) - a}\right)^{-1} = \frac{1}{f'(a)} \tag{9}$$

$\therefore\ g'(b) = \frac{1}{f'(a)}$ and the theorem is proved. □

Example 12. *(The Sigmoid function) Consider the function* $f : \mathbb{R} \to \mathbb{R}$ *given by* $f(x) = \frac{1}{1+e^{-x}}$*, whose graph is shown below.*



Using the chain rule, we have that

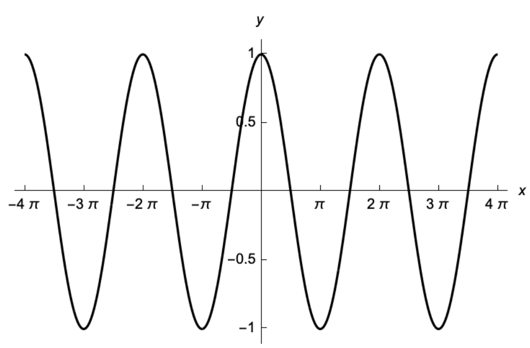
$$f'(x) = -\frac{1}{(1+e^{-x})^2}(-e^{-x}) = \frac{e^{-x}}{(1+e^{-x})^2}$$

2 Maximum and minimum points

The derivative of $f : X \to \mathbb{R}$ at point $a \in X$ tells us crucial information about the behavior of the function in a neighborhood of a.

Let $f : X \to \mathbb{R}$ be a real valued function and $a \in X$. We say f has a *local maximum* at a if there exists $\delta > 0$, such that $x \in (a - \delta, a + \delta) \Rightarrow f(x) \leq f(a)$. If the strict inequality $f(x) < f(a)$ is true, then a is called *strict local maximum.* Similar definitions are given to *local minimum* and *strict local minimum.*

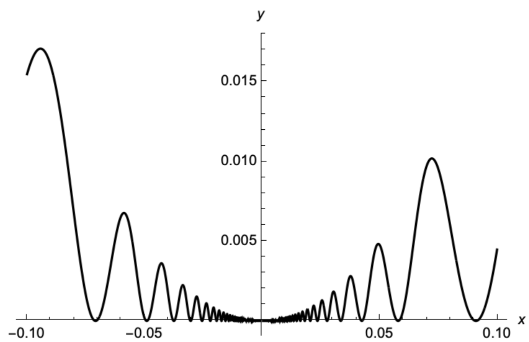
Example 13. *The function* $\cos : \mathbb{R} \to \mathbb{R}$ *has (strict) local maxima at points of the form* $a = 2\pi n$*,* $n \in \mathbb{Z}$*.*



Similarly, $\cos x$ *has (strict) local minima at points of the form* $(2n-1)\pi$, $n \in \mathbb{Z}$.

Example 14. *The constant function given by* $f(x) = C$ *has (non-strict) local maxima and minima at every point of its domain.*

Example 15. *Consider the function* $f : \mathbb{R} \to \mathbb{R}$ *given by* $f(0) = 0$ *and* $f(x) = x^2(1 + \sin \frac{1}{x})$, *whose graph is shown below.*



By definition, $f(x) \geq 0$, $\forall x \in \mathbb{R}$. *Moreover, any neighborhood of* 0 *contains points whose image is* 0. *Hence, the point* 0 *is a (non-strict) local minimum.*

Theorem 16. *Let* $f : X \to \mathbb{R}$ *be differentiable from the right at* $a \in X \cap X'_+$, *i.e.* $f'_+(a)$ *exists. If* $f'_+(a) > 0$ *then we can find* $\delta > 0$ *such that* $x \in (a, a+\delta) \Rightarrow f(x) > f(a)$. *Similarly, if* $f'_+(a) < 0$ *then* $\exists \delta > 0 : x \in (a, a+\delta) \Rightarrow f(x) < f(a)$.

Proof. Follows directly from Corollary 6. □

A similar result is valid in the case $f'_-(a) > 0$ or $f'_-(a) < 0$.

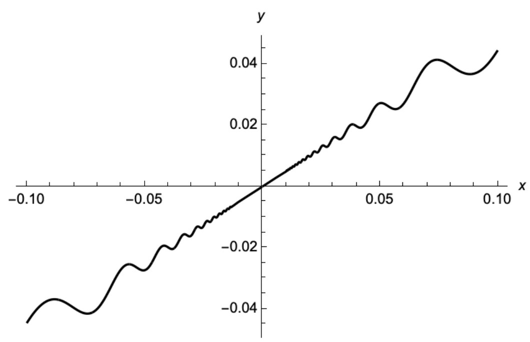
Corollary 17. *Let* $f : X \to \mathbb{R}$ *be differentiable at* $a \in X \cap X'_+ \cap X'_-$. *If* $f'(a) > 0$ *then we can find* $\delta > 0$ *such that for all* $x, y \in X$, *we have* $a - \delta < x < a < y < a + \delta \Rightarrow f(x) < f(a) < f(y)$.

Notice that the corollary above *is not* saying that f is locally increasing.

Corollary 18. *Let* $f : X \to \mathbb{R}$ *be differentiable at* $a \in X \cap X'_+ \cap X'_-$. *If* f *has a local maximum or minimum at* $a \in X$ *then* $f'(a) = 0$.

Example 19. *The converse of Corollary 18 is false. The function* $f(x) = x^3$ *and* $a = 0$ *gives a counter-example.*

Example 20. *Consider the continuous function* $f(x) = x^2 \sin \frac{1}{x} + \frac{x}{2}$ *if* $x \neq 0$ *and* $f(0) = 0$.



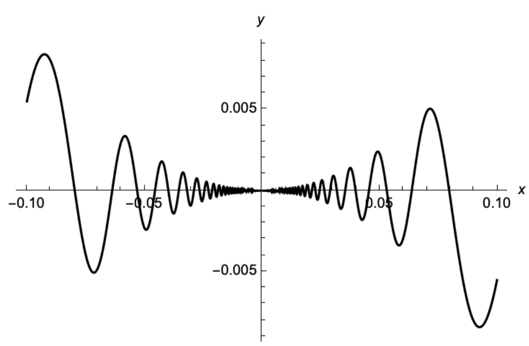
We have $f'(0) = \frac{1}{2} > 0$, *but* f *is not increasing in any neighborhood* I *of* 0. *Indeed,* $f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x} + \frac{1}{2}$, *so we can pick* $x \in I$ *sufficiently small such that* $\sin \frac{1}{x} = 0$ *and* $\cos \frac{1}{x} = 1$, *for this* $x \in I$ *we have* $f'(x) = -\frac{1}{2} < 0$, *so* f *can't be increasing in* I.

