$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x)$$

Analysis 1

Continuity, Differentiability & Integration

Giuseppe Bassi

Academic Year 2023/2024

Abstract

These notes originate from a personal undertaking to consolidate and summarize the core material of an Analysis I course taught by Professor Giuseppe Savaré at Bocconi University. They offer a concise re-elaboration of the main topics covered, rather than an exhaustive, textbook-level treatment. Hence, they are not intended to serve as an official reference and do not provide comprehensive coverage of the entire syllabus. I encourage readers to suggest improvements or report any errors they may encounter so that future editions can be improved.

Version 1.0 Date: September 23, 2025

Contents

1 Real Numbers

In this section, we present some fundamental results on real numbers. We will not construct them[1]. Instead, we *assume* the existence and uniqueness of a set $\mathbb{R}$ satisfying three sets of axioms:

i) *Algebraic Axioms:* we endow the set $\mathbb{R}$ with multiplication and addition, obtaining the triple $(\mathbb{R}, +, \cdot)$. Specific axioms define how operators behave on $\mathbb{R}$, ensuring that this triple is a *field*;

ii) *Ordering Axioms:* we endow the *field of real numbers* with a total order relation $\leq$. Specific axioms clarify how to compare real numbers and how $\leq$ interacts with $+$ and $\cdot$. The quadruple $(\mathbb{R}, +, \cdot, \leq)$ is an *ordered field*;

iii) *Completeness Axiom:* The completeness axiom is the final axiom needed to distinguish $(\mathbb{R}, +, \cdot, \leq)$ and $(\mathbb{Q}, +, \cdot, \leq)$.

We also assume familiarity with some basic concepts from set theory, elementary logic (especially the principle of induction), fundamental properties of functions, and elementary functions.

1.1 Initial Definitions

Let us start our discussion from some useful definitions:

Definition 1.1 (Upper Bound). A set $X \subseteq \mathbb{R}$ is *bounded from above* (resp. below) if there exists a real number C, called *upper bound* for X (resp. lower bound), such that $x \leq C$ (resp. $x \geq C$) for every $x \in X$.

Definition 1.2 (Bounded Set). A set $X \subseteq \mathbb{R}$ is *bounded* if it is bounded from above and from below. Equivalently, X is bounded if there exists $M \in \mathbb{R}$ such that $|x| \leq M$ for every $x \in X$.

Similarly, a function is said bounded if its image is bounded. For example, the function $f : \mathbb{N}^+ \to \mathbb{R}$ defined as $f(n) = \frac{1}{n}$ is bounded, while the function $f : \mathbb{R}^+ \to \mathbb{R}$ defined as $f(x) = \frac{1}{x}$ is unbounded.

Definition 1.3 (Maxima and minima). Given a set X, its largest number is called *maximum* and its smallest number is called *minimum*.

Similarly, the *maximum* of a function $f : X \to Y$ is the maximum element of its range $\operatorname{Im} f \subseteq Y$. A point $x^\star$ for which f attains its maximum (resp. minimum) is called *maximizer* (resp. *minimizer*) for f.

Definition 1.4 (Supremum and Infimum). Given a set $X \subseteq \mathbb{R}$, its supremum is the smallest upper bound S, and its infimum is the greatest lower bound. By definition, the following are equivalent:

i) S is the supremum for X;

ii) $x \leq S$ for every $x \in X$ <u>and</u> $S \leq C$ for every upper bound C of X;

iii) $x \leq S$ for every $x \in X$ <u>and</u>, for every $\varepsilon > 0$, there exists $x \in X$, such that $S - \varepsilon < x$. This third definition is the one useful for the exercises.

We conclude this section with three additional definitions which will be useful when dealing with limits and continuity:

Definition 1.5 (Neighbourhoods). The bounded, open interval $I_r(\ell) := (\ell - r, \ell + r) = \{x \in \mathbb{R} : |x - \ell| < r\}$ is called a *neighbourhood* of ℓ of radius $r > 0$.

Remark 1.6. It's sometimes useful to consider **pointed** neighbourhoods, where we discard the center itself from $I_r(\ell)$, that is, $\dot{I}_r(\ell) := (\ell - r, \ell + r) \setminus \{\ell\} = \{x \in \mathbb{R} : 0 < |x - \ell| < r\}$

Remark 1.7. for $\ell = \pm\infty$ you can define $I_m(\pm\infty)$ as the interval $(m, +\infty)$ or $(-\infty, m)$.

Definition 1.8 (Accumulation points). A point $x_0 \in \mathbb{R}$ is an *accumulation* point for $A \subseteq \mathbb{R}$ if one of the following <u>equivalent</u> conditions hold:

i) Every <u>pointed</u> neighbourhood (no matter how small) of x_0 contains points of A.

ii) Every neighbourhood of x_0 contains points of A <u>different</u> from x_0.

iii) Every neighbourhood of x_0 contains infinitely many points of A.

[1] The constructive approach relies on set theory and logic to define the natural numbers $\mathbb{N}$ axiomatically (via the Peano Axioms), and then build $\mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$ through a series of rigorous and tedious constructions.

iv) There exists a sequence of points of $A \setminus \{x_0\}$ converging to x_0 *(Convergence will be defined later in these notes).*

Remark 1.9. $+\infty$ is an accumulation point of A if for every $y \in \mathbb{R}$ there exists $x \in A$ such that $x > y$.

Definition 1.10 (Isolated Points). $x_0 \in A$ is an isolated point if there exists a (small) radius r for which $I_r(x_0)$ contains just x_0 as elements of A.

Definition 1.11 (Interior). Given a set $D \subseteq \mathbb{R}$, a point $x_0 \in D$ is in its *interior* if there exists a neighbourhood of x_0 completely included in D. In other words, x_0 is in the interior of D if x_0 is not an accumulation point of $\mathbb{R} \setminus D$.

Remark 1.12. We remark the relationship between isolated points and accumulation points given a set $A \subseteq \mathbb{R}$. Isolated point for A *must* belong to A. Hence, an isolated point is a point in A which is not an accumulation point for A. Meanwhile, accumulation points may belong or not to A.

1.2 Completeness

In this section we focus on the *completeness* axiom, which is the crucial property that distinguishes $\mathbb{R}$ and $\mathbb{Q}$, enabling us to rigorously define concepts such as continuity and limits, and laying the foundation for the theory of infinitesimal changes, namely, real analysis.

Definition 1.13 (Completeness). If A, B are *nonempty* sets of real numbers such that $a \leq b$ for every $a \in A$, $b \in B$, then there exists a separating element $x \in \mathbb{R}$ such that $a \leq x \leq b$.

Completeness is not the only way to define real numbers. Indeed, there is an alternative but completely equivalent axiom:

Proposition 1.14 (Existence of Supremum). *The completeness property is equivalent to the supremum property, that is, the existence of a supremum for any set bounded from above.*

Proof.

We first prove that the *supremum property* implies the existence of a separating element for every two nonempty sets A and B. Consider two **nonempty** sets A, B and let a and b denote generic elements belonging to A and B respectively *(so a property which holds for a is true for every element in A).* If $\boldsymbol{a \leq b}$, then A is bounded from above by every element of B. So A has a supremum S. Therefore $\boldsymbol{S \geq a}$. If we prove also $\boldsymbol{S \leq b}$ then S would be a separating element. By contradiction, if $\exists\, b' \in B$ such that $b' < S$, then b' would be an upper bound of A smaller than S.

Next, we prove that the existence of a separating element implies the *supremum property.* If A is a **nonempty** set bounded from above, the set B of upper bounds of A is nonempty, so there exists a separating element S such that $\boldsymbol{a \leq S \leq b}$. S is an upper bound since $a \leq S$, but S is the least upper bound since $S \leq b$. $\square$

We now state, without proof, an important property of real numbers, which is a direct result of the order axioms and the completeness axiom.

Lemma 1.15 (Archimedean Property). *For every real number x, there exists a natural number n larger than x.*

Corollary 1.16. *For every $0 < a < b$, there exists $n \in \mathbb{N}$ such that $na > b$.*

Theorem 1.17 (Density of $\mathbb{Q}$ in $\mathbb{R}$). *$\mathbb{Q}$ is a dense set in $\mathbb{R}$, i.e., every two real numbers can be separated by a rational number. Formally,*

$$\forall x < y \in \mathbb{R},\ \exists\, q \in \mathbb{Q}\ :\ x < q < y.$$

Proof.

We distinguish three cases. First, assume $0 \leq a < b$. Since $\frac{1}{b-a} > 0$, by the archimedean property there exists $n \in \mathbb{N}^+$ such that $n > \frac{1}{b-a}$, which implies $\frac{1}{n} < b - a$.
Moreover, by the archimedean property there exists $k \in \mathbb{N}^+$ such that $k > an$. Let m be the smallest natural number for which $m > an$. Therefore $\boldsymbol{\frac{m}{n} > a}$ and $\boldsymbol{\frac{m-1}{n} \leq a}$. However,

$$\frac{m}{n} = \frac{m-1}{n} + \frac{1}{n} \leq a + \frac{1}{n} < a + b - a = b \implies a < \frac{m}{n} < b.$$

Hence, we found a rational number separating a and b.
If $a < 0 < b$, just pick $q := 0$. Finally, if $a < b \leq 0$, we can consider $a' := -a$ and $b' := -b$. Then $0 \leq b' < a'$. By the previous result, we can find $q' \in \mathbb{Q}$ such that $b' < q' < a'$. To conclude, set $q := -q'$. □

The next proposition is historically important. It is a constructive example of a number that belongs to $\mathbb{R} \setminus \mathbb{Q}$.

Proposition 1.18 ($\sqrt{2}$ is irrational). *$\sqrt{2}$ does not belong to $\mathbb{Q}$, but it belongs to $\mathbb{R}$*

Proof.
By contradiction, we assume that there exists $q \in \mathbb{Q}$ such that $q^2 = 2$. Write $q = \frac{m}{n}$ with $m, n \in \mathbb{Z}$, $\gcd(m, n) = 1$, and $n > 0$. Among all such representations, let q be one with the smallest denominator. Now

$$\boldsymbol{q} = q \cdot \frac{q-1}{q-1} = \frac{q^2 - q}{q-1} = \frac{2-q}{q-1} = \frac{\boldsymbol{2n - m}}{\boldsymbol{m - n}}.$$

Noting that $1 < q < 2$, we see that $0 < n < m < 2n$. Hence the denominator $m - n$ is strictly smaller than n, and we have found another representation of a solution to $q^2 = 2$ with a smaller denominator. This contradicts the minimality assumption.
Consider the sets $A = \{a \in \mathbb{R}^+ \mid a^2 < 2\}$ and $B = \{b \in \mathbb{R}^+ \mid b^2 > 2\}$. They are nonempty and $a \leq b$ $\forall\, a \in A, b \in B$. By the completeness axiom, there exists a separating element ξ such that $a \leq \xi \leq b$, which implies $a^2 \leq \xi^2 \leq b^2$. If $\xi^2 < 2$ there exists $q \in \mathbb{Q}$ such that $\xi^2 < q^2 < 2$ (thanks to the density of $\mathbb{Q}$ in $\mathbb{R}$), so $q \in A$ and therefore ξ wouldn't be the separating element. Similarly if $\xi^2 > 2$, so the only way is that $\xi^2 = 2$. This proves that $\sqrt{2} \in \mathbb{R}$ and $\mathbb{Q}$ does not satisfy the completeness axiom. □

2 Sequences

Real-valued sequences are functions from $\mathbb{N}$ to $\mathbb{R}$, assigning to each natural number n a corresponding real value a_n.[2] The aim of this section is to establish the foundational tools needed to study the behaviour of sequences as n grows large, specifically to rigorously define the concepts of "approaching" and "converging".

2.1 Limits of Sequences

Definition 2.1 (Limit). the sequence a_n converges to ℓ as n approaches ∞, and we write $\lim_{n\to\infty} a_n = \ell$, if

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

The same definition via neighbourhoods would be: for every neighbourhood $\mathcal{I} = I_\varepsilon(\ell)$ of ℓ there exists a neighbourhood $\mathcal{J} = I_N(+\infty)$ of $+\infty$ such that $a_n \in \mathcal{I}$ for every $n \in \mathbb{N} \cap \mathcal{J}$.

Remark 2.2. a_n converges to ℓ if for every tolerance $\varepsilon > 0$, a_n will be eventually contained in the interval $(\ell - \varepsilon, \ell + \varepsilon)$, that is, there are at most a finite number of indices for which a_n does not belong to that interval. This implies that the asymptotic behaviour of a sequence does not change if we modify, insert or suppress a finite number of terms.

Definition 2.3 (Diverging sequence). The sequence a_n diverges to $+\infty$ (resp. $-\infty$) as n approaches ∞, and we write $\lim_{n\to\infty} a_n = +\infty$ (resp. $-\infty$), if

$$\forall\, m > 0\ \exists\ N_m \in \mathbb{N} : a_n > m\ \forall\ n \geq N_m \quad (\text{resp. } a_n < -m)$$

Definition 2.4 (Irregular sequence). A sequence is called regular if it has a (finite or infinite) limit; otherwise it is irregular. For instance, $a_n = (-1)^n$ is a bounded irregular sequence, and $b_n = (-1)^n \cdot n$ is an unbounded irregular sequence.

Definition 2.5 (Landau notation). Landau notation is a compact way to estimate the asymptotic growth of sequences and functions. We will first define them for sequences, but they can be naturally extended to functions.

[2] A more precise definition is as follows: a real-valued sequence is a real function that is eventually defined on $\mathbb{N}$. "Eventually" means that the function may be undefined for at most a finite number of natural numbers, but beyond a certain point, it is always defined. For example, the sequence $a_n = \sqrt{n - 2024}$ is defined for $n \geq 2024$, while $b_n = \sqrt{2024 - n}$ is not a sequence, as it is only defined for a finite number of natural numbers.

i) $a_n = o(b_n)$ as $n \to +\infty$, which reads "a_n is little oh of b_n", if $\lim_{n\to+\infty} \frac{a_n}{b_n} = 0$.

ii) $a_n \sim b_n$, which reads "a_n is asymptotically equivalent to b_n", if $\lim_{n\to+\infty} \frac{a_n}{b_n} = 1$.

iii) $a_n = O(b_n)$, which reads "a_n is big o of b_n", if $\frac{a_n}{b_n}$ is bounded eventually[3].

Some remarks:

- We should require a_n and b_n eventually nonzero.
- $a_n = o(b_n)$ in particular implies $a_n = O(b_n)$. Similarly, $a_n \sim b_n$ implies $a_n = O(b_n)$. Therefore, big o is a weak and imprecise symbol.
- The asymptotic notation is used to reduce intricate sequences into simpler ones, maintaining the principal part. For example, $n^2 + n + \log(n) \sim n^2$
- Asymptotic equivalence and little o are transitive: if $a_n \sim b_n$ and $b_n \sim c_n$, then $a_n \sim c_n$; if $a_n = o(b_n)$ and $b_n = o(c_n)$, then $a_n = o(c_n)$.
- Asymptotic equivalence behaves well with products, quotients and powers. If $a_n \sim a'_n$ and $b_n \sim b'_n$, then
$$a_n b_n \sim a'_n b'_n \qquad \frac{a_n}{b_n} \sim \frac{a'_n}{b'_n} \qquad a_n^\alpha \sim a_n'^\alpha$$
- Asymptotic equivalence behaves badly with sums when the principal part cancels out. For example, $n^3 + n^2 \sim n^3 + 1$ and $n^3 + n \sim n^3$, but $(n^3 + n^2) - (n^3 + n) = n^2 - n$ is not asymptotically equivalent to $(n^3 + 1) - n^3 = 1$.
- Asymptotic equivalence and little o are linked by $a_n \sim b_n \iff a_n = b_n + o(b_n)$. In fact,
$$a_n \sim b_n \iff \lim_{n\to\infty} \frac{a_n}{b_n} = 1 \iff \lim_{n\to\infty} \frac{a_n - b_n}{b_n} = 0 \iff a_n - b_n = o(b_n) \iff a_n = b_n + o(b_n).$$
This motivates why it's always better to work with little o, which is "safer" (e.g. when the principal part cancels out).
- if $a_n = o(1)$ as $n \to \infty$, then $\lim_{n\to\infty} a_n = 0$. Thus, $o(1)$ denotes a generic sequence converging to 0.

We now prove some key results for limits of sequences:

Proposition 2.6 (Uniqueness of limits). *If a limit exists, then it's unique.*

Proof.
By contradiction, suppose there are two distinct limits.
Case 1: both limits are real numbers. Let $\ell_1 \neq \ell_2$ and fix $\varepsilon > 0$. We set $N := \max(N_1, N_2)$, so that for $n > N$ we get $|a_n - \ell_1| < \varepsilon$ and $|a_n - \ell_2| < \varepsilon$. This means that a_n is distant from ℓ_1 less than ε, and distant from ℓ_2 less than ε. To find a contradiction, it's sufficient to take $\varepsilon < \frac{|\ell_1 - \ell_2|}{2}$.[a] To prove formally this intuition, we can employ the triangle inequality $|a + b| \leq |a| + |b|$:
$$|\ell_1 - \ell_2| = |(\ell_1 - a_n) + (a_n - \ell_2)| \overset{!}{\leq} |\ell_1 - a_n| + |a_n - \ell_2| < 2\varepsilon.$$
Since ℓ_1 and ℓ_2 are fixed but the inequality holds for every $\varepsilon > 0$, we have $|\ell_1 - \ell_2| = 0$, implying $\ell_1 = \ell_2$.
Case 2: $\ell_1 \in \mathbb{R}$ and $\ell_2 = +\infty$. Given $\varepsilon > 0$, by ℓ_1 being a limit there exists N_1 such that for all $n > N_1$,
$$|a_n - \ell_1| < \varepsilon \Rightarrow a_n < \ell_1 + \varepsilon.$$
Let $M := \ell_1 + \varepsilon$. Since $\ell_2 = +\infty$, there exists N_2 such that for all $n > N_2$, $a_n > M = \ell_1 + \varepsilon$. With $N := \max\{N_1, N_2\}$ we get, for all $n > N$, both $a_n < \ell_1 + \varepsilon$ and $a_n > \ell_1 + \varepsilon$, a contradiction.
The case $\ell_2 = -\infty$ is analogous: from $|a_n - \ell_1| < \varepsilon$ we have $a_n > \ell_1 - \varepsilon$, while $a_n \to -\infty$ gives $a_n < \ell_1 - \varepsilon$ eventually, again a contradiction.

[3] In some books $a_n = O(b_n)$ if $\lim_{n\to+\infty} \frac{a_n}{b_n} = \ell \in \mathbb{R}$, which is stronger. For example, $(-1)^n = O(1)$ with the first definition, but not for the second.

Case 3: $\ell_1 = -\infty$ and $\ell_2 = +\infty$. Taking $M := 0$, from $a_n \to +\infty$ there exists N_2 with $a_n > 0$ for all $n > N_2$, and from $a_n \to -\infty$ there exists N_1 with $a_n < 0$ for all $n > N_1$. With $N := \max\{N_1, N_2\}$ we obtain a contradiction. □

[a]Comment: what we just did is the typical scheme for a proof with limits. You know that something, depending on the tolerance ε, eventually exists, and you choose the tolerance ϵ to be k times some quantity of the problem, to find some contradiction.