3 Derivative as a function

Let $f : I \to \mathbb{R}$ be a differentiable function defined on a interval I. We associate to f its derivative function $f' : I \to \mathbb{R}$, whose value at each $x \in I$ is $f'(x)$.

When f' is continuous, we say f is *continuously differentiable*. The set of all continuously differentiable functions on a interval I is denoted by $C^1(I)$. In case $I = (-\infty, +\infty)$, we simply write $f \in C^1$ and say f is of class C^1.

Example 21. *The function defined by* $f(x) = x^2 \sin \frac{1}{x}$ *if* $x \neq 0$ *and* $f(0) = 0$ *is differentiable but* $f \notin C^1$.



At $x = 0$ we have $f'(0) = 0$. However, $f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$ and $\lim_{x \to 0} f'(x)$ doesn't exists. Therefore, f' is not continuous at 0.

If $f : I \to \mathbb{R}$ is of class C^1, then we can apply the Intermediate Value Theorem to f' to conclude that: Given $a, b \in I$ such that $f'(a) < y < f'(b)$ for some $y \in \mathbb{R}$, then there exists $c \in I$ such that $y = f'(c)$.

The following theorem strengthens the above by removing the continuity assumption of f'.

Theorem 22. *(Darboux's theorem) Let $f : [a, b] \to \mathbb{R}$ be differentiable. If $f'(a) < y < f'(b)$, then there exists $c \in I$ such that $y = f'(c)$.*

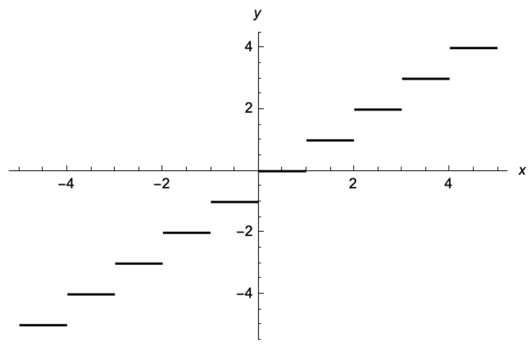
Proof. It suffices to prove the result when $y = 0$ and then consider $g(x) = f(x) - yx$. From the fact that $f'(a) < 0 < f'(b)$, we know that $f(x) < f(a)$ in a neighborhood of a, and $f(x) < f(b)$ in a neighborhood of b. That implies that f achieves its minimum (see corollary 60) at a point $c \in (a, b)$, by 18 we must have $f'(c) = 0$. □

Example 23. *The corollary above says that the Dirichlet function $f(x) = 1$, if $x \in \mathbb{Q} \cap [0, 1]$, $f(x) = 0$ if $x \in (\mathbb{R} - \mathbb{Q}) \cap [0, 1]$ can't be the derivative of a function defined on $[0, 1]$.*

Corollary 24. *Let $f : I \to \mathbb{R}$ be a differentiable function on an interval I. Then f' doesn't have jump discontinuities.*

Proof. We claim that given a point $a \in I$, if the one sided limits $\lim_{x \to a^+} f'(x)$, $\lim_{x \to a^-} f'(x)$ exist, then $f'(x)$ is continuous at a. Suppose $R = \lim_{x \to a^+} f'(x)$ exists but $R \neq f'(a)$, say $R > f'(a)$. Take $y \in \mathbb{R}$ such that $f'(a) < y < R$. Then there exists $\delta > 0$ such that $x \in (a, a + \delta) \Rightarrow f'(x) > y$. In particular, $f'(a) < R < f'(a + \frac{\delta}{2})$ but there is no $c \in (a, a + \frac{\delta}{2})$ such that $f'(c) = R$, a contradiction. Using a similar argument, we conclude the equivalent result if $\lim_{x \to a^-} f'(x)$ exists. □

Example 25. *The corollary above says that the floor function $f(x) = \lfloor x \rfloor$, can't be the derivative of a function defined on $\mathbb{R}$.*



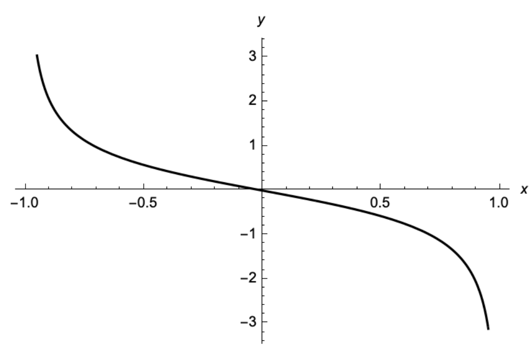
Theorem 26. *(Rolle) Let $f : [a, b] \to \mathbb{R}$ be continuous satisfying $f(a) = f(b)$. If f is differentiable on (a, b) then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. If f is constant then $f'(x) = 0$, so suppose f not constant. Since f is continuous on $[a, b]$, it achieves its maximum and minimum in $[a, b]$. Since $f(a) = f(b)$, the maximum/minimum can't be at an endpoint, otherwise the function would be constant. Hence, the function has at least one maximum or minimum in the interior (a, b), at that point the derivative must be zero by Corollary 18. □

Notice that we didn't use $f'(a)$ or $f'(b)$ in the proof, hence the requirement that f be differentiable in (a, b) and not in $[a, b]$.

Example 27. *The absolute value function $f(x) = |x|$ when defined on $[-1, 1]$ is continuous and satisfies $f(-1) = f(1)$, but there is no point $c \in [-1, 1]$ such that $f'(c) = 0$. This is not a counter-example to Theorem 26, because f is not differentiable at $0 \in [-1, 1]$.*

Example 28. *The function $f(x) = \sqrt{1 - x^2}$ is continuous on $[0, 1]$ but it's differentiable only in $(0, 1)$, since it's derivative $f'(x) = -\frac{x}{\sqrt{1-x^2}}$ is discontinuous at ± 1, as the picture below suggests.*

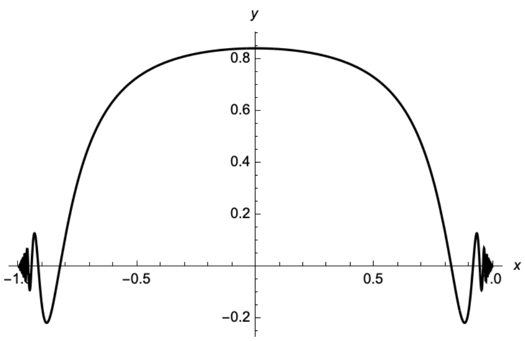


Still, Rolle's theorem guarantees the existence of a point $c \in [0,1]$ *with* $f'(c) = 0$. *Indeed,* $c = 0$ *in this case.*

Example 29. *(The headphone function) The function* $f : [-1,1] \to \mathbb{R}$ *defined by*

$$f(x) = \begin{cases} 0, & \text{if } |x| = 1 \\ (1-x^2)\sin\frac{1}{1-x^2}, & \text{if } |x| \neq 1 \end{cases}$$

is another example of function continuous on $[-1,1]$ *but differentiable only in* $(-1,1)$.



Theorem 30. *(Lagrange's Mean Value Theorem) Let* $f : [a,b] \to \mathbb{R}$ *be continuous. If* f *is differentiable on* (a,b) *then there exists* $c \in (a,b)$ *such that*

$$f'(c) = \frac{f(b)-f(a)}{b-a}.$$

Proof. Set $g(x) = \frac{f(b)-f(a)}{b-a}(x-a) + f(a)$. Then g satisfies $g(a) = f(a)$ and $g(b) = f(b)$. If we set $h(x) = f(x) - g(x)$, the function h satisfies $h(a) = h(b)$, hence by Rolle's theorem $h'(c) = 0$ for some $c \in (a,b)$. The result follows. □

Corollary 31. *Let* $f : [a,b] \to \mathbb{R}$ *be continuous such that* $f'(x) = 0$ *for every* $x \in (a,b)$. *Then* f *is constant.*

Corollary 32. *Let* $f, g : [a,b] \to \mathbb{R}$ *be continuous functions such that* $f'(x) = g'(x)$ *for every* $x \in (a,b)$. *Then* $f(x) = g(x) + C$, *for some constant* $c \in \mathbb{R}$.

Corollary 33. *Any function* $f : I \to \mathbb{R}$ *defined on a interval such that* $x \in I \Rightarrow |f'(x)| \leq C$ *for some* $C \in \mathbb{R}$, *is Lipschitz.*

Corollary 34. *Let $f : I \to \mathbb{R}$ be differentiable in an interval I. Then $f'(x) \geq 0$ if and only if f is nondecreasing in I. In case $f'(x) > 0$, then f is increasing. Equivalent statements are true if $f'(x) \leq 0$ and f nonincreasing.*

Proof. Suppose $f'(x) \geq 0$ and $x, y \in I$ such that $x \leq y$. By the Mean Value Theorem, $f(y) - f(x) = f'(c)(y - x) \geq 0$, and we conclude that $f(x) \leq f(y)$. Conversely, if f is nondecreasing then for every $x \in I$ such that $x + h \in I$, we have that the ratio $\frac{f(x+h)-f(x)}{h}$ is always nonnegative, hence its limit when $h \to 0$ is also nonnegative. The same argument *mutatis mutandis* applies in the strict inequality. □

Example 35. *As a nice application of the Mean Value theorem we show that* $\lim(\sqrt{n+1} - \sqrt{n}) = 0$. *Consider the function $f : [n, n+1] \to \mathbb{R}$ given by $f(x) = \sqrt{x}$. Using the Mean Value Theorem we can find $c \in (n, n+1)$ such that*

$$f'(c) = \frac{\sqrt{n+1} - \sqrt{n}}{(n+1) - n},$$

or equivalently

$$\sqrt{n+1} - \sqrt{n} = \frac{1}{2c} \leq \frac{1}{2n}.$$

Using the Squeeze theorem we conclude that $\lim(\sqrt{n+1} - \sqrt{n}) = 0$.

4 Taylor's Theorem

Let $f : I \to \mathbb{R}$ be a real valued function defined on an interval I. The n-th derivative of f, if exists, is defined inductively by setting $f''(x) = (f')'(x)$ and $f^{(n)}(x) = (f^{(n-1)})'(x)$ for $n \in \mathbb{N}$. By convention, we set $f^0(x) = f(x)$.

We say that f is of class C^k in I, denoted by $f \in C^k(I)$, if $f^{(k)}$ exists and is continuous in I. When $I = \mathbb{R}$, we simply write $f \in C^k$. Recall that $f \in C^0$, means f is continuous, so the definition makes sense even if k is zero.

In case $f \in C^k(I)$ for every $k \in \mathbb{N}$, we say that f is *smooth* and write $f \in C^\infty(I)$. Equivalently, a function f is smooth if $f^{(n)}$ exists for every $n \in \mathbb{N}$.

The following example generalizes example 21.

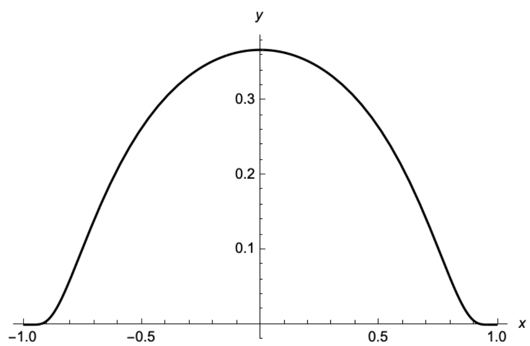
Example 36. *The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = |x|x$ is C^1 but it's not C^2. Indeed, we can easily check that its derivative is given by*

$$f'(x) = \begin{cases} 2x, \ x \geq 0 \\ -2x, \ x < 0 \end{cases}$$

which is continuous everywhere. Whereas, f'' has a jump discontinuity at zero, so $f \notin C^2$. More generally, the function $g(x) = |x|x^n$ is in C^n but $g \notin C^{n+1}$.

Example 37. *(Standard Mollifier) Consider the function defined by:*

$$f(x) = \begin{cases} e^{-\frac{1}{1-|x|^2}}, \ |x| < 1 \\ 0, \ |x| \geq 1 \end{cases}$$



We can easily see that $f \in C^\infty$ and the set where $f \neq 0$ is bounded, hence has compact closure. This type of function and its higher dimensional generalization are extensively used in the field of differential equations.