Proposition 2.7 (Converging sequences form a vector space). *The sum of two converging sequences converges to the sum of their limits. Moreover, $a_n \to a$ and $\lambda \in \mathbb{R}$, then $\lambda \cdot a_n \to \lambda \cdot a$.*

Proof.

Let $c_n := a_n + b_n$ and $c := a + b := \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$. We want to prove that $|c_n - c| < \varepsilon$ eventually. The intuition here is that eventually, a_n will be within ε of a, and similarly, b_n will be within ε of b. Thus, $a_n + b_n$ must be within 2ε of $a + b$, and since ε is arbitrarily small, we conclude the proof.
To formally prove this, we set $N = \max(N_a, N_b)$ and apply the triangle inequality:

$$|c_n - c| = |a_n + b_n - a - b| \overset{!}{\leq} |a_n - a| + |b_n - b| < \varepsilon + \varepsilon = 2\varepsilon \quad \forall\, n > N.$$

For the second part, if $\lambda = 0$, the theorem is trivial. Otherwise, we have: $|\lambda \cdot a_n - \lambda \cdot a| = |\lambda| \cdot |a_n - a| < |\lambda| \cdot \varepsilon = \varepsilon'$, where ε' is a positive tolerance arbitrarily small as ε gets small. □

Remark 2.8. similar results hold for other operations. The limit of the sum, difference, product and ratio of converging sequences is the sum, difference, product and ratio of the two limits. We can extend these results also to diverging sequences, but in such case pay attention to indeterminate forms![4]

Proposition 2.9 (Converging sequences are bounded). *if a sequence converges to a finite limit, then it's bounded.*

Proof.

Since we know that the limit exists, take $\varepsilon = 1$. So $\ell - 1 < a_n < \ell + 1$ for every $n > N$. Consider M as the maximum of $|a_n|$ for $n \leq N$. The maximum exists since it's a finite list. So we have a_n is bounded by $\max(M, \ell + 1)$ and $\min(-M, \ell - 1)$. □

Theorem 2.10 (Limit of monotone sequences). *Every monotone increasing sequence is regular, and $\lim\limits_{n\to\infty} a_n = \sup\limits_{n\in\mathbb{N}} a_n$. In particular, a_n converges to a finite limit if it is bounded from above, and diverges to $+\infty$ otherwise. Dual statements hold for decreasing sequences.*

Proof.

Suppose $S := \sup a_n \in \mathbb{R}$. We want to show that eventually

$$-\varepsilon < a_n - S < \varepsilon \quad \forall\, \varepsilon > 0.$$

Since S is an upper bound, we have $\boldsymbol{a_n - S \leq 0 < \varepsilon}$, so one inequality is already proven. Since S is the least upper bound, there exists an index N such that $a_N > S - \varepsilon$ for a fixed ε. Because the sequence is increasing, it follows that

$$a_n \geq a_N > S - \varepsilon \quad \forall\, n > N,$$

which implies $-\varepsilon < a_n - S$. This establishes the other inequality, valid for all $n > N$. Combining the two inequalities, we obtain $|a_n - S| < \varepsilon$, hence $\lim_{n\to\infty} a_n = S$.
If instead $(a_n)_{n\in\mathbb{N}}$ is not bounded, then for every $m \in \mathbb{R}$ there exists $N \in \mathbb{N}$ such that $a_N > m$. Since the sequence is increasing, it follows that $a_n > m$ for all $n > N$, which is exactly the definition of $a_n \to +\infty$. □

We conclude this section with a lemma which is more a general strategy when dealing with limits.

[4]For example, if $a_n \to 0$ and $b_n \to +\infty$, $a_n \cdot b_n$ may converge, diverge or none of them (consider the pairs $(a_{1,n}, b_{1,n}) = (\frac{1}{n}, n^2)$, $(a_{2,n}, b_{2,n}) = \left(\frac{1}{n^2}, n\right)$, $(a_{3,n}, b_{3,n}) = (\frac{-1}{n}, n^2)$, $(a_{4,n}, b_{4,n}) = (\frac{1}{n}, 15n)$, $(a_{5,n}, b_{5,n}) = \left(\frac{(-1)^n}{n}, n\right)$

Lemma 2.11. *Let $a < b$ and let $(a_n)_{n\in\mathbb{N}}$ be a sequence converging to a. Then there exists $x \in (a,b)$ such that $a_n < x$ for all sufficiently large n.*

Proof. Let $\varepsilon := \frac{b-a}{2}$. Since $a_n \to a$, there exists $N \in \mathbb{N}$ such that $|a_n - a| < \frac{b-a}{2}$ for every $n > N$. In particular, $a_n < \frac{a+b}{2}$ for all $n > N$. Thus, choosing $x := \frac{a+b}{2}$ concludes the proof. □

2.2 Theorems involving inequalities and limits

Lemma 2.12 (Squeeze). *If eventually $a_n \le b_n \le c_n$ and $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = \ell$, then $\lim_{n\to\infty} b_n = \ell$.*

Proof.
By definition, for every $\varepsilon > 0$,

$$|a_n - \ell| < \varepsilon \quad \forall\, n > N_a \quad \text{and} \quad |c_n - \ell| < \varepsilon \quad \forall\, n > N_c.$$

Take $N := \max\{N_a, N_c\}$. Thus, for every $n > N$, $a_n > \ell - \varepsilon$ and $c_n < \ell + \varepsilon$. Combining these two inequalities, we can squeeze b_n for every $n > N$ as

$$\ell - \varepsilon < a_n \le b_n \le c_n \le \ell + \varepsilon \implies |b_n - \ell| < \varepsilon \implies \lim_{n\to\infty} b_n = \ell.$$

□

Theorem 2.13. *Non-strict inequalities are preserved by limits. if $a_n \le b_n$ eventually $\lim_{n\to\infty} a_n \le \lim_{n\to\infty} b_n$.*

Proof.
Let $a := \lim_{n\to\infty} a_n$ and $b := \lim_{n\to\infty} b_n$. Eventually $a_n > a - \varepsilon$ and $b_n < b + \varepsilon$. Since $a_n \le b_n$ eventually,

$$a - \varepsilon < a_n \le b_n < b + \varepsilon \implies a - b < 2\varepsilon.$$

ε is arbitrarily close to 0, hence $a - b \le 0 \iff a \le b$. □

Theorem 2.14 (Permanence of the sign). *1) If a sequence converges to a positive real number, then it's eventually positive. 2) If a sequence is eventually non-negative, the limit is non-negative.*

Proof.
1) Given a sequence $(a_n)_{n\in\mathbb{N}}$, let $a := \lim\limits_{n\to\infty} a_n > 0$. Fixing $\varepsilon := a/2$, we deduce $a_n > a - \varepsilon = a/2 > 0$ eventually.
2) Assume $a_n \ge 0$ eventually and suppose, by contradiction, that $a < 0$. Taking $\varepsilon := -a/2 > 0$, we eventually have $a_n < a + \varepsilon = a - a/2 = a/2 < 0$, which contradicts the initial assumption. □

Lemma 2.15 (Cantor property). *if $a_n < b_n$, a_n is increasing, b_n is decreasing and $\lim_{n\to\infty} b_n - a_n = 0$, there exists a unique x which belongs to $[a_n, b_n]$ for every $n \in \mathbb{N}$, that is, $x \in \bigcap_{n\in\mathbb{N}} [a_n, b_n] = \sup a_n = \inf b_n$.*

[Proof]

2.3 Counterexamples and urban legends

In this section we provide a list of false statements with corresponding counterexamples.

i) If a sequence is bounded, then it eventually converges. The obvious example is $a_n = (-1)^n$.

ii) If the $\sup\limits_{n\in\mathbb{N}} a_n = +\infty$ then $a_n \to +\infty$. Two clear counterexamples are $a_n = (-1)^n \cdot n$ and $b_n = (1+(-1)^n)\cdot n$.

iii) If $a_n \to +\infty$, then a_n is eventually increasing. A clear counterexample is $a_n = n + 2 \cdot (-1)^n$.

iv) A convergent sequence is eventually monotonic. A clear counterexample if $a_n = \frac{(-1)^n}{n}$

v) If $a_n \to 0$, then $a_n \to 0^+$ or $a_n \to 0^-$ (if these symbols are not clear to you, they will be after dealing with limits of functions). A clear counterexample is $a_n = \frac{(-1)^n}{n}$.

vi) If $a_n \to 0^+$, then a_n is eventually decreasing. A clear counterexample is the sequence $a_n = \frac{1}{n+2\cdot(-1)^n}$.

vii) If $a_n > b_n$ eventually, $a_n \to a$ and $b_n \to b$, then $a > b$. A clear counterexample is $a_n = \frac{1}{n}$ and $b_n = 0$, where $a = b$. Moral: strict inequalities, like $a_n > 0$, become weak inequalities when taking limits. So $a_n > 0 \implies \lim_{n\to\infty} a_n \geq 0$.

2.4 Exercises

Example 1: principal part

$$\lim_{n\to\infty} \frac{n + 3n\sqrt{n} + 4}{2n - 4n\sqrt[3]{n} + 7}$$

Solution: Brutal mode:

$$\frac{n + 3n\sqrt{n} + 4}{2n - 4n\sqrt[3]{n} + 7} \sim \frac{3n^{3/2}}{-4n^{4/3}} = -\frac{3}{4}n^{1/6} \to -\infty.$$

Formal mode:

$$\frac{n + 3n\sqrt{n} + 4}{2n - 4n\sqrt[3]{n} + 7} = \frac{n^{3/2}\left(\frac{1}{\sqrt{n}} + 3 + \frac{4}{n\sqrt{n}}\right)}{n^{4/3}\left(\frac{2}{\sqrt[3]{n}} - 4 + \frac{7}{n\sqrt[3]{n}}\right)} = \frac{n^{3/2}}{n^{4/3}} \cdot \frac{o(1) + 3 + o(1)}{o(1) - 4 + o(1)} = n^{1/6}\frac{3}{-4} \to -\infty$$

Example 2: squeeze theorem

$$\lim_{n\to\infty} \frac{\sin(\log(n^2+7)) + \cos(n!)}{\sqrt{n}}$$

Solution: Notice that $-2 \leq \sin(\alpha) + \cos(\beta) \leq 2$ for every $\alpha, \beta \in \mathbb{R}$. Therefore, dividing by $\sqrt{n}$ and applying the squeeze theorem:

$$\underbrace{\boxed{-\frac{2}{\sqrt{n}}}}_{0} \leq \underbrace{\frac{\sin(\log(n^2+7)) + \cos(n!)}{\sqrt{n}}}_{0} \leq \underbrace{\boxed{\frac{2}{\sqrt{n}}}}_{0}$$

Example 3: orders of infinity

$$\lim_{n\to\infty} \frac{4n^4 3^n - 4^n + 2^n \log n}{3 \cdot 2^{2n-1} - 4^n \log n}$$

Solution:

$$\frac{4n^4 3^n - 4^n + 2^n \log n}{3 \cdot 2^{2n-1} - 4^n \log n} = \frac{4^n\left(\frac{4n^4 3^n}{4^n} - 1 + \frac{\log n}{2^n}\right)}{4^n\left(\frac{3}{2} - \log n\right)} = \frac{o(1) - 1 + o(1)}{\frac{3}{2} - \log n} \sim \frac{1}{\log n} \to 0.$$

Example 4: never cancel the principal part

$$\lim_{n\to\infty} \sqrt{n^4 + n^2 + 1} - n^2$$

Wrong solution: Since $\sqrt{n^2 + n^2 + 1} \sim n^2$, then $\sqrt{n^4 + n^2 + 1} - n^2 \to 0$.
First solution: rationalization

$$\sqrt{n^4+n^2+1} - n^2 = (\sqrt{n^4+n^2+1} - n^2) \cdot \frac{\sqrt{n^4+n^2+1} + n^2}{\sqrt{n^4+n^2+1} + n^2} = \frac{n^4 + n^2 + 1 - n^4}{\sqrt{n^4+n^2+1} + n^2} \sim \frac{n^2}{2n^2} = \frac{1}{2}.$$

Second solution: completing the square

$$\begin{aligned} \sqrt{n^4+n^2+1} - n^2 &= \sqrt{\left(n^2 + \frac{1}{2}\right)^2 + \frac{3}{4}} - n^2 = \left(n^2 + \frac{1}{2}\right)\sqrt{1 + \frac{3/4}{(n^2+1/2)^2}} - n^2 \\ &= \left(n^2 + \frac{1}{2}\right)\sqrt{1 + o(1)} - n^2 = n^2\left(\sqrt{1+o(1)} - 1\right) + \frac{1}{2}\sqrt{1+o(1)} = \frac{1}{2}. \end{aligned}$$

3 Subsequences

In this section we explore subsequences, which are sequences derived by removing certain elements from the original sequence. The study of subsequences is primarily practical and serves as the foundation for a series of key theorems in analysis. The first of these is the **Bolzano-Weierstrass** theorem, which we will later use to prove **Weierstrass's theorem**, a fundamental result that guarantees the existence of minima and maxima for continuous functions on closed and bounded intervals.
Following this, subsequences will be applied to investigate the relationship between continuity and uniform continuity, particularly through the **Heine-Cantor** theorem. This theorem is crucial for proving that continuous functions on a closed interval are integrable, a result that plays a central role in the theory of integration.

Definition 3.1 (Subsequence). Given a sequence $(a_n)_{n\in\mathbb{N}}$ and a *strictly increasing* map $\varphi : \mathbb{N} \to \mathbb{N}$, $k \mapsto \varphi(k)$, a subsequence $(a'_k)_{k\in\mathbb{N}}$ of $(a_n)_{n\in\mathbb{N}}$ is the composition of the map φ with the map $n \mapsto a_n$:

$$a'_k = a_{\varphi(k)}$$

The following are three ways to get an intuition of what a subsequence is:

i) To create a subsequence we select an infinite list of some natural numbers and we create a sequence considering just the integers of $(a_n)_{n\in\mathbb{N}}$ whose index corresponds to the natural numbers we selected. And then we shift the indices such that the n-th element of our new sequence has index n.

ii) A subsequence is a restriction of the map $a : n \to a_n$ for just some (but infinitely many) values of n

iii) A subsequence of $(a_n)_{n\in\mathbb{N}}$ can be obtained just eliminating some elements of the sequence. You can eliminate infinitely many, but the resulting sequence must have infinitely many terms.

Lemma 3.2 (Monotone subsequences). *Every sequence of real numbers has a monotone subsequence.*

Proof.
Given a sequence $(a_n)_{n\in\mathbb{N}}$ of real numbers, we say that a_k is a **peak** if $a_k \geq a_n$ for every $n > k$. We distinguish two cases:
Case A. The set of peaks is infinite. Then the subsequence consisting of all peaks is infinite and non-increasing, hence monotone.
Case B. The set of peaks is finite. Let N be the index of the last peak. Then for every $m > N$ there exists an index $i > m$ such that $a_i > a_m$ (otherwise m would itself be a peak). For each $m > N$, let $i(m)$ be the smallest index i with this property. Define inductively a subsequence by setting $n(1) := N$ and $n(k+1) := i(n(k))$. Then the subsequence $(a_{n(k)})$ is infinite and strictly increasing. □

Corollary 3.3. *Every sequence of real numbers has a subsequence which has a limit.*

Corollary 3.4 (Bolzano-Weierstrass). *Every **bounded** sequence of real numbers has a **convergent** subsequence.*

Proof.
Since the sequence is bounded, there exists a constant $M > 0$ such that $-M < a_n < M$ for every index $n \in \mathbb{N}$. By Lemma 3.2, there exists a monotone subsequence a'_k. Clearly $-M < a'_k < M$ for every index k, so a'_k is a monotone bounded sequence and therefore convergent. □

Corollary 3.5. *Every unbounded sequence from above (resp. below) has a subsequence diverging to $+\infty$ (resp. $-\infty$).*

Corollary 3.6 (Limits of subsequences). *A sequence a_n has no limit if **and only if** there exist two subsequences a'_k and a''_k converging to two different (possibly infinite) limits.*

Corollary 3.7. *If a sequences converges to a limit, every subsequence of it converges to the same limit.*

Theorem 3.8 (Cauchy). *A sequence $(a_n)_{n\in\mathbb{N}}$ is converging if and only if for every $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ such that, for any choice of $m, n > N$, we have $|a_m - a_n| < \varepsilon$.*[5]

[5] A sequence that satisfies this condition is called a Cauchy sequence. The fact that every Cauchy sequence converges to some real number in $\mathbb{R}$ is a profound result, closely tied to the completeness axiom. In fact, the standard construction of the real numbers is based on Cauchy sequences of rational numbers. In this construction, each equivalence class of Cauchy sequences of

[Proof]

Exercise: Prove that there exists a sequence $(x_n) \in (-2, 2)$ converging to 2 such that $\lim_{n\to+\infty} f(x_n)$ exists, where $f : [-2, 2] \to [-1, 1]$.

Solution: Let $a_n := 2 - 1/n$. Clearly $(a_n) \to 2$. The sequence $b_n := f(a_n)$ is bounded, thus, by Bolzano-Weierstrass' theorem, we can find a subsequence (b'_n) of (b_n) which has a limit. Then the corresponding subsequence (x_n) of (a_n) is a sequence converging to 2 such that $f(x_n)$ is convergent.

4 Series

In this chapter, we will focus on series, which refers to the sum of the elements of a sequence. Although a series involves summing up an infinite number of terms, it may still converge under certain conditions. For instance, the series $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots$ converges to 1, while other series, such as $\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots$, diverge to infinity. Most of the key results apply to series with nonnegative terms. For series with terms of mixed signs, we can either apply the powerful Leibniz criterion or reduce the problem to the standard case by examining the convergence of the absolute values of the terms. As a final remark, the study of real-valued series will be more complete after having covered Taylor series in the second part of the course, since many interesting series can be directly evaluated using this powerful tool. For example, the alternating series $1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} + \dots$ converges to $\frac{\pi}{4}$. But where does this π come from?

Definition 4.1 (Partial sums). Given a sequence $(a_n)_{n\in\mathbb{N}}$, we define the *sequence of partial sums* as $S_0 := a_0$ and $S_{n+1} := S_n + a_{n+1}$.

Definition 4.2 (Series). A *series* is the sum of all elements of a sequence a_n, that is,

$$\sum_{n=0}^{\infty} a_n := \lim_{n\to\infty} \sum_{i=0}^{n} a_i = \lim_{n\to\infty} S_n.$$

It's said convergent if $\lim_{n\to\infty} S_n \in \mathbb{R}$, divergent to $\pm\infty$ if the sequence S_n diverges, not converging[6] if S_n has no limit. Similarly to limits of sequences, the character of a series (i.e. convergence or divergence) is not affected by modifying, inserting or suppressing a finite number of terms, or shifting the indices.

Remark 4.3. Another property inherited from limits of sequences is the following: $\sum(a_n + b_n) = \sum a_n + \sum b_n$ as long as the two series are converging or one is converging and the other is diverging to $\pm\infty$. Moreover, $\sum q \cdot a_n = q \sum a_n$ for every $q \in \mathbb{R}$.

We now state the zero test, a *necessary* but not sufficient condition useful to prove that a series is *not* converging.

Theorem 4.4 (Zero test). *A series converges only if the sequence tends to zero.*

Proof.
Suppose that the series converges, and set $S := \lim_{n\to\infty} S_n \in \mathbb{R}$. Applying the shift property,

$$S = \lim_{n\to\infty} S_n = \lim_{n\to\infty} S_{n-1} + \lim_{n\to\infty} a_n = S + \lim_{n\to\infty} a_n \implies \lim_{n\to\infty} a_n = 0.$$

□

Proposition 4.5 (Geometric Series). *A geometric series of ratio x is the sum $\sum_{n=0}^{\infty} x^n$ and it converges whenever $|x| < 1$.*

rational numbers with a certain tail behaviour, that is, each class of sequences that get arbitrarily close to one another, is a real number. Therefore, one could almost argue that the convergence of every Cauchy sequence in $\mathbb{R}$ is the defining property that distinguishes $\mathbb{R}$ from $\mathbb{Q}$. In our construction, the equivalence between converging sequences and Cauchy sequences is a theorem based on the Bolzano-Weierstrass Theorem.

[6] In some books the term divergent series generically means "not convergent". This can be misleading since it would imply that both $a_n = (-1)^n$ and $b_n = n$ have divergent series, even if in one case the sequence of partial sums one has no limit, and in the other has limit $+\infty$.

Proof.
If $x = 1$, then the series diverges for the zero test. If $x \neq 1$, note that $xS_n = x + x^2 + \ldots + x^{n+1}$. Therefore

$$(1-x)S_n = 1 - x^{n+1} \implies S_n = \frac{1-x^{n+1}}{1-x}.$$

Passing the limit as $n \to \infty$ we conclude that the series converges if and only if $|x| < 1$ and

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

□

Theorem 4.6 (Nonnegative series). *A series of nonnegative terms converges if and only if S_n is bounded, and diverges to $+\infty$ otherwise.*

Proof.
Notice that S_n is increasing if and only if $(a_n)_{n\in\mathbb{N}}$ is nonnegative: $S_{n+1} = S_n + a_n \geq S_n$ if and only if $a_n \geq 0$. Therefore, S_n is a monotone sequence, hence it admits an (extended) limit. By Theorem 2.10 it converges to a finite number if and only if S_n is bounded, and diverges to $+\infty$ otherwise. □

Theorem 4.7 (Comparison Test). *If $0 \leq a_n \leq b_n$ eventually,*

i) if $\sum b_n$ converges, then $\sum a_n$ converges.

ii) if $\sum a_n$ diverges, then $\sum b_n$ diverges.

Proof.
Let k be the smallest index such that $0 \leq a_n \leq b_n \ \forall\ n > k$. Therefore,

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{k} a_n + \sum_{n=k+1}^{\infty} a_n, \quad \sum_{n=1}^{\infty} b_n = \sum_{n=1}^{k} b_n + \sum_{n=k+1}^{\infty} b_n.$$

Since $\sum_{n=1}^{k} a_n$ and $\sum_{n=1}^{k} b_n$ are finite, the behaviour of $\sum a_n$ and $\sum b_n$ depends only on the remaining series.
Let A_n be the sequence of partial sums of a_n starting from $n > k$, and B_n the sequence of partial sums of b_n starting from $n > k$. Thus, $\sum_{n=k+1}^{\infty} a_n = \lim_{n\to\infty} A_n$, and the same holds for B_n.
But since $0 \leq a_n \leq b_n \ \forall\ n > k$, it is clear that $0 \leq A_n \leq B_n \ \forall\ n > k$. As A_n and B_n are series of nonnegative terms, they either converge or diverge. Moreover, since limits preserve non-strict inequalities,

$$\lim_{n\to\infty} A_n \leq \lim_{n\to\infty} B_n.$$

This proves both claims. □

Theorem 4.8 (Asymptotic Comparison 1). *If $a_n \sim qb_n$ and they are nonnegative sequences, then the series $\sum a_n$ and $\sum b_n$ have the same behaviour.*

Proof.
If $\lim_{n\to\infty} \frac{a_n}{b_n} = q > 0$, then taking $\varepsilon = \frac{q}{2}$ we obtain that eventually

$$\frac{q}{2} < \frac{a_n}{b_n} < \frac{3q}{2},$$

so that

$$\frac{q}{2} b_n < a_n < \frac{3q}{2} b_n.$$

Setting $b_n' = \frac{3q}{2} b_n$, and noting that b_n and b_n' have the same behaviour, we see that if $\sum b_n$ converges, then $\sum b_n'$ converges. By pointwise comparison, $\sum a_n$ also converges.
On the other hand, setting $b_n'' = \frac{q}{2} b_n$, and noting that b_n and b_n'' have the same behaviour, we see that if $\sum b_n$ diverges, then $\sum b_n''$ diverges. By pointwise comparison, $\sum a_n$ also diverges.

□

Theorem 4.9 (Asymptotic Comparison 2). *If $a_n = o(b_n)$ and they are nonnegative sequences, then 1) if $\sum b_n$ converges, $\sum a_n$ converges. 2) If $\sum a_n$ diverges, $\sum b_n$ diverges.*

Proof.
If $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$, then eventually $0 \le \frac{a_n}{b_n} < 1 \implies a_n < b_n$, so we can apply the pointwise comparison test to conclude. $\square$

Theorem 4.10 (Ratio Test). *Given a sequence a_n of **positive** terms, if $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = q < 1$, then the series converges. If $q > 1$, the series diverges. If $q = 1$, the test is inconclusive.*

Proof.
If $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = q < 1$, then taking $\varepsilon := (1-q)/2$ we have

$$\frac{a_{n+1}}{a_n} - q < \varepsilon = \frac{1-q}{2} \iff a_{n+1} < a_n \cdot \frac{q+(1-q)}{2} = a_n \cdot \frac{q+1}{2} \ \forall\, n > N.$$

Defining $x := (q+1)/2$, we deduce $a_{n+N} \le a_{n-1+N} \cdot x$ for every $n \ge 0$. By induction, one proves $a_{n+N} \le a_N \cdot x^n$. Notice that $x < 1$ since $q < 1$, so the series $\sum a_N \cdot x^n$ converges. Using the comparison test, we conclude that $\sum a_n$ converges as well.
If $q > 1$, then taking $\varepsilon := (q-1)/2$ we have

$$-\varepsilon < \frac{a_{n+1}}{a_n} - q \iff a_{n+1} > a_n(q-\varepsilon) = a_n \cdot \frac{q+1}{2}.$$

Again, by induction one proves $a_{n+N} \ge a_N \cdot \left(\frac{q+1}{2}\right)^n$. Given that $\sum a_N \cdot \left(\frac{q+1}{2}\right)^n$ diverges, we conclude that $\sum a_n$ diverges as well. $\square$

Theorem 4.11 (Root Test). *Given a sequence a_n of **positive** terms, if $\lim_{n\to\infty} \sqrt[n]{a_n} = q < 1$, then the series converges. If $q > 1$, the series diverges. If $q = 1$, the test is inconclusive.*

Remark 4.12. The root test is stronger than the ratio test. The only reason to use the ratio test is that in some cases it may lead to simpler computations.

Proof.
Let $q = \lim_{n\to\infty} \sqrt[n]{a_n}$. If $q < 1$ by Lemma 2.11 we can find $x \in (q, 1)$ such that $\sqrt[n]{a_n} < x$ eventually. Hence, $a_n < x^n$ eventually and in particular $\sum a_n$ converges by comparison test with $\sum x^n$. Similarly, if $q > 1$ by Lemma 2.11 we can find $x \in (1, q)$ such that $\sqrt[n]{a_n} > x$ eventually. Hence, $a_n > x^n$ eventually and in particular a_n does not converge to zero, hence the series diverges by zero test. $\square$

The previous tests are useful to test convergence, but they may fail, as the following example shows:

Example 4.13 (Harmonic Series). A generalized Harmonic Series of parameter γ is defined as $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^\gamma}$ and it's the most important example of a divergent series for which none of the previous tests help. If we try to apply the root test and the ratio test, they both give us 1, which is an inconclusive result:

$$\lim_{n\to+\infty} \sqrt[n]{\frac{1}{n^\gamma}} = \lim_{n\to\infty} \left(\frac{1}{\sqrt[n]{n}}\right)^\gamma = 1 \qquad \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \left(\frac{n+1}{n}\right)^\gamma = 1.$$

To study the Harmonic Series, we need a new theorem, called **Cauchy Condensation**:

Theorem 4.14 (Cauchy Condensation). *Given a **decreasing** sequence a_n of **positive** terms, then $\sum_{n=1}^{+\infty} a_n$ and $\sum_{h=0}^{+\infty} 2^h a_{2^h}$ have the same behaviour.*

Proof.

$$\sum_{n=1}^{\infty} a_n = \underbrace{a_1}_{\le a_1} + \underbrace{a_2 + a_3}_{\le a_2 + a_2} + \underbrace{a_4 + a_5 + a_6 + a_7}_{\le a_4 + a_4 + a_4 + a_4} + \cdots \le a_1 + 2a_2 + 4a_4 + \cdots = \sum_{h=0}^{\infty} 2^h a_{2^h}.$$

On the other hand,

$$\sum_{h=0}^{\infty} 2^h a_{2^h} = \underbrace{a_1 + a_2}_{\le a_1+a_1} + \underbrace{a_2 + a_4 + a_4 + a_4}_{\le a_2+a_2+a_3+a_3} + \underbrace{a_4 + a_8 + a_8 + \ldots + a_8}_{\le 2a_4+2a_5+2a_6+2a_7} + \underbrace{a_8 + 15a_{16}}_{\le 2a_8+2a_9+\cdots+2a_{15}} + \cdots \le 2\sum_{n=1}^{\infty} a_n.$$

Combining these two results we conclude that

$$\sum_{n=1}^{+\infty} a_n \le \sum_{h=0}^{+\infty} 2^h \cdot a_{2^h} \le 2\sum_{n=1}^{+\infty} a_n \implies \text{the two series have the same behaviour.}$$

□

Example 4.15. Cauchy condensation can be used to prove that the generalized harmonic series converges if and only if $\gamma > 1$. Instead of the series of a_n, we analyse the series of $2^h a_{2^h} = \frac{2^h}{(2^h)^\gamma} = \left(\frac{1}{2^{\gamma-1}}\right)^h$. This series is a geometric series with ratio $2^{1-\gamma}$, thus it converges if and only if $-1 < 2^{1-\gamma} < 1 \iff 1-\gamma < 0 \iff \gamma > 1$. In particular, the harmonic series ($\gamma = 1$) is the limiting case of divergence.

Theorem 4.16 (Absolute Convergence). *If $\sum |a_n|$ is convergent, then $\sum a_n$ is convergent as well and it's said that the series is absolutely convergent.*[7] *Moreover, $|\sum a_n| \le \sum |a_n|$.*

Proof.
Let p_n be the subsequence of the positive elements of a_k and q_k the subsequence of the negative elements of a_k. It's trivial that $S_n = \sum a_k = \sum p_k + \sum q_k$. Moreover $0 \le \sum p_k \le \sum |a_n|$ and $0 \le \sum -q_k \le \sum |a_n|$, so by comparison test, both $\sum p_k$ and $\sum -q_k$ are convergent, thus also $\sum a_k = \sum p_k - \sum -q_k$ is convergent. □

Theorem 4.17 (Leibniz Criterion). *Given a **decreasing** sequence of **positive** real numbers a_n which converges to zero, the series with **alternating sign** $\sum(-1)^n a_n$ is converging to a real number S.*[8]

Proof.
We show that **the sequence of intervals $[S_{2n-1}, S_{2n}]$ satisfies the conditions of Cantor Theorem** and S is the unique element of their intersection.

i) $S_{2n} = S_{2n-1} + a_{2n} > S_{2n-1}$, since $a_{2n} > 0$.

ii) S_{2n} is decreasing: $S_{2n+2} = S_{2n} + a_{2n+2} - a_{2n+1} \le S_{2n}$, since $a_{2n+2} \le a_{2n+1}$.

iii) With the same trick we prove that S_{2n+1} is increasing.

iv) $S_{2n} - S_{2n-1} = a_{2n} \to 0$.

Therefore $\exists! \, \boldsymbol{x \in [S_{2n-1}, S_{2n}] \ \forall\, n} \in \mathbb{N}$, and $x = \sup S_{2n-1} = \inf S_{2n}$. Since S_{2n} and S_{2n-1} are monotone, from the property of the limit of monotone sequences we deduce that $\lim_{n\to\infty} S_{2n-1} = \lim_{n\to\infty} S_{2n} = x$. So we have that $|S_n - \ell| < \varepsilon \ \forall n > p$, with n even, and $|S_n - \ell| < \varepsilon \ \forall n > d$, with n odd. Taking $m = \max\{p, d\}$ we get $\boldsymbol{|S_n - \ell| < \varepsilon \ \forall n > m}$, which is exactly our thesis. □

The following corollary is even more useful for exercises:

Corollary 4.18. *The sequence of even partial sums is decreasing, converging to S and can be bounded by $0 \le S_{2n} - S \le a_{2n+1}$. The sequence of odd partial sums is increasing, converging to S and can be bounded by $0 \le S - S_{2n-1} \le a_{2n}$.*

[7] Absolute convergence is stronger than convergence, that is, some sequences may converge not absolutely. For example, the alternating harmonic series $1 - 1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} + \cdots = \log(2)$, even if it doesn't converge absolutely. Therefore, it's a sufficient condition which is not necessary.

[8] Any of these assumptions is necessary for the theorem. $a_n = \frac{n+1}{n}$ shows that the convergence to zero is required. $a_n = \frac{(-1)^n}{n}$ shows that positivity is also required. $\sum(-1)^n a_n$ diverge. Monotonicity is also fundamental. A counterexample is $\frac{1}{1} - \frac{1}{4} + \frac{1}{3} - \frac{1}{16} + \frac{1}{5} - \frac{1}{36} + \ldots - \frac{1}{n^2} + \frac{1}{n+1} + \ldots$, which satisfies all the requirements except monotonicity, and it indeed diverges. Note also that Leibniz criterion is a sufficient but not necessary condition for convergence: $a_n = \frac{1}{n+(-1)^n}$ is not monotonic, but $\sum(-1)^n a_n$ converges.

Proof. Just by looking at the previous proof,

$$\boldsymbol{S_{2n-1} \leq S_{2n+1} = S_{2n} - a_{2n+1} \leq S \leq S_{2n} = S_{2n-1} + a_{2n}} \implies \begin{cases} S - S_{2n-1} \leq a_{2n} \\ \text{and} \\ S_{2n} - S \leq a_{2n+1} \end{cases}.$$

□

4.1 Exercises

In these examples you are asked to study the behaviour of a list of series. Remember: before applying comparison test, asymptotic comparison, ratio or root tests, assure the sequence is nonnegative.[9] Otherwise, use the absolute convergence.

Example 1: comparison test

$$\sum_{n=1}^{\infty} \frac{1}{n!} \leq \sum_{n=1}^{\infty} \frac{1}{2^{n-1}} = 2,$$

where the inequality comes from $n! \geq 2^{n-1}$ for every $n \in \mathbb{N}$, which can be proved by induction.

Example 2: ratio test

$$\sum_{n=0}^{\infty} (n!)^{\alpha} \sin\left(\frac{e^n + n^5}{(2n+1)! - 2^n}\right).$$

Solution: Since $\lim_{n\to\infty} \frac{e^n}{(2n+1)!} = 0$, asymptotic comparison implies that $(n!)^{\alpha} \sin\left(\frac{e^n+n^5}{(2n+1)!-2^n}\right) \sim \frac{(n!)^{\alpha} e^n}{(2n+1)!} = b_n$ as $n \to +\infty$. We can then apply the ratio test to this b_n:

$$\frac{b_{n+1}}{b_n} = \frac{((n+1)!)^{\alpha} e^{n+1}}{(2n+3)!} \cdot \frac{(2n+1)!}{(n!)^{\alpha} e^n} = \frac{e(n+1)^{\alpha}}{(2n+3)(2n+2)} \sim \frac{e}{4} \cdot n^{\alpha-2}, \text{as } n \to +\infty.$$

Therefore, the series converges if $\alpha \leq 2$.

Example 3: asymptotic comparison

$$S(x) = \sum_{n=1}^{\infty} n\left(\frac{1}{n^x} - \sin\left(\frac{1}{n^{2x-1}}\right)\right), \quad x \geq \frac{1}{2}$$

Solution: let's study the general term a_n for different values of x:

- If $x = 1/2$, then $a_n \sim \sqrt{n} - n\sin(1)$, thus the series diverges for the zero test.
- If $1/2 < x < 1$, then

  $$a_n \to \frac{1}{n^{x-1}} - \frac{1}{n^{2x-2}} = \frac{1}{n^{x-1}}\left(1 - \frac{1}{n^{x-1}}\right) = n^{1-x}\left(1 - n^{1-x}\right) \to -\infty,$$

  thus the series diverges for the zero test.
- If $x = 1$, then

  $$a_n = 1 - n\sin\left(\frac{1}{n}\right) \sim 1 - n\left(\frac{1}{n} - \frac{1}{6n^3}\right) \sim \frac{1}{6n^2},$$

  where in the second passage we Taylor expanded the sine[10]. Thus, the series converges by asymptotic comparison with $1/6n^2$.
- If $x > 1$, then

  $$a_n \sim \frac{1}{n^{x-1}} - \frac{1}{n^{2x-2}} = \frac{1}{n^{x-1}}\left(1 - \frac{1}{n^{x-1}}\right) \sim \frac{1}{n^{x-1}},$$

  thus the series converges for $x > 2$ and diverges for $1 < x \leq 2$.

[9]For example, if $a_n = (-1)^n \frac{\sqrt{n}+(-1)^n}{n}$, we may be tempted to use the asymptotic comparison with $b_n = (-1)^n \frac{1}{\sqrt{n}}$. In fact, you can show that $a_n \sim b_n$, and b_n converges from Leibniz criterion. BUT that would lead to a wrong conclusion, since $a_n = b_n + \frac{1}{n}$, thus its series diverges. Conclusion: never apply asymptotic comparison to sequences of mixed signs.