Example 38. *Since $\sin' x = \cos x$ and $\cos' x = -\sin x$, we deduce that $\sin x, \cos x \in C^\infty$. Similarly, $e^x, \log x$ and any polynomial are examples of smooth functions.*

Let $f : I \to \mathbb{R}$ be a real valued function defined on an interval $I \subseteq \mathbb{R}$ having derivatives up to order n at $a \in I$, i.e. $f^{(n)}(a)$ exists. The polynomial $p(x)$ defined by

$$p(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \dots \frac{f^{(n)}(a)}{n!}(x-a)^n \qquad (10)$$

is called the *Taylor polynomial* of order n of f at a.

Equivalently, the n-th order Taylor polynomial of f at a is the unique polynomial $p(x)$ of degree n, such that $f^{(k)}(a) = p^{(k)}(a)$ for $k = 1, 2, \dots, n$.

Theorem 39. *(Taylor's Theorem) Let $f : I \to \mathbb{R}$ be a real valued function having derivatives up to order n at $a \in I$, and $p(x)$ be the n-th order Taylor polynomial at a. Then the function $r : I \to \mathbb{R}$, defined by $r(x) = f(x) - p(x)$, i.e.*

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + r(x),$$

satisfies $\lim\limits_{x \to a} \frac{r(x)}{(x-a)^n} = 0$.

Proof. Recall that the case $n = 1$ was proved in theorem 7. Suppose $n = 2$, we use the Mean Value Theorem to obtain c between x and a such that:

$$\frac{r(x)}{(x-a)^2} = \frac{r(x) - r(a)}{(x-a)^2} = \frac{r'(c)(x-a)}{(x-a)^2} = \frac{r'(c)}{x-a} = \frac{[r'(c) - r'(a)](c-a)}{(c-a)(x-a)}$$

$\therefore \lim\limits_{x \to a} \frac{r(x)}{(x-a)^2} = 0$, since $r^{(2)}(a) = 0$ and $\left|\frac{c-a}{x-a}\right| \leq 1$. Using the same argument, we can prove the result for any value n. □

Corollary 40. *(L'Hôpital's rule) Let $f, g : I \to \mathbb{R}$ be real valued functions having derivatives up to order n at $a \in I$, such that $f^{(k)}(a) = g^{(k)}(a) = 0$, for $k = 0, 1, 2, \ldots, n-1$, but $f^{(n)}(a)$ and $g^{(n)}(a)$ are non-zero. Then*

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f^{(n)}(a)}{g^{(n)}(a)}.$$

Proof. By Taylor's formula and the hypothesis of the corollary, we have:

$$\frac{f(x)}{g(x)} = \frac{\frac{f^{(n)}(a)}{n!} + \frac{r(x)}{(x-a)^n}}{\frac{g^{(n)}(a)}{n!} + \frac{s(x)}{(x-a)^n}},$$

for some $r(x), s(x)$, satisfying $\frac{r(x)}{(x-a)^n} \to 0$ and $\frac{s(x)}{(x-a)^n} \to 0$, when $x \to a$. The corollary is then immediate. □

Corollary 41. *Let $f : I \to \mathbb{R}$ be real valued function having derivative up to order n at $a \in \operatorname{int}(I)$, such that $f^{(k)}(a) = 0$, for $k = 1, 2, \ldots, n-1$, but $f^{(n)}(a) \neq 0$. Then if n is odd, the point a is not a local maximum or minimum, and if n is even, two outcomes are possible: $f^{(n)}(a) > 0$ implies the point a is a strict local minimum; $f^{(n)}(a) < 0$ implies the point a is a strict local maximum.*

Proof. Notice that in this case Taylor's formula can be written as

$$f(a+h) - f(a) = h^n \left[\frac{f^{(n)}(a)}{n!} + \frac{r(a+h)}{h^n} \right]$$

for $h \in \mathbb{R}$ such that $a+h \in I$. Since $\frac{r(a+h)}{h^n} \to 0$ when $h \to 0$, for h sufficiently small, say $0 < |h| < \delta$, the expression in the square brackets has the same sign as $f^{(n)}(a)$. Hence, if n is odd, we can always find $h_1, h_2 \in I$ such that $f(a+h_1) - f(a) > 0$ and $f(a+h_2) - f(a) < 0$, so a can't be a local maximum or minimum.

Now, suppose n is even. Then if $f^{(n)}(a) > 0$, the above discussion implies $f(a+h) - f(a) > 0$ for $0 < |h| < \delta$, hence a is a local minimum. Similarly, if $f^{(n)}(a) < 0$ we must have $f(a+h) - f(a) < 0$, and a is a local maximum. □

We can enhance Taylor's Theorem if we require f to be of Class C^n and having the $f^{(n+1)}$ derivative, instead of just having the f^n derivative, which is not necessarily continuous.

Theorem 42. *(Taylor's Theorem with Lagrange Remainder) Let $f : [a, b] \to \mathbb{R}$ be a real valued function such that $f \in C^n$ and $f^{(n+1)}(x)$ exists in (a, b). Then there exists $c \in (a, b)$ such that*

$$f(b) = f(a) + f'(a)(b-a) + \ldots + \frac{f^{(n)}(a)}{n!}(b-a)^n + \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}.$$

Proof. Define $g : [a, b] \to \mathbb{R}$ by

$$g(x) = f(b) - f(x) - f'(x)(b-x) + \ldots + \frac{f^{(n)}(x)}{n!}(b-x)^n + \frac{C}{(n+1)!}(b-x)^{n+1},$$

where C is the unique number such that $g(a) = 0$.

The function g is continuous on $[a, b]$, differentiable in (a, b), and satisfies $g(a) = g(b)$. Therefore, by Rolle's Theorem, there exists $c \in (a, b)$ such that $g'(c) = 0$. On the other hand, a quick computation gives:

$$g'(x) = \frac{C - f^{(n+1)}(x)}{n!}(b-x)^n,$$

We conclude that $C = f^{(n+1)}(c)$, and the theorem becomes the statement $g(a) = 0$. □

Let $f : I \to \mathbb{R}$ be a smooth function, i.e. $f \in C^\infty$, and $a \in I^\circ$. Using Taylor's Theorem with Lagrange remainder, for each $n \in \mathbb{N}$ we have:

$$f(x) = f(a) + f'(a)(x-a) + \ldots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + r_n(x), \qquad (11)$$

where $r_n(x) = \frac{f^{(n)}(c)}{n!}(x-a)^n$ and c is between x and a. It is then natural to ask what happens when we let $n \to +\infty$ in (11).