[10]Taylor expansions can be found later in these notes

Example 4: root test

$$\sum_{n=1}^{\infty} \frac{3^{n^2}}{(n!)^n}.$$

Solution: the series converges and it's sufficient to apply the root test

$$\lim_{n\to+\infty} \sqrt[n]{a_n} = \lim_{n\to+\infty} \sqrt[n]{\frac{3^{n^2}}{(n!)^n}} = \lim_{n\to+\infty} \frac{3^n}{n!} = 0.$$

Example 5: Leibniz Criterion Prove that the following series is convergent: :

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{\sqrt{n+\frac{1}{n}}}{n}.$$

Find an index N after which the partial sums S_n differ by no more than 10^{-2} from the limit (N not necessarily being the smallest solution).
Solution: Let $a_n := \frac{\sqrt{n^2+1}}{n\sqrt{n}}$. We want to apply Leibniz Criterion: it's immediate to check that $a_n \geq 0$ and $\lim\limits_{n\to\infty} \sqrt{\frac{n^2+1}{n^3}} = \sqrt{0} = 0$, since square root is a continuous function. Moreover, a_n is decreasing since

$$a_{n+1} = \sqrt{\frac{(n+1)^2+1}{(n+1)^3}} = \sqrt{\frac{1}{n+1} + \frac{1}{(n+1)^3}} < \sqrt{\frac{1}{n} + \frac{1}{n^3}} = a_n.$$

Therefore, we can apply Leibniz criterion to conclude that the series converges. We can bound S by $0 \leq S_{2n} - S \leq a_{2n+1}$ and $0 \leq S - S_{2n-1} \leq a_{2n}$, that is, $|S - S_n| \leq a_{n+1} < \sqrt{\frac{1}{n} + \frac{1}{n^3}} < \sqrt{\frac{2}{n}}$. Thus $|S - S_n| \leq 10^{-2}$ occurs when $a_{n+1} < \sqrt{\frac{2}{n}} \leq 10^{-2}$. Hence, we require $N \geq 2 \cdot 10^4$.[11]

Orders of Infinity and Series

$\sqrt[n]{n^p} \to 1,\ p > 0$	Ex.: $\lim\limits_{n\to\infty} n^{20/n} = 1$
$\frac{n^p}{a^n} \to 0,\ a > 1, p > 0$	Ex.: $n^{200} = o(1.1^n)$
$\frac{n^p}{a^{\sqrt[q]{n}}} \to 0,\ a > 1, p > 0, q > 1$	Ex.: $n^{200} = o(1.1^{\sqrt[10]{n}})$
$\frac{\log_b n}{n^p} \to 0,\ p > 0, b > 1$	Ex.: $\ln(n) = o(\sqrt[3]{n})$
$\frac{a^n}{n!} \to 0 \quad a > 1$	Ex.: $100^n = o(n!)$
$\frac{n!}{n^n} \to 0$	Ex.: $\lim\limits_{n\to\infty} \frac{(n-3)^{n-5}}{(n+10)!} = 0$
$\sum\limits_{n=0}^{+\infty} q^n = \frac{1}{1-q}$ with $\lvert q\rvert < 1$	Ex.: $\sum\limits_{n=0}^{+\infty} 10^{-n} = \frac{10}{9}$
$\sum\limits_{n=0}^{+\infty} \frac{1}{n^\gamma}$ converges iff $\gamma > 1$	Ex.: $\sum\limits_{n=4}^{+\infty} \frac{1}{\sqrt[3]{n}-3}$ diverges
$\sum\limits_{n=0}^{+\infty} \frac{a^n}{n!} = e^a,\ a \in \mathbb{R}$	Ex.: $\sum\limits_{n=0}^{+\infty} \frac{2^n}{n!} = e^2$

Fundamental Limits

$\lim\limits_{x\to\pm\infty} \left(1 + \frac{a}{x}\right)^x = e^a$	
$\lim\limits_{x\to 0} (1 + ax)^{1/x} = e^a$	
$\lim\limits_{x\to 0} \frac{e^x - 1}{x} = 1$	$e^x = 1 + x + o(x)$
$\lim\limits_{x\to 0} \frac{a^x - 1}{x} = \ln a$	
$\lim\limits_{x\to 0} \frac{\ln(x+1)}{x} = 1$	$\ln(x+1) = x + o(x)$
$\lim\limits_{x\to 0} \frac{\log_a(1+x)}{x} = \frac{1}{\ln a}$	
$\lim\limits_{x\to 0} \frac{(1+x)^a - 1}{x} = a$	$(1+x)^a = 1 + ax + o(x)$
$\lim\limits_{x\to 0} \frac{\sin x}{x} = 1$	$\sin x = x + o(x),\ x \to 0$
$\lim\limits_{x\to 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}$	$\cos x = 1 - \frac{x^2}{2} + o(x^2)$
$\lim\limits_{x\to 0} \frac{\tan x}{x} = 1$	$\tan x = x + o(x)$

5 Continuity

In this brief section we will define what continuity for real valued functions is and we will state an important result, called **Transference principle**.

Definition 5.1 (Continuity). f is continuous at $x_0 \in dom(f)$ if

$$\forall\, \varepsilon > 0\ \exists\, \delta > 0 : x \in dom(f), |x - x_0| \leq \delta \implies |f(x) - f(x_0)| \leq \varepsilon.$$

The same definition via neighbourhoods would be: for every neighbourhood $\mathcal{Y} = I_\varepsilon(f(x_0))$ of $f(x_0)$ there exists a neighbourhood $\mathcal{X} = I_\delta(x_0)$ of x_0 such that $f(x) \in \mathcal{Y}$ for every $x \in dom(f) \cap \mathcal{X}$.

[11] The smallest solution would be $N = 2500$.

Remark 5.2. The function is continuous if it is continuous at every point of its domain.

Remark 5.3. Continuity is a local property: to verify it at x_0, it suffices to work within a sufficiently small neighbourhood of x_0.

Remark 5.4. Any function is continuous in an isolated point of its domain, where there's nothing to check.

Remark 5.5. we will later see that a more compact way to state continuity of f at x_0, which only holds if x_0 is an accumulation point of $dom(f)$, is $\lim_{x \to x_0} f(x) = f(x_0)$.

Definition 5.6 (Uniform Continuity). $f : I \subseteq \mathbb{R} \to \mathbb{R}$ is uniformly continuous over I if

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in I, |x - y| \leq \delta \implies |f(x) - f(y)| \leq \varepsilon. \tag{5.1}$$

Compare this definition with the usual (weaker) continuity:

$$\forall x \in I, \forall \varepsilon > 0, \exists \delta > 0, \forall y \in I : |x - y| \leq \delta \implies |f(x) - f(y)| \leq \varepsilon.$$

A uniformly continuous function over I is always continuous over I. If I is a compact set, that is, a closed and bounded interval, then the converse is also true:

Theorem 5.7 (Heine). *A continuous function in a closed bounded interval is uniformly continuous.*[12]

Definition 5.8 (Lipschitz Continuity). A function is L-Lipschitz if there exists a constant $L \in \mathbb{R}$ such that, for every $x_1, x_2 \in dom(f)$,

$$|f(x_1) - f(x_2)| \leq L|x_1 - x_2|.$$

Lipschitz is a strong version of continuity:

Theorem 5.9. *Lipschitz functions are continuous.*

Proof.
We will prove that f is continuous at x_0. Since f is Lipschitz, for every $x_0, x \in dom(f)$, $|f(x) - f(x_0)| \leq L|x - x_0| \overset{?}{<} \varepsilon$. If we can find δ_ε such that ? holds for every $\varepsilon > 0$, we would conclude that f is continuous. But ? is equivalent to $|x - x_0| \overset{?}{<} \frac{\varepsilon}{L}$. Therefore, a correct choice is $\delta := \frac{\varepsilon}{L}$: for every $\varepsilon > 0$, if $|x - x_0| < \frac{\varepsilon}{L} = \delta$, then $|f(x) - f(x_0)| \leq L|x - x_0| < \varepsilon$, concluding the proof. $\square$

Example 5.10. $\sin(x)$ is 1-Lipschitz in $\mathbb{R}$. x^2 is Lipschitz in $[0, 100]$ but not in $\mathbb{R}$. In fact, $|x^2 - y^2| = |x + y||x - y| \leq 200|x - y|$. $\sqrt{x}$ is not Lipschitz in $(0, 1)$, but it's Lipschitz in $[1, +\infty)$.

We now state a highly important theorem known as *transference principle* that has been asked consistently in past oral exams, converse implication included.

Theorem 5.11 (Transference Principle). *A function $f : \mathbb{R} \supseteq dom(f) \to \mathbb{R}$ is continuous at $x_0 \in dom(f)$ if and only if for every sequence $(a_n)_{n \in \mathbb{N}}$ of points* ***in its domain*** *converging to x_0, then the sequence $b_n := f(a_n)$ converges to $f(x_0)$. In other words,*

$$\lim_{n \to \infty} f(a_n) = f\left(\lim_{n \to \infty} a_n\right).$$

Proof.
Assume f continuous at $x_0 \in dom(f)$. Take $(a_n)_{n \in \mathbb{N}}$ with $\lim_{n \to \infty} a_n = x_0$. Fixing $\varepsilon > 0$, we want to prove that $|f(a_n) - f(x_0)| \leq \varepsilon$ eventually.
The proof consists of combining together the definitions of continuity (C) and limit (L):

$$\forall \varepsilon > 0 \, \exists \delta > 0 : |x - x_0| \leq \delta \implies |f(x) - f(x_0)| \leq \varepsilon \tag{C}$$

$$\forall \delta > 0 \, \exists N \in \mathbb{N} : n > N \implies |a_n - x_0| \leq \delta. \tag{L}$$

Applying (C) with $x := a_n$ (since $a_n \in dom(f)$) from (L) we deduce

$$\forall \varepsilon > 0 \, \exists N \in \mathbb{N} : n > N \implies |f(a_n) - f(x_0)| \leq \varepsilon.$$

[12] This theorem, an application of the Bolzano-Weierstrass Theorem, allows you to prove that a continuous function is Riemann-Integrable, even with a finite number of discontinuities.

The converse implication can be proved by contrapositive. We assume that f is not continuous and we show sequence $(a_n)_{n\in\mathbb{N}}$ such that $\lim\limits_{n\to\infty} a_n = x_0$ but $\lim\limits_{n\to\infty} f(a_n) \neq f(x_0)$. We negate the continuity:

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in dom(f) : |x - x_0| \leq \delta \implies |f(x) - f(x_0)| \leq \varepsilon$$

Becomes

$$\exists \varepsilon > 0, \forall \delta > 0, \exists x_\delta \in dom(f) : |x_\delta - x_0| \leq \delta \text{ and } |f(x_\delta) - f(x_0)| > \varepsilon.$$

Therefore, take $\delta = \frac{1}{n}$ and we can find a sequence $x_n \in dom(f)$ such that $|x_n - x_0| \leq \frac{1}{n}$ and a value $\varepsilon > 0$ such that $|f(x_n) - f(x_0)| > \varepsilon$, with ε not depending on n. Therefore $\lim\limits_{n\to\infty} f(x_n) \neq f(x_0)$. □

Remark 5.12. In future courses, where continuity will become harder and harder to define, the transference principle will become the actual *definition* of continuity. For instance, in analysis 2 it is used to check continuity in $\mathbb{R}^d$.

Proposition 5.13 (Stability of continuity). *Continuity is stable under product, quotient and linear combination, that is, if f and g are continuous and defined in a neighbourhood of x_0, then $\alpha f + \beta g$, $f \cdot g$ and f/g (here we assume $g \neq 0$ locally) are there continuous.*

Proof.

Take an arbitrary sequence $(a_n)_{n\in\mathbb{N}}$ that converges to x_0. To prove all the properties at once, define $h(x) := H(f(x), g(x))$, where H denotes the desired operation between the two functions (either the linear combination, the product, or the quotient). We want to prove that $h(a_n) \to h(x_0)$. The transference principle implies

$$f(a_n) \to f(x_0) \quad \text{and} \quad g(a_n) \to g(x_0).$$

By applying the corresponding property of limits, we obtain

$$h(a_n) = H(f(a_n), g(a_n)) \to H(f(x_0), g(x_0)) = h(x_0).$$

□

Proposition 5.14 (Continuity and compositions). *Given $h(x) = g(f(x))$ well defined locally with respect to x_0, then if f is continuous at x_0 and g is continuous at $f(x_0)$, then h is continuous at x_0.*

Proof.

Take a sequence $(a_n)_{n\in\mathbb{N}} \subseteq dom(f)$ converging to x_0. Continuity of f implies $f(a_n) \to f(x_0)$, by Theorem 5.11. Let $b_n := f(a_n)$ and $b := f(x_0)$. Then $b_n \to b$. Continuity of g at $b = f(x_0)$ implies

$$(g \circ f)(a_n) = g(f(a_n)) = g(b_n) \to g(b) = g(f(x_0)) = (g \circ f)(x_0),$$

by Theorem 5.11. This concludes the proof, once again by applying Theorem 5.11 to the function $g \circ f$. □

Remark 5.15. The theorem cannot be reversed. If $f(x) = |x|$ and $g(x) = \text{sign}(x)$, then $f(g(x)) = 1$ is continuous, but g is not continuous.

6 Limits of Functions

6.1 Definition of limit

Definition 6.1 (Extension of f to x_0). Given f and x_0 an accumulation point of $dom(f)$, we define the *natural* extension of f to x_0 as the <u>continuous</u> (*at* x_0) function

$$f^\star(x) := \begin{cases} f(x) & \text{if } x \in dom(f) \setminus \{x_0\}, \\ \ell = \lim_{x\to x_0} f(x) & \text{if } x = x_0 \end{cases}$$

The definition of limit ℓ should be such that it's possible to define $f^\star(x)$ as a continuous function <u>at x_0</u>.

Definition 6.2 (General definition of limit). Given a function $f : dom(f) \to \mathbb{R}$, $x_0 \in \overline{\mathbb{R}}$ an accumulation point of $dom(f)$ and $\ell \in \overline{\mathbb{R}}$, we write

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

if for every neighbourhood $\mathcal{Y} = I(\ell)$ there exists a pointed neighbourhood $\mathcal{X} = \dot{I}(x_0)$ such that $f(x) \in \mathcal{Y}$ for every $x \in \mathcal{X} \cap dom(f)$.[13]

Given such definition, we can easily recover the following:

i) *Finite limit as* $x \to \pm\infty$: the function $f(x)$ converges to $\ell \in \mathbb{R}$ as $x \to +\infty$ (resp. $-\infty$) if

$$\forall\, \varepsilon > 0, \exists B > 0 : x \geq B \implies |f(x) - \ell| \leq \varepsilon, x \in dom(f) \qquad \text{(resp. } x \leq -B\text{)}.$$

ii) *Finite limit as* $x \to x_0$*:* the function $f(x)$ converges to $\ell \in \mathbb{R}$ as $x \to x_0$ if

$$\forall\, \varepsilon > 0, \exists \delta > 0 : 0 < |x - x_0| \leq \delta \implies |f(x) - \ell| \leq \varepsilon, x \in dom(f)$$

iii) *Infinite limit as* $x \to \pm\infty$*:* the function $f(x)$ diverges to $-\infty$ as $x \to +\infty$ if[14]

$$\forall\, M > 0, \exists\, B > 0 : x \geq B \implies f(x) < -M, x \in dom(f).$$

iv) *Infinite limit as* $x \to x_0$*:* the function $f(x)$ diverges to $\pm\infty$ as $x \to x_0$ if

$$\forall\, M > 0\ \exists\, \delta > 0 : 0 < |x - x_0| < \delta \implies f(x) \lesseqgtr \pm M.$$

Definition 6.3 (Right Limit). Given a function $f : dom(f) \to \mathbb{R}$ and x_0 an accumulation point of $dom(f) \cap (x_0, +\infty)$, we call g the restriction of f to $dom(g) := dom(f) \cap (x_0, +\infty)$ and we set $\lim\limits_{x \to x_0+} f(x) := \lim\limits_{x \to x_0} g(x)$.

Remark 6.4. The definition is similar to the standard one, just change the neighbourhood: not $-\delta < x - x_0 < \delta$ but $0 < x - x_0 < \delta$.

Remark 6.5. For left limits $\lim\limits_{x \to x_0-} f(x) := \lim\limits_{x \to x_0} g(x)$ with $dom(g) := dom(f) \cap (-\infty, x_0)$. The neighbourhood in this case is $-\delta < x - x_0 < 0$.

Lemma 6.6. $\lim_{x \to x_0} f(x) = \ell$ *exists if and only if both right and left limits exist and coincide:* $\lim_{x \to x_0+} f(x) = \lim_{x \to x_0-} f(x) = \ell$*.*

Definition 6.7 (Continuity via limits). Given a function f and $x_0 \in dom(f)$ an accumulation point of $dom(f)$, then f is continuous at x_0 if and only if $\lim_{x \to x_0} f(x) = f(x_0)$.[15]

Remark 6.8. If you consider x_0 as a right/left accumulation point, then f is continuous at x_0 from the right/left if and only if $\lim_{x \to x_0\pm} f(x) = f(x_0)$, (choosing the right sign)

6.2 Main results with limits of functions

Here, whenever we say "locally" it means in a pointed neighbourhood (no matter how small) of $x_0 \in \overline{\mathbb{R}}$. We take for granted that calculus rules on limits of functions are the same as those for limits of sequences.

i) **Uniqueness:** If the limit is well-defined and exists, then it's unique.

ii) **Restriction:** If $f : dom(f) \to \mathbb{R}$, $E \subset dom(f)$, and x_0 is also an accumulation point of E, then if $\lim\limits_{x \to x_0} f(x) = \ell$, it follows that $\lim\limits_{E \ni x \to x_0} f(x) = \ell$.[16]

iii) **Locality:** If f and g locally coincide, they will both have or not have limit at x_0, and if they have it, the limits coincide.

[13] Recall that $\overline{\mathbb{R}}$ is the extended real line, explained at the beginning of these notes. And notice that $\mathcal{X}$ is a pointed neighbourhood, meaning that we don't care at all what happens exactly at x_0: we only care what happens in its proximity.

[14] Pay attention to choose the correct inequality thinking about the sign of infinity, here I've shown just an example.

[15] As previously mentioned, this theorem is tautological but allows to define continuity easier. However, it fails to define continuity for points which are not accumulation points of the domain.

[16] This means that you can consider only a subset of x values approaching x_0. For example, you can focus on x values from the left, resulting in the left-hand limit.

iv) **Local boundedness:** if $\lim_{x\to x_0} f(x) = l \in \mathbb{R}$, then f is locally bounded around x_0.

v) **Permanence of inequalities 1:** If $\lim_{x\to x_0} f(x) < \lim_{x\to x_0} g(x)$, then f and g locally satisfy $f(x) < g(x)$ for every $x \in dom(f) \cap dom(g) \setminus \{x_0\}$. In particular, if $\lim_{x\to x_0} f(x) > 0$, then f is locally positive.

vi) **Permanence of inequalities 2:**[17] If $f(x) \leq g(x)$ locally and the limits of both f and g exist, possibly being infinite, then $\lim_{x\to x_0} f(x) \leq \lim_{x\to x_0} g(x)$. Specifically, if $f(x) \geq 0$ locally, then $\lim_{x\to x_0} f(x) \geq 0$, assuming the limit exists.

vii) **Squeeze Theorem:** If $f(x) \leq g(x) \leq h(x)$ locally and $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} h(x) = l \in \overline{\mathbb{R}}$, then $\lim_{x\to x_0} g(x) = l$.

Theorem 6.9 (Limits of monotone functions 1). *If x_0 is an accumulation point of $dom(f) = D$ from the left (similarly from the right) then*

$$\lim_{x\to x_0-} f(x) = \begin{cases} \sup\limits_{x\in D\cap(-\infty,x_0)} f(x) & \text{if } f \text{ is increasing} \\ \inf\limits_{x\in D\cap(-\infty,x_0)} f(x) & \text{if } f \text{ is decreasing} \end{cases}.$$

[proof]

Theorem 6.10 (Limits of monotone functions 2). *Let f be a monotonically increasing function and $x_0 \in dom(f)$ be an accumulation point both from the left and from the right of $dom(f)$. Then both left and right limits exist, are finite and*

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

A dual result holds if f is decreasing. As a consequence, ***monotone functions may have only jump discontinuities****.*

[proof]

Theorem 6.11 (Transference Principle (2)). *Given a function f and an accumulation point x_0 of $dom(f)$, we have $\lim_{x\to x_0} f(x) = \ell$ if and only if, for every sequence $(a_n)_{n\in\mathbb{N}}$ of points of its domain, excluding x_0, which converges to x_0, we have $\lim_{n\to\infty} f(a_n) = \ell$.*

Proof.
Let us define the extension $f^\star(x)$ of f as

$$f^\star(x) = \begin{cases} f(x) & \text{if } x \in \text{dom}(f) \setminus \{x_0\} \\ \ell & \text{if } x = x_0 \end{cases} \tag{$\star$}$$

By definition, if $\lim_{x\to x_0} f(x) = \ell$, then $f^\star(x)$ is continuous. Therefore, from the transference principle applied on $f^\star(x)$, we find a sequence $(a_n)_{n\in\mathbb{N}} \in dom(f) \setminus \{x_0\}$ such that $f^\star(a_n) \to \ell$. In particular, $f(a_n) \to \ell$. As for the converse, by the transference principle, if for every sequence $(a_n)_{n\in\mathbb{N}} \in dom(f) \setminus \{x_0\}$ converging to x_0 we have $f(a_n) \to \ell$, then $f^\star(x)$ defined as in $(\star)$ is continuous, which implies that $\lim_{x\to x_0} f^\star(x) = \ell = \lim_{x\to x_0} f(x)$. □

6.3 Digression: relationship between limits, continuity and composition

6.3.1 Transference Principle

In this section, which is more suggestive than mandatory, we adopt a new perspective to build an intuition on various concepts previously introduced, aiming to make them more coherent and logical, rather than viewing them as isolated results. The student is encouraged not to memorize every theorem as part of a list, but to link concepts to understand the "general philosophy" beneath them. The starting point of our discussion is

[17] Be cautious in applying these theorems. If $\lim_{x\to x_0} f(x) \leq \lim_{x\to x_0} g(x)$, it does not provide specific information regarding the ordering of the functions when the limits are equal. Likewise, if $f(x) < g(x)$, we cannot deduce that $\lim_{x\to x_0} f(x) < \lim_{x\to x_0} g(x)$, as limits do not maintain strict inequalities. For example, $\sin(x) < x$ for $x > 0$, but $\lim_{x\to 0+} \sin(x) = \lim_{x\to 0+} x$.

the definition of limit (we consider the finite case for simplicity). According to this definition, for a given ε, we examine if every point $x \in dom(f) \setminus \{x_0\}$ and sufficiently close to x_0 (at a distance less than δ) satisfies the condition $|f(x) - f(x_0)| < \varepsilon$. For instance, if $f(x) = x^2$ and $x_0 = 0$, we verify that for any $\varepsilon > 0$, each x satisfying $|x - 0| < \delta = \sqrt{\varepsilon}$ also satisfies $|f(x) - 0| < \varepsilon$.

The **restriction property** of limits indicates that if this condition holds for all such x, it naturally holds for any subset of these x. An interesting subset to consider is a countable one. For example, let $S = \{x \in \mathbb{R} : x = \sin(1/n), n \in \mathbb{N}\}$, a countable set of points $x_n = \sin(1/n)$ arbitrarily close to 0. Given that $\lim_{x\to 0} f(x) = 0$, we can find a δ small enough so that for $|x_n| < \delta$, $|f(x_n) - 0| < \varepsilon$. Choosing $\delta = \frac{1}{n}$ changes the task to finding a sufficiently large n such that $|f(x_n) - 0| < \varepsilon$, that is, $f(x_n) \to 0$ as $n \to \infty$.

This interpretation suggests that if $\lim_{x\to x_0} f(x) = \ell$, then $\lim_{n\to\infty} f(x_n) = \ell$ for any sequence $(x_n)_{n\in\mathbb{N}}$ of points in $dom(f) \setminus \{x_0\}$ converging to x_0. And that's almost the transference principle in its second form!!

However, the transference principle is somewhat stronger, as it also says the converse: if for every sequence $(x_n)_{n\in\mathbb{N}}$ of points converging to x_0, $f(x_n) \to \ell$ as $n \to \infty$, then also $\lim_{x\to x_0} f(x) = \ell$. The intuition behind this result is that even though we might only test convergence for a countable set of points, since the criterion is met for every possible sequence, no point in the set $\mathcal{X}$ escapes this check, ensuring that $f(x)$ gets close to ℓ for any $x \in \mathcal{X}$. Therefore, the **transference principle** in its second form is basically a strong form of the restriction property.

6.3.2 How to use the Transference Principle

This turns out to be really useful to prove that some limits do not exist. Consider the limit $\lim_{x\to+\infty} \sin x$, which does not exist. To show this, it suffices to identify two sequences, $a_n \to +\infty$ and $b_n \to +\infty$, where $\sin(a_n) \to \ell_1$ and $\sin(b_n) \to \ell_2 \neq \ell_1$. We can choose $a_n = \pi/2 + 2n\pi$ and $b_n = -\pi/2 + 2n\pi$. Another example is $\lim_{x\to+\infty} \sin(\log x + 1)x$. Observing the function's behaviour, we find it oscillates between reaching $+\infty$ and 0 (when $\sin(\log x) = -1 \iff \log x = -\pi/2 + 2n\pi \iff x = e^{-\pi/2+2n\pi}$). Thus, the sequences $a_n = e^{n\pi}$, where $f(a_n) = (\sin(n\pi) + 1)e^{n\pi} = e^{n\pi} \to +\infty$, and $b_n = e^{-\pi/2+2n\pi}$, where $f(b_n) = 0 \to 0$, are sufficient to show the limit's non-existence.

6.3.3 Substitution Theorem

Until now, we have applied the restriction property using a countable subset defined as the image of a sequence. This can also extend to subsets defined by the image of a function. Consider $S := \{y \in \mathbb{R} : y = \sin(1/x), x \in \mathbb{R}\}$, consisting of points arbitrarily close to zero. Let the function $g(x) = \frac{\sin x}{x}$. Following the same logic, because $\lim_{y\to 0} g(y) = 1$, it follows that:

$$\lim_{x\to+\infty} g(\sin(1/x)) = \lim_{y\to 0} g(y) = \lim_{y\to 0} \frac{\sin y}{y} = 1.$$

What we have just stated is the theorem for the limit of a composition of functions, which is basically what justifies the typical substitutions we always perform in solving various limits.

The assumptions are quite heavy and are necessary to well define the composition. Suppose $x_0, y_0 \in \overline{\mathbb{R}}$ are accumulation points of $dom(f)$ and $dom(g)$ respectively, such that $\lim_{x\to x_0} f(x) = y_0$, and the composition $g \circ f$ is well defined in a small pointed neighbourhood of x_0 in $dom(f)$. Then[18]

$$\boldsymbol{\lim_{y\to y_0} g(y) = \ell \qquad \Longrightarrow \qquad \lim_{x\to x_0} g(f(x)) = \ell.} \tag{6.1}$$

This theorem enables us to solve $\lim_{y\to y_0} g(y)$ instead of directly addressing $\lim_{x\to x_0} g(f(x))$. By finding the first, we infer that both limits must coincide. For instance, consider the limit $\lim_{x\to\pi/2} \frac{\sin(\cos(x))}{\cos(x)}$. Since $\lim_{y\to 0} \frac{\sin(y)}{y} = 1$ exists, we can perform the substitution $y = \cos(x)$ and transform the first limit into the second.[19]

[18] If we add the condition that f is locally invertible around x_0 and $\lim_{y\to y_0} f^{-1}(y) = x_0$, then $\lim_{x\to x_0} g(f(x))$ exists if and only if $\lim_{y\to y_0} g(y)$ exists and in such case (6.1) holds, meaning that the local invertibility allows to invert the theorem.

[19] For annoying technical reasons, we should require also $y_0 \notin dom(g)$, otherwise the functions $g(0) = 1$ and $g(x) = 0$ elsewhere, and $f(x) = 0$, break this theorem. In fact, $g(x) \to 0$ as $x \to 0$, but since $f(x) = 0$ everywhere, $f(g(x))$ is continuous and therefore $f(g(x)) \to f(g(0)) = 1 \neq \lim_{x\to 0} g(0)$.

6.3.4 Continuity

Recall that each time we compute a limit around x_0, we exclude from our analysis what happens at x_0. Continuity integrates this excluded value, asserting that the limit at x_0 matches the function's value at that point. Therefore, the transference principle, in its first form, can be seen as an application of the restriction property for limits, replacing $\lim_{x\to x_0} f(x)$ with $f(x_0)$ due to their equivalence under continuity. For a better understanding, let's compare the core statements of the two transference principles. The second form states that for any sequence a_n converging to x_0:

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

The first form, which assumes continuity at x_0, says

$$\lim_{n\to\infty} f(a_n) = f\left(\lim_{n\to\infty} a_n\right) = f(x_0) \overset{!}{=} \lim_{x\to x_0} f(x).$$

Adding the equality (!), which comes from the definition of continuity, it seems clear that the first form of the transference principle is equivalent to the second, except for the fact that continuity directly equates $\lim_{x\to x_0} f(x)$ to $f(x_0)$. With the same philosophy, it's now clear that, for continuous functions, (6.1) can be simplified as follows:

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

which, in some sense, is a transference principle with $f(x)$ instead of a_n.

6.4 Exercises

Example 1: fundamental limits

$$\begin{aligned}\lim_{x\to 0} \frac{\log(\sqrt[3]{1+\sin x})}{\arctan(2x)} \cdot \frac{x}{\sqrt{1+3\sin x}-1} &= \lim_{x\to 0} \frac{1}{3} \cdot \underbrace{\frac{\log(1+\sin x)}{\sin x}}_{1} \cdot \underbrace{\frac{\sin x}{x}}_{1} \cdot \underbrace{\frac{2x}{\arctan(2x)}}_{1} \cdot \frac{1}{2} \cdot \frac{x}{\sqrt{1+3\sin x}-1} \\ &= \frac{1}{6} \lim_{x\to 0} \frac{x}{\sqrt{1+3\sin x}-1} \cdot \frac{\sqrt{1+3\sin x}+1}{\sqrt{1+3\sin x}+1} = \frac{1}{6} \lim_{x\to 0} \frac{2x}{(1+3\sin x)-1} \\ &= \frac{1}{9} \lim_{x\to 0} \frac{x}{\sin x} = \frac{1}{9}.\end{aligned}$$

Example 2: fundamental limits

$$\begin{aligned}\lim_{x\to+\infty} (1-\cos(1/x)) \ln(x^2+e^x) &= \lim_{x\to+\infty} \frac{1}{2x^2} \ln\left(e^x\left(1+\frac{x^2}{e^x}\right)\right) \\ &= \lim_{x\to+\infty} \frac{1}{2x^2} \left(\log(e^x) + \log\left(1+\frac{x^2}{e^x}\right)\right) = \lim_{x\to+\infty} \frac{1}{2x^2}\left(x+\frac{x^2}{e^x}\right) = 0.\end{aligned}$$

Example 3: logarithm

$$\lim_{x\to 0+} (x+x^2)\log(x) = \lim_{x\to 0+} x(1+x)\log(x) = \lim_{x\to 0+} x(1+0^+)\log(x) = \lim_{x\to 0+} x\log(x) \overset{y=1/x}{=} \lim_{y\to+\infty} \frac{-\log(y)}{y} = 0.$$

Example 4: exponential trick

$$\begin{aligned}\lim_{n\to+\infty} \left(\cos\left(1/\sqrt{n}\right)\right)^n &= \lim_{n\to+\infty} e^{n\log(\cos(1/\sqrt{n}))} = \exp\left(\lim_{n\to+\infty} n\log(\cos(1/\sqrt{n}))\right) \\ &= \exp\left(\lim_{n\to+\infty} n\log\left(1-\frac{1}{2n}\right)\right) = \exp\left(\lim_{n\to+\infty} -\frac{n}{2n}\right) = \sqrt{\frac{1}{e}}.\end{aligned}$$

Example 5: exponential trick

$$\lim_{n\to+\infty} n(\sqrt[n]{n}-1) = \lim_{n\to+\infty} \frac{n^{1/n}-1}{1/n} = \lim_{n\to+\infty} \frac{e^{\log n/n}-1}{1/n} = \lim_{n\to+\infty} \underbrace{\frac{e^{\log n/n}-1}{\log n/n}}_{1} \cdot \log n = +\infty.$$

Example 6: exponential trick

$$\lim_{x\to 0}(\cos x)^{1/\tan^2 x}\cdot(1+x)^{1/\tan x} = \lim_{x\to 0} e^{\log(\cos x)/\tan^2 x}\cdot e^{\log(1+x)/\tan x} = \exp\left(\lim_{x\to 0}\frac{\log(\cos(x))}{\tan^2(x)}+\frac{\log(1+x)}{\tan x}\right)$$
$$= \exp\left(\lim_{x\to 0}\frac{\log(\cos(x))}{x^2}+1\right) = \exp\left(\lim_{x\to 0}\frac{\log(1-x^2/2)}{x^2/2}\cdot\frac{-1}{2}+1\right)$$
$$= e^{-1/2+1} = \sqrt{e}.$$

Example 7: change of variable

$$\lim_{x\to 2}\frac{\sqrt[3]{10-x}-2}{x-2} \overset{y=x-2}{=} \lim_{y\to 0}\frac{\sqrt[3]{8-y}-2}{y} = \lim_{y\to 0} 2\frac{\sqrt[3]{1-y/8}-1}{y} \overset{z=y/8}{=} \lim_{z\to 0}\frac{1}{4}\cdot\frac{\sqrt[3]{1-z}-1}{z} = \frac{1}{4}\left(-\frac{1}{3}\right) = -\frac{1}{12}.$$

Example 8: exponential limit

$$\lim_{x\to 3}\left(\frac{x}{3}\right)^{\frac{1}{x-3}} \overset{y=\frac{1}{x-3}}{=} \lim_{y\to +\infty}\left(1+\frac{1}{3y}\right)^{y} \overset{z=3y}{=} \lim_{z\to +\infty}\left(1+\frac{1}{z}\right)^{z/3} = e^{1/3}.$$

Example 9: squeeze theorem

$$\lim_{x\to 0}\frac{x-\tan x}{x^2} \geq \lim_{x\to 0}\frac{\sin x-\tan x}{x^2} = \lim_{x\to 0}\frac{\sin x(1-1/\cos x)}{x^2} = \lim_{x\to 0}\frac{\sin x(\cos x-1)}{x^2\cos x} = \lim_{x\to 0}-\frac{1}{2}\frac{\sin x}{\cos x} = 0.$$

Moreover, $x-\tan x<0$ in a right neighbourhood of 0. Thus, The right limit is zero. $\frac{x-\tan x}{x^2}$ is odd, the left limit must also be zero.

7 Continuity on Intervals

Continuity, while interesting as a local property at a point x_0, especially due to the transference principle that links it closely with limits of sequences, becomes a fundamental concept when extended to an entire interval. In fact, it serves as the foundation for two highly significant theorems: **Bolzano's theorem**, which ensures that the image of a continuous function has no gaps, and **Weierstrass's theorem**, which guarantees the existence of a minimum and maximum for continuous functions on a closed and bounded interval.

7.1 Bolzano's Theorem and its corollaries

Bolzano's theorem, also known as **intermediate value theorem**, states that a continuous function $f:[a,b]\to\mathbb{R}$ attains all real values between $f(a)$ and $f(b)$.

Theorem 7.1 (Bolzano). *If $f:[a,b]\to\mathbb{R}$ is continuous, then for every $y^\star\in[f(a),f(b)]$ there exists $x^\star\in[a,b]$ such that $f(x^\star)=y^\star$.*

Proof.

Step 1: we can restrict our analysis to the case $f(a)<0<f(b)$. This is because $x^\star$ defined as above exists if and only if there exists $x^\star$ such that $g(x^\star)=0$, where $g(x):=f(x)-y^\star$. Indeed, $g(x^\star)=0=f(x^\star)-y^\star \implies f(x^\star)=y^\star$. Since $g:[a,b]\to\mathbb{R}$ is continuous and $g(a)<0<g(b)$, if we prove the theorem in this case, we can recover the original statement via translation as just shown.

Step 2: We define two sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ such that $a_0=a$ and $b_0=b$, and an auxiliary sequence $c_n:=\frac{1}{2}(a_{n-1}+b_{n-1})$. Then:

$$\begin{cases} a_n=a_{n-1} \text{ and } b_n=c_n & \text{if } f(c_n)>0\\ a_n=c_n \text{ and } b_n=b_{n-1} & \text{if } f(c_n)<0\\ x^\star=c_n & \text{if } f(c_n)=0\end{cases}$$

Step 3: If $\exists\, n\in\mathbb{N}$ such that $f(c_n)=0$, it's clear that the theorem is true since we found a desiring value $c_n=x^\star\in[a,b]$, and we conclude the algorithm. Otherwise, suppose that the algorithm doesn't stop after a finite number of steps. These sequences satisfy the assumptions of the Cantor's Theorem:

i) a_n is increasing and b_n is decreasing by construction;

ii) The length of $[a_n, b_n]$ is infinitesimal because $b_n - a_n = \dfrac{b_{n-1} - a_{n-1}}{2}$. Then it's easy to show by induction that

$$b_n - a_n = \frac{b_0 - a_0}{2^n} = \frac{b-a}{2^n} \to 0 \quad \text{as } n \to +\infty.$$

iii) $a_n < b_n$ by construction.

Step 4: Therefore, by the Cantor's Property, the two sequences converge to the same limit $x^\star$. Continuity of f in $[a,b]$ yields

$$\lim_{n\to\infty} f(a_n) = f(x^\star) = \lim_{n\to\infty} f(b_n),$$

by Theorem 5.11 applied to both $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$.
Step 5: Notice that $f(a_n) < 0$ and $f(b_n) > 0$ by construction, since the choice of the update preserves this property at every step. Therefore, by the permanence of sign through limits, $f(x^\star) = \lim_{n\to\infty} f(a_n) \leq 0$ and $f(x^\star) = \lim_{n\to\infty} f(b_n) \geq 0$, therefore $f(x^\star) = 0$. □

Corollary 7.2 (Existence of zeros). *If $f : [a,b] \to \mathbb{R}$ is continuous and $f(a) \cdot f(b) < 0$, then there exists $x^\star \in [a,b]$ such that $f(x^\star) = 0$.*

Corollary 7.3 (intersection of continuous functions:). *If $f, g : [a,b] \to \mathbb{R}$ are continuous such that $f(a) < g(a)$ and $f(b) > g(b)$, then they intersect at least once.*

Corollary 7.4 (Fixed points). *Every continuous function $f : [a,b] \to [a,b]$ has a fixed point, i.e. a point $x^\star \in [a,b]$ such that $f(x^\star) = x^\star$.*[20]

Corollary 7.5 (Image of continuous functions). *Continuous functions map intervals into intervals.*[21] *More explicitly, if f is defined on an interval, possibly unbounded, then f takes all the values in $[\inf f, \sup f]$.*

Proof.
Since $\inf f$ is a lower bound, we can find an element $a \in I$ such that $f(a) \leq y$. If $f(a) = y$ we are done; otherwise, suppose $\boldsymbol{f(a) < y}$. Similarly, we can find an element $b \in I$ such that $\boldsymbol{f(b) > y}$. Suppose $a < b$ (otherwise we swap the roles of a and b). Since $f(a) < y < f(b)$, with $[a,b] \subset I$ and f continuous, we can apply the **intermediate value theorem**, obtaining $x^\star \in [a,b] \subset I$ such that $\boldsymbol{f(x^\star) = y}$. **[Check]** □

Corollary 7.6 (Compactness through continuity). *Continuous functions map compact intervals into compact intervals.*

Remark 7.7. Any other version is false in general. For instance, $f(x) = x^2$ maps the open interval $(-1,1)$ into the non-open interval $[0,1)$, and the map $f(x) = 1/x$ maps the bounded interval $(0,1)$ into the unbounded interval $(0,+\infty)$.