The series $f(a)+f'(a)(x-a)+\ldots+\frac{f^{(n)}(a)}{n!}(x-a)^n+\ldots = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$, is called the *Taylor Series* of f at $a \in I$. Notice that it's not entirely clear that the Taylor Series of f at a has to coincide with $f(x)$, in fact, it's possible for the Taylor Series to diverge and even if it converges, it could converge to a number other than $f(x)$.

A function $f : I \to \mathbb{R}$ is called *Analytic* if for every $a \in I$, there exists $\delta > 0$ such that

$$|x-a| < \delta \Rightarrow f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

In other words, a function is analytic if it coincides with its Taylor series in a neighborhood of every point of its domain. Notice that it follows from (11) that a function is analytic if and only if for every $x \in I$, we have $\lim_{n\to\infty} r_n(x) = 0$.

Example 43. *Any polynomial $p(x)$ is clearly analytic, since $p^{(n)}(x)$ vanishes for sufficiently large $n \in \mathbb{N}$.*

Example 44. *The exponential function $f(x) = e^x$ is perhaps one of the most famous analytic functions. We use Taylor's theorem (with $a = 0$), to obtain:*

$$e^x = 1 + x + \frac{x^2}{2} + \ldots + \frac{x^n}{n!} + e^{c_n}\frac{x^n}{n!}$$

with $|c_n| < |x|$. Since $\lim \frac{x^n}{n!} = 0$, the Taylor series for e^x at 0 converges to e^x. Moreover, notice that $e^{x+a} = e^x e^a$, hence the Taylor series for e^x converges at any point $a \in \mathbb{R}$, and e^x is analytic.

Example 45. *Let $x \in \mathbb{R}$, then*

$$1 + x + x^2 + \ldots + x^{n-1} + \frac{x^n}{1-x} = \frac{1}{1-x}.$$

Consider the function $f : (0,1) \to \mathbb{R}$ given by $f(x) = \frac{1}{1-x}$. Then using Taylor's Theorem we obtain $r_n(x) = \frac{x^n}{1-x}$ in this case, so $\lim_{n\to\infty} r_n(x) = 0$, which implies $f(x) = \sum_{n=0}^{\infty} x^n$. Hence, $f(x)$ agrees with its Taylor Series at 0.

Example 46. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = \cos x$. Using Taylor's theorem around the origin (with $a = 0$), we can write*

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + r_{2n+1}(x)$$

where $r_n(x) = [\cos x^{(n)}](c)\frac{x^n}{n!}$. Notice that

$$0 \leq |r_n(x)| \leq \frac{|x|^{2n+1}}{(2n+1)!},$$

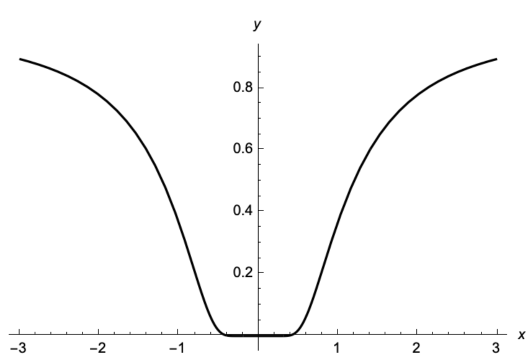
and recall that by example 51, $\lim_{n\to\infty} \frac{|x|^{2n+1}}{(2n+1)!} = 0$. We conclude that $\lim_{n\to\infty} r_n(x) = 0$ and it follows that

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + \ldots .$$

Hence, the Taylor series of $\cos x$ at 0 converges to $\cos x$ at every point $x \in \mathbb{R}$. The same argument can be applied if if the Taylor series is not centered at zero ($a \neq 0$). In conclusion, the function $\cos x$ is analytic.

Example 47. *Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by*

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, \ x \neq 0 \\ 0, \ x = 0 \end{cases}$$



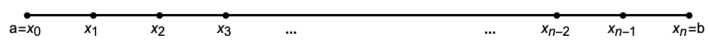
Using the fact that $\lim_{x\to 0} \frac{e^{-\frac{1}{x^2}}}{x^n} = 0$ *for any* $n \geq 0$*, we can see that* $f^{(n)}(0) = 0$*, and the function* f *is smooth. However, the Taylor series at* 0 *is identically zero, since* $\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = 0$*. In particular, since* $x \neq 0 \Rightarrow f(x) \neq 0$*, it's impossible for* $f(x)$ *to be analytic on* $\mathbb{R}$*.*

VII Integrals

1 Definition and first properties

Let $[a,b] \subseteq \mathbb{R}$ be a closed interval. A *partition* of $[a,b]$ is a finite subset $P = \{x_0, x_1, \ldots, x_n\}$ of $[a,b]$, such that $x_0 = a$ and $x_n = b$.

By convention, the elements of a partition are written in increasing order, $P = \{a = x_0 < x_1 < x_2 < \ldots < x_n = b\}$.



Let P, Q be partitions of $[a,b]$. We say that the partition Q is a *refinement* of the partition P if $P \subseteq Q$. More precisely, Q is obtained from P by adding a finite number of points.

Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Set $m = \inf f$ and $M = \sup f$, then:

$$m \leq f(x) \leq M, \ \forall x \in [a,b].$$

If $P = \{x_0, x_1, \ldots, x_n\}$ is a partition of $[a,b]$, we denote

$$m_i := \inf\{f(x); x_{i-1} \leq x \leq x_i\} \text{ and } M_i := \sup\{f(x); x_{i-1} \leq x \leq x_i\},$$

and define the *oscillation* of f at $[x_{i-1}, x_i]$ by

$$\omega_i := M_i - m_i.$$

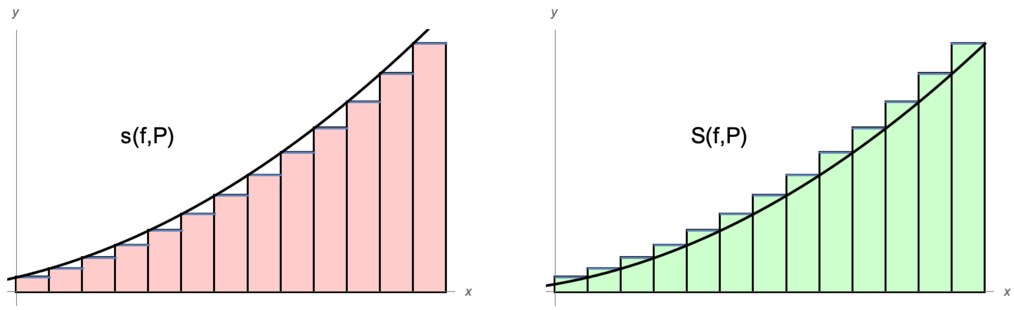
If f is continuous, the values m_i, M_i, ω_i are achieved by Weierstrass Extreme Value Theorem.