7.2 Monotone functions and continuity

Corollary 7.5 shows that continuous functions map intervals into intervals. However, even discontinuous functions, like $f(x) = x+1$ for $x \in [-1,0]$ and $f(x) = x-1$ for $x \in (0,1]$, can map intervals into intervals. Which condition should we add to be sure that the previous corollary can be inverted? The answer is monotonicity. If f is monotone, an equivalent definition of continuity is mapping intervals into intervals. But, before that, let's remark the relationship between monotonicity and injectivity.

Theorem 7.8 (Injectivity and monotonicity). *A continuous function is injective (therefore invertible) if and only if is strictly monotone.*

[20] A simple algorithm to find a fixed point in such case is to take $a_0 \in [a,b]$ and then define the recursion $a_n := f(a_{n-1})$. This algorithm may fail depending on a_0, but if a_n converges, then it converges to a fixed point (you can prove this result with the transference principle). A sufficient condition on f to be sure that the algorithm converges, and thus outputs a fixed point, is to have f monotone.

[21] Combining this result with the Weierstrass' theorem, we can conclude that continuous functions map closed and bounded intervals into closed and bounded intervals.

Proof.

If f is strictly monotone and continuous, clearly f is injective. To prove the converse, we prove the contrapositive. If it's not strictly monotone, then it's not injective. If it's not strictly monotone and there are two point a, b for which $f(a) = f(b)$, we've finished. The 4 cases remaining are that there are 3 points $a < b < c$ such that $f(a) < f(c) < f(b)$ or $f(c) < f(a) < f(b)$ or $f(b) < f(c) < f(a)$ or $f(b) < f(a) < f(c)$.[a]. In each of these 4 cases we can apply Theorem 7.1 to conclude. For instance, in the first case we can prove the existence of $d \in [a, b]$ such that $f(d) = f(c)$. $\square$

[a]This reasoning is less trivial than it may seem...

Theorem 7.9 (Continuity of monotone functions). *Given a monotone function f on an interval I, f is continuous if and only if its image $J = f(I)$ is an interval.*

Proof.

If $f : I \to \mathrm{Im}f$ is continuous, then $\mathrm{Im}f$ is an interval by Corollary 7.5. Let us prove the converse by contrapositive. Assume f not continuous at $x_0 \in I$.
Case 1: x_0 lies in the interior of I. For simplicity, let us assume that f is increasing. The decreasing case is analogous. Theorem 6.10 guarantees that f may only have jump discontinuities. Hence we set $L_- := \lim_{x\to x_0-} f(x)$ and $L_+ := \lim_{x\to x_0+} f(x)$, with $-\infty < L_- < L_+ < +\infty$. Hence, infinitely many points in (L_-, L_+), with the possible exception of x_0, do not belong to $f(I)$, so it cannot be an interval.
Case 2: x_0 is not an interior point. For simplicity, let us assume that $x_0 = \min I$ (the other case is analogous). Since f is increasing, we have $f(x_0) < \lim_{x\to x_0+} f(x) = L+ \in J$, so J does not contain the points between $f(x_0)$ and $L+$. $\square$

Theorem 7.10 (Continuity of the inverse). *The inverse function $f^{-1} : J \to I$ of a continuous function $f : I \to J = f(I)$ is continuous (and strictly monotone).*[22]

Proof.

We consider the case when f is increasing: the other case is analogous. Since f is continuous, $J = f(I)$ is an interval. Since f is strictly increasing, f is injective. Therefore $f^{-1}(y)$ is a monotone function and it's defined $\forall\, y \in J$. Since its image $f^{-1}(J) = I$ is an interval, by the previous theorem $f^{-1}(y)$ is continuous. $\square$

7.3 Weierstrass' Theorem

Weierstrass's theorem is considered one of the most important theorems in mathematical analysis. It states that a continuous function on a **closed** and **bounded** interval attains both its minimum and maximum within that interval.

Lemma 7.11. *If $f : D \to \mathbb{R}$, there exists a sequence $(a_n)_{n\in\mathbb{N}}$ of points of D such that $f(a_n)$ converges to* $\inf_{x\in D} f(x)$.[23]

Proof.

Set $m := \inf_{x\in D} f(x)$. If m is finite, by definition of infimum, $m \le f(x)\ \forall x \in D$ and for every $\varepsilon > 0$ there exists $x_\varepsilon \in D$ such that $f(x_\varepsilon) < m + \varepsilon$. Setting $\varepsilon = 1/n$ we obtain a sequence a_n depending on n such that $m \le f(a_n) < m + 1/n$, i.e. $f(a_n)$ converges to m (by the squeeze theorem). If $m = -\infty$, then $\forall\, n \in \mathbb{N}^+\ \exists\, a_n \in D$ such that $f(a_n) < -n$. By comparison test, $f(a_n)$ diverges to $-\infty$. $\square$

Theorem 7.12 (Weierstrass). *A continuous function $f : [a, b] \to \mathbb{R}$ is bounded over $[a, b]$ and there exists x_{min} and x_{max} in $[a, b]$ such that $f(x_{min}) = \inf_{x\in[a,b]} f(x)$ and $f(x_{max}) = \sup_{x\in[a,b]} f(x)$.*[24]

[22]Here, we are implicitly assuming that f is strictly monotone, in order to have injectivity. This theorem guarantees that the n-th root and functions such as $\arcsin(x)$ and $\arccos(x)$ are continuous. Just pay attention to restrict f to an interval I where it's strictly monotone.
[23]In this proof we don't require that a_n are distinct points of the domain of f. Actually, if $f(x) = 1$ for $x \ne 0$ and $f(0) = 0$, then a sequence may be $a_n := 0$.
[24]As an exercise, show with counterexamples that all the three assumptions are necessary. **[Do it]**

Proof.

We only prove that $\inf_{x\in[a,b]} f \in f([a,b])$.[a] By Lemma 7.11 there exists a sequence $(a_n)_{n\in\mathbb{N}} \in [a,b]$ such that $f(a_n) \to \inf_{x\in[a,b]} f(x)$. Since $(a_n)_{n\in\mathbb{N}} \in [a,b]$, by **Bolzano-Weierstrass** Theorem 3.4 there exists a subsequence $(a'_k)_{n\in\mathbb{N}}$ which converges to some limit $\alpha = x_{\min}$. Since $[a,b]$ is closed and bounded, $x_{\min} \in [a,b]$.[b]

Since f is continuous and $x_{\min} \in [a,b]$, we can apply the **transference principle** 5.11 to the subsequence, that is

$$f(x_{\min}) = f\left(\lim_{k\to\infty} a'_k\right) \overset{1}{=} \lim_{n\to\infty} f(a'_k) \overset{2}{=} \lim_{n\to\infty} f(a_n) = \inf_{x\in[a,b]} f(x),$$

where (1) is the transference principle and (2) is the fact that if a sequence (in this case $f(a_n)$) converges to a limit ($\inf f$), then any of its subsequences (in this case $f(a'_k)$) converges to the same limit. □

[a]To prove the existence of the maximum, notice that $\sup(f) = -\inf(-f)$, so you can prove that $\inf -f \in -f([-b,-a])$ and then deduce that $\sup f \in f([a,b])$.

[b]To be more specific, $a \le a'_k \le b$ for every $k \in \mathbb{N}$, and by the permanence of inequalities in limits we have $a \le \alpha \le b$. This passage should be modified when we prove the extensions of Weierstrass (it's been asked as an exercise during orals). For example, if the interval is $(-\infty,+\infty)$ and we know that $\lim_{x\to\pm\infty} f(x) > \inf_{x\in\mathbb{R}} f(x)$, we should say that these two limits imply the existence of points $a,b \in \mathbb{R}$ such that $f(x) > \inf f + \varepsilon$ for $f < a$ and $f > b$ and some $\varepsilon > 0$. In this way, we conclude that $(a'_k)_{n\in\mathbb{N}}$ eventually belongs to $[a,b]$. And now on we can continue as before.

Remark 7.13 (Extensions of Weierstrass). Given a continuous function f in an interval $I = [a,b)$ opened from the right (possibly unbounded), 1) If $\boldsymbol{\lim_{x\to b-} f(x) > \inf f}$ then f attains its **minimum** in I. 2) If $\boldsymbol{\lim_{x\to b-} f(x) < \sup f}$ then f attains its **maximum** in I. The theorem can be re-adapted for different intervals, as shown in the examples.

Remark 7.14. Usually you don't know anything about inf and sup. So the common sufficient condition is that $\boldsymbol{\lim_{x\to b-} f(x) = +\infty} \implies$ exists minimum, and $\boldsymbol{\lim_{x\to b-} f(x) = -\infty} \implies$ exists maximum

Example 7.15. To prove that $\ln x$ has a maximum in $(0,1]$ you can say that $\lim_{x\to 0+} f(x) = -\infty$ and conclude.

To prove that $f(x) = x^4 - 2x^3 + 3x^2 + 1$ has a minimum you can say that $\lim_{x\to-\infty} f(x) = +\infty$ and $\lim_{x\to+\infty} f(x) = +\infty$ and conclude (in this case both right and left limits since $(-\infty,+\infty)$ is opened both fromt he right and from the left.

To prove that xe^{-x} has a maximum in $(0,+\infty)$ you can say that $\lim_{x\to 0-} f(x) = 0$ and $\lim_{x\to+\infty} f(x) = 0$ and $\sup f \ge f(1) > 0$ and conclude.

Exercise: Let $f : (0,1] \to \mathbb{R}$ be a continuous function such that $f(1) = 1$ and $\lim_{x\to 0+} xf(x) = -\infty$. Prove that the equation

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

has at least one solution in the interval $(0,1)$.

Solution: the equation is equivalent to $xf^2(x) - x - 1 = 0$ and $f(x) \ne \pm 1$. Let $g(x) := xf^2(x) - x - 1$. g is a continuous function, since product and sum of continuous functions. Note that $g(1) = -1$ and

$$\lim_{x\to 0+} g(x) = \lim_{x\to 0+} xf^2(x) - x - 1 = \lim_{x\to 0+} xf(x) \cdot \frac{xf(x)}{x} - 1 = (-\infty)\cdot\frac{-\infty}{0^+} - 1 = +\infty.$$

In particular, there exists $\delta \in (0,1)$ such that $g(x) \ge 100$ for every $x \in (0,\delta]$. Now $g(\delta) \ge 100 > 0$ and $g(1) = -1 < 0$. Thus we can apply Bolzano's Theorem on $[\delta,1]$ finding $\overline{x} \in (\delta,1)$ such that $g(\overline{x}) = 0$. Suppose by contradiction that $f(\overline{x}) = \pm 1$. Then $g(\overline{1}) = -1 \ne 0$. Therefore, $\overline{x}$ is such that $g(\overline{x}) = 0$ and $f(\overline{x}) \ne \pm 1$, so it's a solution of the initial equation in the interval $(0,1)$.

8 Differential Calculus

There are two main goals behind the concept of derivative. The first is to study the slope of a function, that is, the ratio of its rate of change. The second is to locally approximate the function by an affine function, namely a linear function plus a constant term.

8.1 Derivative of real functions

Definition 8.1 (Linear Functions). A function $f : X \to Y$ between vector spaces on a field $\mathbb{F}$ is *linear* if $f(\alpha x_0 + \beta x_1) = \alpha f(x_0) + \beta f(x_1)$, for every $x_0, x_1 \in X$, $\alpha, \beta \in \mathbb{F}$.
If $f : \mathbb{R} \to \mathbb{R}$, linear functions attain the functional form $\ell(x) := px$, where $p \in \mathbb{R}$ is called slope of ℓ.

Definition 8.2 (Affine Functions). A function $f : X \to Y$ between vector spaces on a field $\mathbb{F}$ is *affine* if $f(\alpha x_0 + \beta x_1) = \alpha f(x_0) + \beta f(x_1)$ for every $x_0, x_1 \in X$, $\alpha, \beta \in \mathbb{F}$ and $\alpha + \beta = 1$.
If $f : \mathbb{R} \to \mathbb{R}$, *affine* functions correspond to linear functions up to a constant term: $\ell(x) := px + q$.

Definition 8.3 (Chord). Given a real function f, the chord between the two points $P_0 = (x_0, y_0 = f(x_0))$ and $P_1 = (x_1, y_1 = f(x_1))$ is the affine function passing through them:

$$l_{P_0,P_1}(x) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0) + f(x_0).$$

The slope of the chord is the ratio between the increment of $f(x)$ and the increment of x when x moves from x_0 to x_1.

Definition 8.4 (Difference quotient). fixing a point x_0, the function that maps x to the slope of the chord between $(x_0, f(x_0))$ and $(x, f(x))$ is called difference quotient, defined as

$$\varphi(x) := \frac{f(x) - f(x_0)}{x - x_0} \tag{8.1}$$

Definition 8.5 (Derivative at a point). Let $f : D \subseteq \mathbb{R} \to \mathbb{R}$ be a function defined in a neighbourhood of x_0. Then f is **differentiable**[25] at x_0 if the limit of the difference quotient exists and is finite, i.e.

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \in \mathbb{R}. \tag{8.2}$$

In such case, the limit is called **derivative** of f at x_0 and can be denoted in many ways: $f'(x)$, $\dot{f}(x_0)$, $Df(x_0)$, $\frac{df}{\mathrm{d}x}(x_0)$, $\frac{dy}{\mathrm{d}x}(x_0)$.

Remark 8.6. we can define the **derivative** of f as the function f' which maps $x \mapsto f'(x)$. In general, the domain of f' is a subset of the domain of f. If they coincide, then f is said (everywhere) **differentiable**.

Definition 8.7 (Tangent). If f is differentiable at x_0, the tangent at x_0 to the graph of f is the affine function

$$\ell_{x_0}(x) = f'(x_0)(x - x_0) + f(x_0). \tag{8.3}$$

It's clear that (8.2) is equation (8.1) as x_1 approaches x_0. Geometrically, $f'(x_0)$ denotes the slope of the tangent line at x_0 to the graph of f.

Remark 8.8. If f is differentiable in a neighbourhood $I(x_0)$ and twice differentiable at x_0, then f is locally above the tangent if $f''(x_0) > 0$ and locally below if $f''(x_0) < 0$.

Definition 8.9 (First Order Approximation). Aiming to approximate the function around x_0, it can be proved[26] that (8.3) is the best[27] approximation of f by affine function, i.e. a polynomial of degree 1.

Definition 8.10 (Left / Right Derivatives). f is right-differentiable *(resp. left-differentiable)* at x_0 if it's defined in a right neighbourhood of x_0 and

$$f'_+(x_0) = \lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0+} \frac{f(x_0 + h) - f(x_0)}{h} \in \mathbb{R}. \tag{8.4}$$

Theorem 8.11. *If x_0 is an interior point of $dom(f)$, then f is differentiable at x_0 if and only if the left and right derivatives of f at x_0 exist finite and coincide. In such case, $f'(x_0) = f'_+(x_0) = f'_-(x_0)$.* [28]

[25] In analysis 2 you will see that differentiability and derivability are two distinct concepts, the first being stronger than the second. Brutally, derivability is the existence of the derivative. Differentiability is the ability to approximate the function with an affine function, that is, there exists a function $\omega : \mathbb{R} \to \mathbb{R}$ continuous at 0, with $\omega(0) = 0$, such that $f(x) = f(x_0) + f'(x_0)(x - x_0) + (x - x_0)\omega(x - x_0)$. However, we will prove that for real valued functions the two properties are equivalent.

[26] Here, we anticipate some results which will be more clear after studying Taylor Expansions.

[27] For "best" linear approximation we mean such that $f(x) - \ell(x) = o(x - x_0)$, or, equivalently, $\lim_{x \to x_0} \frac{f(x) - \ell(x)}{x - x_0} = 0$.

[28] For instance, $f(x) = |x|$ has a point of non differentiability at $x = 0$ since $f'_-(0) = -1 \neq 1 = f'_+(0)$. Notice that $f'_+(x_0)$ can exist even if $\lim_{x \to x_0+} f'(x)$ doesn't.

Definition 8.12 (K-th order derivative). if f' is differentiable in $domf$, then f is twice differentiable and the derivative of f' is called second derivative of f and denoted by f'', $D^2 f$, $\frac{d^2 f}{\mathrm{d}x^2}$. Iterating this process we can define the derivative of order k denoted by $D^k f$, $f^{(k)}$ or $\frac{d^k f}{\mathrm{d}x^k}$.

Definition 8.13 (Class C^k). f is of class C^k, and we write $f \in C^k(I)$, if f is differentiable k times and its derivative of order k is continuous in an interval I.

Definition 8.14 (Local Extrema). $x_0 \in dom(f)$ is a **local minimizer** (resp. maximizer) of f if there exists a neighbourhood $I_\delta(x_0)$ such that $f(x_0) \leq f(x)$ (resp. $f(x_0 \geq f(x))$ for every $x \in dom(f) \cap I_\delta(x_0)$.

Definition 8.15 (Critical Points). A point x_0 for which $f'(x_0) = 0$. *(also called **stationary** point)* There are four types of critical points: local minimizers, local maximizers, inflection points with horizontal tangent line, and pathological points, like $x_0 = 0$ in the function $f(x) = x^4 \sin(1/x)$ for $x \neq 0$, and $f(0) = 0$.

Definition 8.16 (Inflection Points). A point x_0 where f'' changes sign (*supposing that f is twice differentiable in a small neighbourhood of* x_0), and the function changes convexity. Therefore f crosses the tangent at x_0.

8.2 Calculus rules on derivatives

Proposition 8.17 (Linearity). *the derivative is a linear operator, that is,* $(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0)$

Proof.
The proof follows immediately from the corresponding property of limits:

$$\lim_{x \to x_0} \frac{\alpha f(x) + \beta g(x) - \alpha f(x_0) - \beta g(x_0)}{x - x_0)} = \alpha \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} + \beta \lim_{x \to x_0} \frac{g(x) - g(x_0)}{x - x_0}.$$

□

Theorem 8.18 (Leibniz Rule). $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$

Proof.
We work on the numerator, adding and substracting $f(x_0)g(x)$:

$$\begin{aligned}(fg)'(x_0) &= \lim_{x \to x_0} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} + \left(\frac{f(x_0)g(x) - f(x_0)g(x)}{x - x_0} \right) \\ &= \lim_{x \to x_0} \frac{g(x)[f(x) - f(x_0)] + f(x_0)[g(x) - g(x_0)]}{x - x_0} \\ &= \underbrace{\lim_{x \to x_0} g(x) \frac{f(x) - f(x_0)}{x - x_0}}_{g(x_0)f'(x_0)} + \underbrace{\lim_{x \to x_0} f(x_0) \frac{g(x) - g(x_0)}{x - x_0}}_{f(x_0)g'(x_0)} = f'(x_0)g(x_0) + f(x_0)g'(x_0).\end{aligned}$$

□

Theorem 8.19 (Chain Rule). $D(g \circ f)(x_0) = D(g(f(x_0))) \cdot Df(x_0)$

Remark 8.20. Assumptions: $f : dom(f) \to dom(g)$, $g : dom(g) \to \mathbb{R}$, f is differentiable at x_0, g is differentiable at $f(x_0)$, **and** x_0 is an interior point of $dom(f)$ and $f(x_0)$ is an interior point of $dom(g)$.