We define the *lower sum* of f with respect to P by

$$s(f;P) = m_1(x_1 - x_0) + \cdots + m_n(x_n - x_{n-1}) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}),$$

and likewise, the *upper sum* of f with respect to P by

$$S(f;P) = M_1(x_1 - x_0) + \cdots + M_n(x_n - x_{n-1}) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}).$$



By definition, we have

$$m(b-a) \leq s(f;P) \leq S(f;P) \leq M(b-a) \text{ and } S(f;P)-s(f;P) = \sum_{i=1}^{n} \omega_i(x_i - x_{i-1}).$$

When $f \geq 0$, the number $s(f;P)$ represents an approximation of the area under the graph of f using rectangles that are below the graph, whereas $S(f;P)$ represents an approximation using rectangles above the graph of f.

Let $\mathcal{P} = \{P; P \text{ is a partition of } [a,b]\}$ and $f : [a,b] \to \mathbb{R}$ be a bounded function. The *lower integral* and *upper integral* are defined respectively by:

$$\underline{\int_a^b} f(x)\mathrm{dx} := \sup_{P\in\mathcal{P}} s(f;P) \text{ and } \overline{\int_a^b} f(x)\mathrm{dx} := \inf_{P\in\mathcal{P}} S(f;P),$$

Theorem 1. *Let $P, Q \in \mathcal{P}$. Then*

$$P \subseteq Q \Rightarrow s(f;P) \leq s(f;Q) \text{ and } S(f;Q) \leq S(f;P)$$

Proof. It's enough to prove the result when $Q = P \cup \{a\}$. Suppose $P = \{x_0 < x_1 < \ldots < x_n\}$ and $x_{k-1} < a < x_k$ for some $k \leq n$. Define

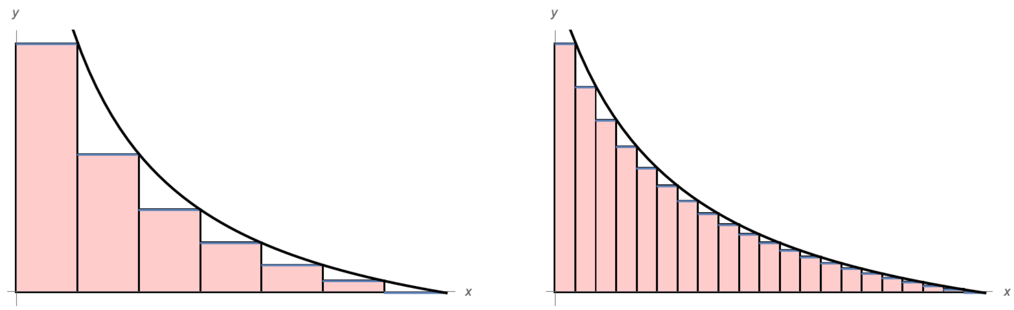
$$m' := \inf_{x\in[x_{k-1},a]} f(x) \text{ and } m'' := \inf_{x\in[a,x_k]} f(x).$$

Notice that m_k is less than or equal to m', m''. We have:

$$\begin{aligned} s(f;Q) - s(f;P) &= m'(a - x_{k-1}) + m''(x_k - a) - m_k(x_k - x_{k-1}) \\ &= (m'' - m_k)(x_k - a) + (m' - m_k)(a - x_{k-1}) \qquad (12) \\ &\geq 0 \end{aligned}$$

A similar argument shows that $S(f;Q) \leq S(f;P)$. $\square$

The figure below illustrates theorem 1 for a partition P and a refinement $Q \supseteq P$, when $f(x) = \frac{1}{x}$. The sum of the highlighted rectangles represent $s(f;P)$ and $s(f;Q)$ respectively. It's easy to see that $s(f;Q) \geq s(f;P)$.



Corollary 2. *For any partitions $P, Q \in \mathcal{P}$ we have*

$$s(f;P) \leq S(f;Q)$$

Proof. Apply Theorem 1 to P and $P \cup Q$ (Q and $P \cup Q$). □

Lemma 3. *Let $X, Y \subseteq \mathbb{R}$ be sets satisfing*

$$x \leq y, \ \forall x \in X, \forall y \in Y,$$

then $\sup X \leq \inf Y$. *Moreover, the equality* $\sup X = \inf Y$ *holds if and only if given $\epsilon > 0$, there are $x \in X, y \in Y$ such that $y - x < \epsilon$.*

Proof. By definition, every $y \in Y$ is an upper bound for X hence $\sup X \leq y$, for every $y \in Y$. On the other hand, $\sup X$ is a lower bound for Y, thus $\sup X \leq \inf Y$. Suppose $\sup X = \inf Y$ and $\epsilon > 0$ is given. Then $\sup X - \frac{\epsilon}{2}$ is not an upper bound, so $\exists x \in X$ such that $\sup X - \frac{\epsilon}{2} < x \leq \sup X$. Similarly, we can find $y \in Y$ such that $\inf Y \leq y < \inf Y + \frac{\epsilon}{2}$. Therefore, $y - x < \inf Y + \frac{\epsilon}{2} - \sup X + \frac{\epsilon}{2} = \epsilon$. Conversely, suppose $\sup X < \inf Y$. If we set $\epsilon = \inf Y - \sup X$, then $y - x \geq \epsilon$. □

Theorem 4. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function, say $m \leq f(x) \leq M$, then:*

$$m(b-a) \leq \underline{\int_a^b} f(x)\mathrm{dx} \leq \overline{\int_a^b} f(x)\mathrm{dx} \leq M(b-a)$$

Proof. The proof of the middle inequality follows directly from lemma 3. The other two inequalities are obvious. □

A bounded function $f:[a,b]\to\mathbb{R}$ is *(Riemann) integrable* if

$$\underline{\int_a^b} f(x)\mathrm{dx} = \overline{\int_a^b} f(x)\mathrm{dx},$$

and we denote this common value by $\int_a^b f(x)\mathrm{dx}$, or simply, by $\int_a^b f$.

Example 5. *The constant function $f:[a,b]\to\mathbb{R}$ given by $f(x)=C$ is clearly integrable since $s(f;P)=S(f;P)=C(b-a)$ for any partition P.*

Example 6. *The Dirichlet function $f:[0,1]\to\mathbb{R}$ given by $f(x)=1$ if $x\in\mathbb{Q}$, and 0 otherwise, is not integrable since $s(f;P)=0$ and $s(f;P)=b-a$ for any partition P.*

Theorem 7. *Let $f:[a,b]\to\mathbb{R}$ be a bounded function. The following are equivalent:*

(1) f is integrable,

(2) For every $\epsilon>0$, there are partitions P and Q of $[a,b]$ such that $S(f;Q)-s(f;P)<\epsilon$,

(3) For every $\epsilon>0$, there is a partition $R=\{x_0<x_1<\ldots<x_n\}$ of $[a,b]$ such that $S(f;R)-s(f;R)=\sum_{k=1}^{n}\omega_k(x_k-x_{k-1})<\epsilon$.

Proof. The fact that (1) ⇒ (2) and (3) ⇒ (1) follows directly from lemma 3. Suppose (2) is true and set $R=P\cup Q$, then

$$s(f;P)\le s(f;R)\le S(f;R)\le S(f;Q),$$

$\therefore\ S(f;R)-s(f;R)<\epsilon$, and (2) ⇒ (3). □

2 Properties of Integrals

Let $f:[a,b]\to\mathbb{R}$ be a bounded function. For simplicity, we adopt the following conventions:

$$\int_a^a f = 0 \text{ and } \int_b^a f = -\int_a^b f$$

Theorem 8. *Let $a < c < b$. Then $f : [a,b] \to \mathbb{R}$ is integrable if and only if $f_{|_{[a,c]}}$ and $f_{|_{[c,b]}}$ are integrable. In the affirmative case, we have*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Proof. Consider the sets

$$\begin{aligned} A &= \{s(f_{|_{[a,c]}}; P); P \text{ is a partition of } [a,c]\}, \\ B &= \{s(f_{|_{[c,b]}}; P); P \text{ is a partition of } [c,b]\}, \\ C &= \{s(f; P); P \text{ is a partition of } [a,b] \text{ and } c \in P\}. \end{aligned}$$

Notice that by Theorem 1, $\underline{\int_a^b} f = \sup C$. It follows that

$$\underline{\int_a^b} f = \sup(A+B) = \sup A + \sup B = \underline{\int_a^c} f + \underline{\int_c^b} f,$$

and similarly,

$$\overline{\int_a^b} f = \overline{\int_a^c} f + \overline{\int_c^b} f.$$

$$\therefore \quad \overline{\int_a^b} f - \underline{\int_a^b} f = \left(\overline{\int_a^c} f - \underline{\int_a^c} f \right) + \left(\overline{\int_c^b} f - \underline{\int_c^b} f \right).$$

We conclude that $\overline{\int_a^b} f = \underline{\int_a^b} f$ if and only if $\overline{\int_a^c} f = \underline{\int_a^c} f$ and $\overline{\int_c^b} f = \underline{\int_c^b} f$. □

Example 9. *(Step functions) Given a set $X \subseteq \mathbb{R}$, consider the function $\chi_A : \mathbb{R} \to \mathbb{R}$ defined by*

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}$$

χ_A is called the characteristic function of $A \subseteq \mathbb{R}$. Let $P = \{x_0 < x_1 < \ldots < x_n\}$ be a partition of $[a,b]$, and $c_1, c_2, \ldots, c_n \in \mathbb{R}$. The function $f(x) = \sum_{j=1}^{n} c_j \chi_{I_j}$, where $I_j = [x_{j-1}, x_j]$, is called a Step function. Since f is constant, in particular integrable, on I_j, theorem 8 guarantees that f is integrable.

Theorem 10. *Let $f, g : [a,b] \to \mathbb{R}$ be integrable. Then*

(1) $f+g$ *is integrable and* $\int_a^b (f+g) = \int_a^b f + \int_a^b g$,

(2) $f \cdot g$ *is integrable,*

(3) If $\exists k > 0$ *such that* $0 < k \leq |g(x)|$ *for every* $x \in [a,b]$, *then* f/g *is integrable,*

(4) If $f \leq g$ *then* $\int_a^b f \leq \int_a^b g$,

(5) $|f|$ *is integrable and* $\left|\int_a^b f\right| \leq \int_a^b |f|$.

Proof. Notice that for P, Q partitions of $[a,b]$ we have:

$$s(f;P)+s(g;Q) \leq s(f;P\cup Q)+s(g;P\cup Q) \leq s(f+g;P\cup Q) \leq \underline{\int_a^b}(f+g),$$

and hence:

$$\underline{\int_a^b} f + \underline{\int_a^b} g \leq \underline{\int_a^b}(f+g).$$

Similarly, we can show that $\overline{\int_a^b} f + \overline{\int_a^b} g \geq \overline{\int_a^b}(f+g)$. We conclude from the inequalities

$$\underline{\int_a^b} f + \underline{\int_a^b} g \leq \underline{\int_a^b}(f+g) \leq \overline{\int_a^b}(f+g) \leq \overline{\int_a^b} f + \overline{\int_a^b} g,$$

that (1) is true.

To prove (2), choose $K > 0$ big enough such that $\max\{|f(x)|, |g(x)|\} \leq K$. Let $P = \{x_i; i = 0, \ldots, n\}$ be a partition of $[a,b]$, and $\omega_i', \omega_i'', \omega_i$ the oscillations of f,g and fg respectively, on the interval $[x_i, x_{i-1}]$. For $x, y \in [x_i, x_{i-1}]$ we have:

$$\begin{aligned}|f(y)g(y) - f(x)g(x)| &= |[f(y)-f(x)]g(y) + [g(y)-g(x)]f(x)| \\ &\leq \omega_i' K + \omega_i'' K = (\omega_i' + \omega_i'')K\end{aligned}$$

It follows that:

$$\sum_{k=1}^{n} \omega_i (x_i - x_{i-1}) \leq \sum_{k=1}^{n} (\omega_i' + \omega_i'')K(x_i - x_{i-1}),$$

and (2) is a direct consequence of Theorem 7(3).