Proof.
Set $f(x) = y$ and $f(x_0) = y_0$. Differentiability of g at y_0 implies

$$g(y) = g(y_0) + g'(y_0)(y - y_0) + o(y - y_0) \iff \lim_{x \to x_0} \frac{g(y) - g(y_0) - g'(y_0)(y - y_0)}{y - y_0} = 0.$$

Therefore, there exists a function $\omega(y)$ **continuous** at y_0 such that $\omega(y_0) = 0$ and $g(y) - g(y_0) - g'(y_0)(y - y_0) = \omega(y)(y - y_0)$. Therefore

$$\begin{aligned}\lim_{x \to x_0} \frac{g(f(x)) - g(f(x_0))}{x - x_0} &= \lim_{x \to x_0} \frac{g(y) - g(y_0)}{x - x_0} = \lim_{x \to x_0} \frac{[g'(y_0) + \omega(y)][y - y_0]}{x - x_0} \\ &= \lim_{x \to x_0} (g'(y_0) + \omega(y)) \cdot \frac{f(x) - f(x_0)}{x - x_0} \\ &\overset{!}{=} (g'(y_0) + \omega(y_0)) \cdot f'(x_0) = g'(f(x_0)) \cdot f'(x_0),\end{aligned}$$

where at ! we used continuity of ω at y_0 and the definition of derivative of f at x_0. □

Theorem 8.21 (Derivative of the inverse). *If f is differentiable at x_0, then f^{-1} is differentiable at $y_0 = f(x_0)$ if and only if $f'(x_0) \neq 0$ and in that case $D(f^{-1}(y_0)) = 1/Df(x_0)$.*

Remark 8.22. assumptions: f is invertible and x_0 is an <u>interior</u> point of $dom(f)$.

Proof. $Df(x_0) \sim \frac{f(x)-f(x_0)}{x-x_0} = \frac{y-y_0}{f^{-1}(y)-f^{-1}(y_0)} \sim \frac{1}{Df^{-1}(y_0)}$ as $x \to x_0$ and $y \to y_0$. **[Improve]** □

8.3 Properties of differentiable functions defined in intervals

Theorem 8.23 (Fermat). *A local extremum x_0 of f is a critical point. (Other assumptions: f is differentiable at x_0 and x_0 is an <u>interior</u> point of $dom(f)$).*

Proof.
First, let's consider the case of a minimizer. Demonstrating that $f'_+(x_0) \geq 0$ and $f'_-(x_0) \leq 0$ leads us to conclude that $f'(x_0) = 0$, as the left and right derivatives must match, given that f is differentiable at x_0. The case for a maximizer is similar, but we would prove that $f'_+(x_0) \leq 0$ and $f'_-(x_0) \geq 0$.
Case A. If x_0 is a minimizer within the interval $I_+ = (x_0, x_0 + \delta)$, then $\frac{f(x)-f(x_0)}{x-x_0} \geq 0$ for all $x \in I_+$, as the numerator is non-negative and the denominator is positive. Hence, $f'_+(x_0) = \lim\limits_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} \geq 0$.
Case B. If x_0 is a minimizer within the interval $I_- = (x_0 - \delta, x_0)$, then $\frac{f(x)-f(x_0)}{x-x_0} \leq 0$ for all $x \in I_-$, since the numerator is non-negative and the denominator is negative. Therefore, $f'_-(x_0) = \lim\limits_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} \leq 0$. □

Theorem 8.24 (Rolle). *If $f(a) = f(b)$ then there exists a point $x^\star \in (a, b)$ such that $f'(x^\star) = 0$. (Other assumptions: f is continuous in $[a, b]$ and differentiable in (a, b).)*

Proof.
By the Weierstrass theorem there exist a global minimizer x_m and a global maximizer x_M of f in $[a, b]$. If at least one of them, say x_m, belongs to (a, b), then $f'(x_m) = 0$ by Fermat's theorem. Otherwise, if $x_m, x_M \in \{a, b\}$, then $f(x_m) = f(a) = f(b) = f(x_M)$, so the global minimum $f(x_m)$ and the global maximum $f(x_M)$ coincide, so $f(x)$ is constant in $[a, b]$ and therefore $f'(x)$ vanishes identically. □

Theorem 8.25 (Mean Value (Cauchy)). *Let $f, g : [a, b] \to \mathbb{R}$. Then there exists a point $x^\star \in (a, b)$ such that*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x^\star)}{g'(x^\star)}.$$

(Other assumptions: f, g are continuous in $[a, b]$, differentiable in (a, b) and $g'(x) \neq 0$ $\forall x \in (a, b)$)

Proof.
By hypothesis, $g'(x) \neq 0$, hence $g(b) \neq g(a)$ by the contrapositive of Rolle's Theorem. The thesis is equivalent to finding $x^\star \in (a, b)$ such that

$$g'(x^\star)(f(b) - f(a)) - f'(x^\star)(g(b) - g(a)) = 0. \tag{8.5}$$

Let us define a similar function:

$$h(x) := g(x)(f(b) - f(a)) - f(x)(g(b) - g(a)).$$

Let us compute $h(a)$ and $h(b)$:

$$h(a) = g(a)f(b) - g(a)f(a) - f(a)g(b) + g(a)f(a) = g(a)f(b) - f(a)g(b)$$
$$h(b) = g(b)f(b) - g(b)f(a) - f(b)g(b) + f(b)g(a) = g(a)f(b) - f(a)g(b)$$

Hence, $h(a) = h(b)$. It is also clear that h is continuous in $[a,b]$ and differentiable in (a,b). By rolle's theorem, there exists $x^\star \in (a,b)$ such that $h'(x^\star) = 0$. It is easy to see that $h'(x^\star) = 0$ is equivalent to (8.5), concluding the proof. □

Theorem 8.26 (Mean Value (Lagrange, MVT)). *There exists a point $x^\star \in (a,b)$ such that $\boldsymbol{f'(x^\star) = \frac{f(b)-f(a)}{b-a}}$. (Other assumptions: f is continuous in $[a,b]$ and differentiable in (a,b).)*

Proof.
Let $g(x) = x$. Applying Cauchy's theorem to f and g we find a point $x^\star$ such that $\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(x^\star)}{g'(x^\star)} \implies f'(x^\star) = \frac{f(b)-f(a)}{b-a}$. □

Theorem 8.27 (L'hôpital's Rule). *If $\lim_{x\to x_0} \frac{f(x)}{g(x)} = l$ produces the indeterminate form $[0/0]$ or $[\infty/\infty]$ then $\boldsymbol{\lim_{x\to x_0} \frac{f'(x)}{g'(x)} = l \implies \lim_{x\to x_0} \frac{f(x)}{g(x)} = l}$ (assumptions: $x_0, l \in \overline{\mathbb{R}}$, f, g are defined and differentiable in a <u>pointed</u> neighbourhood of x_0 -symmetric, right or left-, and $g' \neq 0$ there).*

Proof.
[improve] We'll prove just the case $\lim\limits_{x\to x_0+} f(x) = \lim\limits_{x\to x_0+} g(x) = 0$ and $\dot{I}(x_0) = (x_0, x_0+\delta)$. Extend f and g to x_0 by setting $f(x_0) := 0$ and $g(x_0) := 0$. This modification ensures that the new functions f, g are continuous in $\dot{I}(x_0)$ and differentiable in $I(x_0)$. Consider the interval $[x_0, k] \subseteq [x_0, x_0+\delta)$. Applying Cauchy's theorem to this interval gives a point $x^\star(k) \in (x_0, k)$ such that $\frac{f'(x^\star(k))}{g'(x^\star(k))} = \frac{f(x)-f(x_0)}{g(x)-g(x_0)} = \frac{f(x)}{g(x)}$. Since $x_0 < x^\star(k) < k$, taking the limit as $k \to x_0+$ leads to $x^\star(k) \to x_0+$, and therefore, by the transference principle, $\lim\limits_{x\to x_0+} \frac{f(x)}{g(x)} = \lim\limits_{x\to x_0+} \frac{f'(x^\star(k))}{g'(x^\star(k))} = \lim\limits_{y\to x_0+} \frac{f'(y)}{g'(y)} = l$. *(If $I(x_0)$ is different, simply consider $[k, x_0]$ or $[x_0-k, x_0+k]$).* □

8.4 Applications of theorems involving differential properties

Finding maxima and minima: Given a real differentiable function f, the Weierstrass Theorem assures the existence of a global minimizer $x^\star$ within the interval $[a,b]$. According to Fermat's Theorem, if this minimizer is an interior point, then it must satisfy $f'(x^\star) = 0$. Therefore, if there are only a finite number of points where $f'(x) = 0$, we can identify the minimizer $\overline{x}$ among these. Then, we consider the minimum value among $f(\overline{x}), f(a)$, and $f(b)$: this is the minimum of the function (*In the last passage we checked what happens at the boundaries*).

Theorem 8.28 (Differentiability and continuity). *If f is differentiable at x_0 then it's also continuous at x_0.*

Proof.
f is continuous at x_0 if and only if $\lim\limits_{x\to x_0} f(x) = f(x_0)$, that is, $\lim\limits_{x\to x_0} f(x) - f(x_0) = 0$.

$$\lim_{x\to x_0} f(x) - f(x_0) = \underbrace{\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}}_{f'(x_0)\in\mathbb{R}} \cdot (x-x_0) = f'(x_0) \cdot 0 = 0.$$

□

Theorem 8.29 (Null Derivative). *if f is differentiable in an interval I and $f'(x) = 0\ \forall x \in I$, then f is constant in I.*

Proof.
Take two points $a < b \in I$. Applying Lagrange's theorem we can find $x^\star \in (a, b) \subseteq I$ such that $f'(x^\star)(b - a) = f(b) - f(a)$, but we know that $f'(x^\star) = 0$, so $f(a) = f(b)\ \forall a, b \in I$. □

Theorem 8.30 (Derivative and monotonicity). *If f is differentiable in the **interior** of an interval I, then f is increasing (resp. decreasing) if and only if $f'(x) \geq 0$ (resp. $f'(x) \leq 0$) for every $x \in I$. Moreover, if $f'(x) > 0$, then f is strictly increasing (resp. decreasing)*[29].

Proof.
We consider the case of an increasing function (the other cases are similar). If $f'(x) \geq 0$, we want to prove that if $a, b \in I$ and $a < b$ then $f(a) \leq f(b)$. Applying Lagrange theorem on $[a, b]$ we get the existence of $x^\star \in (a, b)$ such that

$$f(b) - f(a) = f'(x^\star)(b - a) \quad \text{and} \quad f'(x^\star) \geq 0 \implies f(b) \geq f(a).$$

Notice that if $f'(x^\star) > 0$ then the inequality is strict and therefore we get the strict monotonicity.
Conversely, if $f(x_0) \leq f(x_1)$ for $x_0 < x_1$, then $\frac{f(x_1)-f(x_0)}{x_1-x_0} \geq 0$. Passing to the limit as $x_1 \to x_0$ and using the permanence of sign we get $f'(x_0) \geq 0\ \forall x_0 \in I$. □

Theorem 8.31 (Lipschitz Functions). *If f is differentiable in the **interior** of an interval I, the f is **L-Lipschitz** in I if and only if $\boldsymbol{|f'(x)| \leq L}$ for every x in the interior of I. In other words, these conditions are equivalent:*

$$\textit{Bounded derivative} \iff \textit{L-Lipschitz} \iff |f(x_1) - f(x_0)| \leq L|x_1 - x_0|.$$

Proof.
If $|f'(x)| \leq L$, we have to prove that $|f(x_1) - f(x_0)| \leq L|x_1 - x_0|$. Notice that applying **Lagrange's Theorem** to $[x_0, x_1] \subseteq I$ (*supposing* $x_0 < x_1$) we get the existence of $x^\star \in (x_0, x_1)$ such that $|f(x_1) - f(x_0)| = |f'(x^\star)(x_1 - x_0)| = |f'(x^\star)||x_1 - x_0| \leq L|x_1 - x_0|$.
Conversely, if the function is L-Lipschitz, then $\boldsymbol{-L \leq \frac{f(x_1)-f(x_0)}{x_1-x_0} \leq L}$ for every $x_0 \neq x_1$. Passing to the limit as $x_1 \to x_0$ we obtain $-L \leq f'(x_0) \leq L$. □

Theorem 8.32 (Criterion for differentiability). *If f is continuous in a right (resp.left) neighbourhood $I_\delta(x_0) = [x_0, x_0 + \delta)$ of x_0 and it's differentiable in $I - \{x_0\}$, then $\lim_{x \to x_0+} f'(x) = l \implies f'_+(x_0) = l$, with $l \in \overline{\mathbb{R}}$.* [30]

Pay attention to the conditions! f has not to be differentiable at x_0, but it must be there continuous!

Proof.
$f'_+(x_0) = \lim_{x \to x_0+} \frac{f(x)-f(x_0)}{x-x_0}$. Note that this is an indeterminate form $[0/0]$. Therefore we can apply l'Hôpital's Rule, so that

$$f'_+(x_0) = \lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0} \overset{\mathrm{H}}{=} \lim_{x \to x_0+} \frac{f'(x) - f'(x_0)}{1} = \lim_{x \to x_0+} f'(x) = l.$$

□

Theorem 8.33 (First order expansion.). *The best approximation of f by an affine function $\ell(x)$ in a neighbourhood of x_0 (i.e. such that $f(x) - \ell(x) = o(x - x_0)$) exists if and only if f is differentiable at x_0 and in that case $\ell(x) = f'(x_0)(x - x_0) + f(x_0)$.*

[29] Stronger statement: if f is differentiable in an interval I, then f is strictly increasing if and only if $f'(x) \geq 0$ for every point on I and there are no subintervals of I where f' identically vanishes.

[30] This is useful with piecewise functions. If the function is not continuous at x_0, clearly neither is differentiable. If it is, we can calculate the limits of the derivative from the left and from the right, obtaining $f_-(x_0)$ and $f'_+(x_0)$, thanks to this criterion. And now we check whether $f_-(x_0) = f_+(x_0) = l \in \mathbb{R}$. If yes, then $f'(x_0) = l$. If not, f is not differentiable at x_0.

Proof.
Since $f(x) - \ell(x) = o(x - x_0)$, it follows that $l(x_0) = f(x_0)$, hence $\ell(x) = p(x - x_0) + f(x_0)$. Now, we need to identify p such that

$$f(x) - f(x_0) - p(x - x_0) = o(x - x_0) \iff \lim_{x \to x_0} \frac{f(x) - f(x_0) - p(x - x_0)}{x - x_0} = 0 \qquad (*)$$

1) If f is differentiable at x_0, setting $p := f'(x_0)$ results in

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

2) Conversely, if f is differentiable at x_0 and there exists, p such that $(*)$ holds true, then

$$0 = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} - p = f'(x_0) - p \implies p = f'(x_0).$$

□

Theorem 8.34 (Second order approximation). *If f is differentiable in $I(x_0)$ and twice differentiable at x_0, the best*[31] *quadratic approximation of f has the form*

$$T_2(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2.$$

Proof.
Being $o((x-x_0)^2)$ is stronger than being $o(x-x_0)$. Therefore, we must also have $f(x)-T_2(x) = o(x-x_0)$, and from the first-order expansion, we deduce that $T_2(x) = f(x_0) + f'(x_0)(x - x_0) + r(x - x_0)^2$. Our task now is to determine r such that $\lim_{x \to x_0} \frac{f(x)-T_2(x)}{(x-x_0)^2} = 0$. Note that $f(x_0) - T_2(x_0) = 0$, leading to an indeterminate form. Applying l'**Hôpital's rule**, we derive

$$\begin{aligned} 0 = \lim_{x \to x_0} \frac{f(x) - T_2(x)}{(x - x_0)^2} &\overset{\text{H}}{=} \lim_{x \to x_0} \frac{f'(x) - T_2'(x)}{2(x - x_0)} = \lim_{x \to x_0} \frac{f'(x) - f'(x_0) - 2r(x - x_0)}{2(x - x_0)} \\ &= \frac{1}{2} \lim_{x \to x_0} \frac{f'(x) - f'(x_0)}{x - x_0} - r = \frac{1}{2} f''(x_0) - r \implies r = \frac{1}{2} f''(x_0). \end{aligned}$$

□

8.5 Exercises

This section is a selection of problems that involve applying the key results on differentiability by defining clever auxiliary functions.

Example 1: Rolle Let $f : [0, 1] \to \mathbb{R}$ be a continuous function, differentiable on $(0, 1)$ and such that $f(1) = 0$. Show that there exists $\xi \in (0, 1)$ such that $\xi f'(\xi) + f(\xi) = 0$.

Solution: The function $g(x) := xf(x)$ is continuous on $[0, 1]$, differentiable on $(0, 1)$ and $g(0) = g(1) = 0$. Therefore we can apply Rolle's theorem, finding $\xi \in (0, 1)$ such that

$$g'(\xi) = 0 \implies \xi f'(\xi) + f(\xi) = 0.$$

Example 2: Lagrange MVT Prove that for every $x > 0$ there exists $y \in (0, x)$ such that $\sin(x) = x \cos(y)$.

Solution: Applying Lagrange's MVT we get the existence of $y \in (0, x)$ such that

$$\frac{f(x) - f(0)}{x - 0} = f'(y) \implies \frac{\sin(x) - \sin(0)}{x - 0} = \cos(y) \implies \sin(x) = x \cos(y).$$

[31] that is, $f(x) - T_2(x) = o((x - x_0)^2)$, or, equivalently, $\lim_{x \to x_0} \frac{f(x)-T_2(x)}{(x-x_0)^2} = 0$.

Example 3: Lagrange MVT Let $f : [0, +\infty) \to \mathbb{R}$ be a continuous function which is differentiable on $(0, +\infty)$. If $|f'(x)| \leq \sqrt{x}$ for $x > 0$, prove that $\lim\limits_{x\to+\infty} \frac{f(x)}{x^2} = 0$.

Solution: Applying Lagrange's MVT we get the existence of $y \in (0, x)$ such that

$$\left|\frac{f(x) - f(0)}{x - 0}\right| = |f'(y)| \leq \sqrt{y} \implies |f(x) - f(0)| \leq x\sqrt{y} \leq x\sqrt{x}.$$

Therefore:

$$\lim_{x\to+\infty} \frac{f(x)}{x^2} = \lim_{x\to+\infty} \frac{f(x) - f(0)}{x^2} + \frac{f(0)}{x^2} \leq \lim_{x\to+\infty} \frac{|f(x) - f(0)|}{x^2} + \frac{f(0)}{x^2} \leq \lim_{x\to+\infty} \frac{x\sqrt{x}}{x^2} = 0,$$

by comparison test (we should repeat the same process using $f(x) - f(0) \geq -x\sqrt{x}$ to prove that the limit is non-negative).