Item (3) follows from (2), if we can show that $\frac{1}{g}$ is integrable. Let $P = \{x_i; i = 0, \dots, n\}$ be a partition of $[a, b]$, and $x, y \in [x_i, x_{i-1}]$. By hypothesis:

$$\left|\frac{1}{g(x)} - \frac{1}{g(y)}\right| \leq \frac{|g(y) - g(x)|}{k^2}.$$

Once more, the result follows from Theorem 7(3).

Item (4) is trivial, since in this case $s(f; P) \leq s(g; P)$ for every partition, hence $\int_a^b f \leq \int_a^b g$. Finally, to see why (5) is true, consider the inequality:

$$||f(x)| - |f(y)|| \leq |f(x) - f(y)|$$

Which tell us that the oscillation of $|f|$ is always bounded by the oscillation of $|f|$, hence by Theorem 7(3) again, $|f|$ is integrable. The last part follows from the inequality $-|f(x)| \leq f(x) \leq |f(x)|$. □

Corollary 11. *Let $f : [a, b] \to \mathbb{R}$ integrable and bounded, say $|f(x)| \leq K$. Then*

$$\left|\int_a^b f\right| \leq K(b - a).$$

Theorem 12. *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is integrable.*

Proof. By Theorem 68, f is uniformly continuous. Let $\epsilon > 0$ be given, and take $\delta > 0$ such that $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \frac{\epsilon}{b-a}$. Now, choose a partition $P = \{x_i; i = 0, \dots, n\}$ such that $x_i - x_{i-1} < \delta$ for every $i = 1, \dots, n$. If ω_i is the oscillation of f at $[x_{i-1}, x_i]$ then $\omega_i < \frac{\epsilon}{b-a}$ and it follows that

$$\sum_{k=1}^{n} \omega_i(x_i - x_{i-1}) < \frac{\epsilon}{b - a} \sum_{k=1}^{n} (x_i - x_{i-1}) = \epsilon.$$

The proof is complete by Theorem 7(3). □

Theorem 13. *Let $f : [a, b] \to \mathbb{R}$ be monotone. Then f is integrable.*

Proof. The argument is similar to the above theorem, namely it uses Theorem 7(3). Without loss of generality, suppose f increasing. Let $\epsilon > 0$ be given, choose a partition $P = \{x_i; i = 0, \dots, n\}$ such that $x_i - x_{i-1} < \frac{\epsilon}{f(b) - f(a)}$. We have:

$$\sum_{k=1}^{n} \omega_i(x_i - x_{i-1}) < \frac{\epsilon}{f(b) - f(a)} \sum_{k=1}^{n} \omega_i = \epsilon.$$

□

Recall that given an interval $I \subseteq \mathbb{R}$ with end-points a and b, the length of I, denoted by $|I|$, is given by $|I| = b - a$.

A set $X \subseteq \mathbb{R}$ has *measure zero* if given $\epsilon > 0$, it's possible to find a countable open cover of $X \subseteq \bigcup_{n=1}^{\infty} I_n$ by open intervals I_n, such that $\sum_{n=1}^{\infty} |I_n| < \epsilon$.

Example 14. *Any countable set $X \subseteq \mathbb{R}$ has measure zero. Indeed, given any $\epsilon > 0$, take an open interval of length $\frac{\epsilon}{2^n}$ around the n-th number $x_n \in X$, then $\sum_{n=1}^{\infty} |I_n| < \epsilon$. In particular, the set of Rational numbers $\mathbb{Q}$ has measure zero.*

Example 15. *The Cantor set K has measure zero since after the n-th iteration, K is contained in the union of 2^n intervals of length 3^{-n}. Hence, given any $\epsilon > 0$, if we take n sufficiently large, K can be covered by open sets whose length add to a number less than ϵ.*

Theorem 16. *Let $f : [a, b] \to \mathbb{R}$ be bounded function. If the set of discontinuities D of f has measure zero then f is integrable*

Proof. Let $\omega := \sup f - \inf f$, be the oscillation of f in $[a, b]$. Let $\epsilon > 0$ be given, and suppose $D \subseteq \bigcup_{n=1}^{\infty} I_n$, where I_n are open intervals such that $\sum_{n=1}^{\infty} |I_n| < \frac{\epsilon}{2\omega}$. For each $x \in [a, b] - D$, take an interval $J_x \ni x$, such that the oscillation of f in J_x is less than $\frac{\epsilon}{2(b-a)}$, this is possible because f is continuous at x.

Now, $[a, b] \subseteq \left(\bigcup_{n=1}^{\infty} I_n \right) \cup \left(\bigcup_{x \notin D} J_x \right)$, and by Borel-Lebesgue Theorem, there is a finite subcover, say $I_{n_1} \cup \dots I_{n_k} \cup J_{x_1} \cup \dots J_{x_l}$ of $[a, b]$. Form a partition P of $[a, b]$ whose elements are a, b, and each endpoint of I_{n_p} and J_{x_q}, for $p = 1, \dots k,\ q = 1, \dots, l$. We write $[y_{j-1}, y_j]$ for an interval associated to P which is contained in I_{n_p}, for some p, and $[y_{t-1}, y_t]$, otherwise. Let ω_j denote the oscillation of f in the j-th interval of P. We have:

$$
\begin{aligned}
S(f; P) - s(f; P) &= \sum \omega_j (y_j - y_{j-1}) + \sum \omega_t (y_t - y_{t-1}) \\
&< \sum \omega (y_j - y_{j-1}) + \sum \frac{\epsilon}{2(b-a)} (y_t - y_{t-1}) \\
&< \omega \frac{\epsilon}{2\omega} + \frac{\epsilon}{2(b-a)} (b - a) = \epsilon
\end{aligned}
$$

By Theorem Theorem 7(3), f is integrable. □

Example 17. *The Cantor function* $f : [0,1] \to \mathbb{R}$ *given by*

$$f(x) = \begin{cases} 1, & \textit{if } x \in K \\ 0, & \textit{if } x \notin K, \end{cases}$$

is integrable. Indeed, f *is continuous in* $[0,1]-K$ *because it's constant there, but it's discontinuous at every point* a *of* K*, since we can find a sequence* $x_n \in [0,1] - K$ *such that* $x_n \to a$*. By Theorem 16,* f *is integrable.*