Example 4: derivative and monotonicity Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function such that $f(0) = 0$ and $f'(x) \leq 2$ for every $x \in \mathbb{R}$. Prove that $f(x) \leq 2x$ for every $x \geq 0$ and $f(x) \geq 2x$ for every $x < 0$.

Solution: The function $g(x) := f(x) - 2x$ is such that $g'(x) = f'(x) - 2 \leq 0$, so that g is (non-strictly) decreasing, implying that

$$f(a) - 2a = g(a) \geq g(0) = f(0) - 0 = 0 \quad \text{and} \quad 0 = g(0) \geq g(b) = f(b) - 2b, \quad \text{for every } a < 0 < b.$$

This is enough to conclude that $f(a) \geq 2a$ and $f(b) \leq 2b$, for every $b \geq 0$ and $a < 0$.

Example 5: derivative and monotonicity Prove that if $f : \mathbb{R} \to \mathbb{R}$ is differentiable and $\lim\limits_{x\to+\infty} f'(x) = l > 0$, then $\lim\limits_{x\to+\infty} f(x) = +\infty$.

Solution: We can write the limit condition by saying that eventually $f'(x) > \frac{l}{2}$, that is, there exists $\overline{x} \in \mathbb{R}$ such that $f'(x) > \frac{l}{2}$ for every $x > \overline{x}$. Therefore, the function $g(x) := f(x) - \frac{l}{2}x$ is such that $g'(x) = f'(x) - \frac{l}{2} > 0$ for every $x > \overline{x}$. Therefore, $g(x)$ is strictly increasing for $x > \overline{x}$. In other words,

$$g(x) \geq g(\overline{x}) \implies f(x) - \frac{l}{2}x \geq f(\overline{x}) - \frac{l}{2}\overline{x} \implies f(x) \geq \frac{l}{2}x + f(\overline{x}) - \frac{l}{2}\overline{x}, x > \overline{x}.$$

Passing the limit as $x \to +\infty$, we conclude that $\lim\limits_{x\to+\infty} f(x) = +\infty$, by comparison test.

9 Taylor Expansions

In this section, we will consider f defined in a neighbourhood $I(x_0)$. Following the positive result of approximating a differentiable function with an affine function (a polynomial of degree 1), we may wonder whether is it possible to add further terms in order to improve the approximation. With this idea, our goal is to find a list of polynomials $T_1(x), T_2(x), \ldots, T_n(x)$ which increasingly improve the approximation of the function around a fixed point x_0. These polynomials are unique[32] and called Taylor Polynomials centered at x_0, having the form:

$$T_N(x) = \sum_{k=0}^{N} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \ldots + \frac{f^{(N)}(x_0)}{N!}(x - x_0)^N. \tag{9.1}$$

Theorem 9.1 (Taylor - Peano). *If f is N times differentiable at x_0 and $N - 1$ times differentiable in $I(x_0)$, then $T_N(x)$ is the unique polynomial of degree $\leq N$ such that:*

$$T(x_0) = f(x_0), \quad T'(x_0) = f'(x_0), \quad \ldots, \quad T^{(N)}(x_0) = f^{(N)}(x_0) \tag{9.2}$$

$$f(x) = T(x) + \underbrace{o(|x - x_0|^N)}_{\textit{Peano Remainder}} \quad \textit{as } x \to x_0 \quad \textit{(asymptotic expansion of } f\textit{)} \tag{9.3}$$

[32] we haven't proved this result, except for $n = 1$ and $n = 2$.

Theorem 9.2 (Taylor - Lagrange). *If f is $N+1$ times differentiable at $I(x_0)$, then $T_N(x)$ is the unique polynomial such that, for every $x \in I(x_0)$, there exists ξ between x_0 and x such that*

$$f(x) - T_N(x) = \underbrace{\frac{f^{(N+1)}(\xi)}{(N+1)!}(x-x_0)^{N+1}}_{\textit{Lagrange Error}} \tag{9.4}$$

Definition 9.3 (Taylor Expansion). the Taylor Expansion of f centered at x_0 with Peano Remainder can be expressed in these two equivalent ways:

$$f(x) = \sum_{k=0}^{N} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + o((x-x_0)^N) \quad \text{as } x \to x_0 \tag{9.5}$$

$$f(x_0+h) = \sum_{k=0}^{N} \frac{f^{(k)}(x_0)}{k!}h^k + o(h^N) \quad \text{as } h \to 0. \tag{9.6}$$

If $x_0 = 0$, the expansion is sometimes called **MacLaurin**'s Expansion[33].

Definition 9.4 (Taylor Series). If $a_n = \frac{f^{(n)}(x_0)}{n!}$, then $T_N(x)$ defines a sequence of partials sums in N, as we add additional terms to our sum. $\lim_{N\to+\infty} T_N(x)$ is in fact a series, called Taylor Series. Taylor Series have a radius of convergence, that is, a convex set of values of x for which the series converges. Every elementary function can be equivalently written as a Taylor Series, as long as x stays in the radius of convergence. Taylor Series are shown in the next table (*As an exercise, try to obtain these formulas starting from Taylor expansion in Peano's form*)[34].

Taylor series

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \dots$$

$$\sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{7!} + \dots$$

$$\sinh(x) = \frac{e^x - e^{-x}}{2} = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} = x + \frac{x^3}{6} + \frac{x^5}{120} + \frac{x^7}{7!} + \dots$$

$$\cos(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} + \dots$$

$$\cosh(x) = \frac{e^x + e^{-x}}{2} = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} = 1 + \frac{x^2}{2} + \frac{x^4}{24} + \frac{x^6}{720} + \dots$$

$$\ln(x+1) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots, \quad |x| < 1$$

$$\arctan(x) = \sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{2k+1} x^{2k+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots$$

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2}x^2 + \dots, \quad |x| < 1$$

$$\tan(x) = x + \frac{x^3}{3} + \frac{2}{15}x^5 + \dots$$

$$\arcsin(x) = x + \frac{x^3}{6} + \dots$$

$$\arccos(x) = \frac{\pi}{2} - x + \dots$$

9.1 Exercises

This section is a selection of problems that require the use of Taylor Expansions, whether implicitly or explicitly required.

[33]Question: what's the Taylor expansion of a polynomial?

[34]Curiosity: not every function, even if of class C^∞, can be written as a Taylor Series, not even locally. For example, the function $f(x) = e^{-1/x^2}$ for $x > 0$ and $f(x) = 0$ for $x \leq 0$ is a C^∞ function whose Taylor Series centered at 0 is the constant function 0. Therefore, its Taylor series fails to approximate the function for any neighbourhood of 0. Such function is said **smooth** but **non-analytic** at the origin.

Example 1: composition of Taylor Expansions Find the Taylor Expansion of order 3 centered at $x = 0$ of $f(x) := \cos(\ln(2x + 1))$.

Solution: Truncating Taylor Series as a polynomial of degree at most 3 we get

$$\cos(\ln(2x+1)) = \cos\left(2x - \frac{4x^2}{2} + \frac{8x^3}{3}\right) + o(x^3) = 1 - \frac{\left(2x - 2x^2 + \frac{8x^3}{3}\right)^2}{2} + o(x^3)$$
$$= 1 - \left(2x^2 + \cancel{2x^4} + \cancel{\frac{32x^6}{9}} - 4x^3 + \cancel{\frac{16x^4}{3}} - \cancel{\frac{16x^5}{3}}\right) + o(x^3) = 1 - 2x^2 + 4x^3 + o(x^3).$$

Example 2: Approximating functions Find a rational function of the form $\frac{ax+b}{cx+d}$ which well approximates $f(x) = \frac{1}{e^x - 1}$ in a pointed neighbourhood of 0.

Solution: Observe that

$$e^x - 1 = x + \frac{x^2}{2} + \frac{x^3}{6} + o(x^3) = x\left(1 + \frac{x}{2} + \frac{x^2}{6} + o(x^2)\right) \implies f(x) = \frac{1}{x}\frac{1}{1 + \frac{x}{2} + \frac{x^2}{6} + o(x^2)}.$$

The latter ratio can be treated using MacLaurin's expansion $1/(1+y) = 1 - y + y^2 + o(y^2)$ and therefore

$$f(x) = \frac{1}{x}\left(1 - \left(\frac{x}{2} + \frac{x^2}{6}\right) + \left(\frac{x}{2} + \frac{x^2}{6}\right)^2 + o(x^2)\right)$$
$$= \frac{1}{x}\left(1 - \frac{x}{2}\right) = \frac{1}{x}\left(1 - \frac{x}{2} + \frac{x^2}{12} + o(x^2)\right) = \frac{1}{x} - \frac{1}{2} + o(1) = \frac{2-x}{2x}.$$

Example 3: limits Compute the limit $\lim\limits_{x\to 0} \frac{\sqrt[4]{1+4x} - e^{\sin x} + x^2}{1 - \cos(6x)}$.

Solution: Using MacLaurin's expansions we get

$$\sqrt[4]{1+4x} = (1+4x)^{1/4} = 1 + \frac{1}{4}(4x) + \frac{1}{2}\cdot\frac{1}{4}\cdot\left(-\frac{3}{4}\right)(4x)^2 + o(x^2) = 1 + x - \frac{3}{2}x^2 + o(x^2)$$
$$e^{\sin x} = 1 + \sin x + \frac{1}{2}(\sin x)^2 + o(\sin^2 x) = 1 + x + \frac{1}{2}x^2 + o(x^2)$$
$$\sqrt[4]{1+4x} - e^{\sin x} + x^2 = \left(1 + x - \frac{3}{2}x^2 + o(x^2)\right) - \left(1 + x + \frac{1}{2}x^2 + o(x^2)\right) + x^2 = -x^2 + o(x^2)$$
$$1 - \cos(6x) = \frac{1}{2}(6x)^2 + o(x^2) = 18x^2 + o(x^2)$$
$$\lim_{x\to 0} \frac{\sqrt[4]{1+4x} - e^{\sin x} + x^2}{1 - \cos(6x)} = \lim_{x\to 0}\frac{-x^2}{18x^2} = -\frac{1}{18}.$$

Example 4: Lagrange remainder Approximate $\sin(1/10)$ with an error less than 10^{-9}.

Solution: We can employ MacLaurin expansion with Lagrange remainder:

$$\sin(1/10) - T_N(x) = \left(\frac{\mathrm{d}^{N+1}}{\mathrm{d}x^{N+1}}\sin(\xi)\right)\cdot\frac{(1/10)^{N+1}}{(N+1)!}, \text{for some } \xi \in (0, 1/10)$$

It's enough to select $N = 6$, obtaining:

$$|\sin(1/10) - T_6(1/10)| = \left|\left(\frac{\mathrm{d}^7\sin(\xi)}{\mathrm{d}x^7}\right)\cdot\frac{10^{-7}}{7!}\right| = \left|\frac{(-1)^7\cdot\cos(\xi)}{7!}\cdot 10^{-7}\right| = \left|\frac{\cos(\xi)}{10^7\cdot 7!}\right| \le \frac{1}{10^9}.$$

Therefore, approximating $\sin(1/10)$ with $T_6(1/10)$ is a sufficient approximation:

$$T_6(1/10) = \left[x - \frac{x^3}{6} + \frac{x^5}{120}\right]_{x=1/10} = \frac{1}{10} - \frac{1}{6\cdot 10^3} + \frac{1}{120\cdot 10^5} = \frac{1.200.000 - 2.000 + 1}{120\cdot 10^5} = \frac{1.198.001}{12.000.000}.$$

10 Convexity

10.1 Preliminaries

Convexity is a powerful and broad concept in mathematics, which can be studied on multiple perspectives. Here, we will focus our attention to *real convex functions*, but, before doing so, it's important to understand the concept of convexity in its general meaning. Broadly speaking, a set of points in a Euclidean Space[35] is defined to be convex if it contains the line segments connecting each pair of its points. For example, A circle is convex, while a "V" shape is not, since there exists a segment whose extrema lie on it, but the segment is not entirely contained in it. A set S of n points generates a convex set in the sense that there exists a unique minimal convex set containing S, called **convex envelope** of S.

Definition 10.1 (Convex combinations). A *convex combination* of real numbers $x_1, \dots, x_n$ is defined as $\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$, where $\alpha_i \geq 0$ for each i, and $\alpha_1 + \dots + \alpha_n = 1$.

Remark 10.2. it's a particular case of linear combination[36]. The coefficients are sometimes called weights, due to the straightforward interpretation that a convex combination of n real numbers is a weighted sum of them.

Example 10.3. a classic example of convex combination is the **arithmetic mean**, where $\alpha_1 = \dots = \alpha_n = 1/n$. In this case, the convex combination simplifies to $\frac{x_1 + \dots + x_n}{n}$. For any other choice of weights, the convex combination is a **weighted arithmetic mean**.

Remark 10.4. In the case of $n = 2$, it's convenient to express the combination as $\alpha_1 + \alpha_2 = 1$ by setting $\alpha_1 = t$ and $\alpha_2 = 1 - t$ (where $t \in [0, 1]$). This results in the convex combination $x_t = tx_1 + (1 - t)x_2$.

Remark 10.5 (Geometric Interpretation). Given two vectors $P_0, P_1 \in \mathbb{R}^2$, the vector combination $P_t = (1 - t)P_0 + tP_1$ parametrizes (i.e. represents) the line segment $l = [P_0, P_1]$, meaning that, for any point X in l, there exists $t_x \in [0, 1]$ such that $X = (1 - t_x)P_0 + t_x P_1$, and, for any $t_x \in [0, 1]$, the point $X = (1 - t_x)P_0 + t_x P_1$ belongs to l. Therefore, line segments between points P_0 and P_1 can be defined as convex combinations of P_0 and P_1.

Definition 10.6 (Epigraph and Hypograph). Given an interval I and a function $f : I \to \mathbb{R}$, the epigraph is the set of points in $I \times \mathbb{R}$ that lie **above** the function f. Conversely, the hypograph is the set of points **below** the function. The epigraph can be formally defined as $\mathrm{Epi}(f) : \{(x, y) : x \in I, y \geq f(x)\}$.

10.2 Convex Functions in an (open and convex) interval

We will list many equivalent formulations of the concept of convex function. Our goal will be to build a chain of double implications to assure all the characterizations being equivalent to each other. Recall that an open interval has the form (a, b), while a closed interval has the form $[a, b]$. Any of the next result hold as long as we consider open intervals, since boundary points may behave weirdly. Therefore, either we assume I open, or we consider only points in the interior of I. **Achtung!** Never apply these theorems on single points, which are closed intervals! Counterintuitively, $f''(0) > 0$ does not imply f convex at $x = 0$, since this characterization holds only in the interior of an interval!

i) If $x_1, \dots, x_n \in I$, $\alpha_1, \dots, \alpha_n \geq 0$ and $\alpha_1 + \alpha_2 + \ldots + \alpha_2 = 1$, then

$$\boldsymbol{f(\alpha_1 x_1 + \ldots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \ldots + \alpha_n f(x_n)}.$$

Remark 10.7. Special case: if $\alpha_i = 1/n$, it follows that $f\left(\frac{x_1 + \ldots + x_n}{n}\right) \leq \frac{f(x_1) + \ldots + f(x_n)}{n}$. Graphically, the output of the average is below the average of the outputs

ii) If $x_0, x_1 \in I$ and $t \in [0, 1]$, then $\boldsymbol{f((1 - t)x_0 + tx_1) \leq (1 - t)f(x_0) + tf(x_1)}$

Proof. ($1 \implies 2$) It's a special case of (1), setting $n = 2$, $\alpha_1 = 1 - t$, $\alpha_2 = t$, $x_1 = x_0$ and $x_2 = x_1$.

($2 \implies 1$) We didn't prove this implication, which requires a clever use of induction[37]. $\square$

iii) The graph of f is below any chord.

Proof. ($2 \iff 3$) This is essentially a geometric interpretation of the inequality at point 2: if $P_0 = (x_0, y_0)$, $P_1 = (x_1, y_1)$ and P has coordinates $(1 - t)x_0 + tx_1$, then $f((1 - t)x_0 + tx_1)$ is the

[35] A space, like $\mathbb{R}$ or $\mathbb{R}^n$, where we define the distance between two points in the usual way.

[36] it's "convex" because the set of convex combinations of $x_1, \dots, x_n$ coincides with the convex envelope of the set $\{x_1, \dots, x_n\}$.

[37] A proof can be found on Wikipedia's article "Jensen's Inequality"

height of f at P. Conversely, $(1-t)f(x_0)+tf(x_1) = (1-t)y_0+ty_1$ indicates the height of the line segment (chord) connecting P_0 and P_1 at point P. Therefore, the chord lies above the function f in the interval between P_0 and P_1. □

iv) The epigraph of f is convex[38].

Proof. (3 $\iff$ 4) A set is convex whenever it contains the segment $[P_0, P_1]$ joining any two points of it. Since the graph of f is always below any chord, the epigraph always contains any segment $[P_0, P_1]$. The converse is trivial. □

v) for every point $x_0 < x_t < x_1$ in I we have

$$\frac{f(x_t)-f(x_0)}{x_t-x_0} \le \frac{f(x_1)-f(x_0)}{x_1-x_0} \le \frac{f(x_1)-f(x_t)}{x_1-x_t}.$$

Proof. (2 $\iff$ 5) The proof involves mere calculations. Start from the first inequality and parametrize x_t as $x_t = (1-t)x_0 + tx_1$ (which can be done since $x_0 < x_t < x_1$):

$$\frac{f(x_t)-f(x_0)}{[(1-t)x_0+tx_1]-x_0} \le \frac{f(x_1)-f(x_0)}{x_1-x_0} \iff \frac{f(x_t)-f(x_0)}{t\cancel{(x_1-x_0)}} \le \frac{f(x_1)-f(x_0)}{\cancel{x_1-x_0}}$$

$$f(x_t)-f(x_0) \le tf(x_1)-tf(x_0) \iff f(x_t) \le (1-t)f(x_0)+tf(x_1) \iff f((1-t)x_0+tx_1) \le (1-t)f(x_0)+tf(x_1),$$

which is characterization 2. Similar computations are performed for the second inequality:

$$\frac{f(x_1)-f(x_0)}{x_1-x_0} \le \frac{f(x_1)-f(x_t)}{[x_1-(1-t)x_0]-tx_1} \iff \frac{f(x_1)-f(x_0)}{\cancel{x_1-x_0}} \le \frac{f(x_1)-f(x_t)}{(1-t)\cancel{x_1-x_0}}$$

$$(1-t)[f(x_1)-f(x_0)] \le f(x_1)-f(x_t) \iff f(x_t) \le (1-t)f(x_0)+tf(x_1).$$

□

Remark 10.8 (Geometric interpretation). the left-hand side represents the slope of the chord $[P_0, P_t]$, the middle part denotes the slope of the chord $[P_0, P_1]$ and the right-hand side symbolizes the slope of $[P_t, P_1]$

vi) For every $x_0 \in I$, the function $\varphi(x) = \frac{f(x)-f(x_0)}{x-x_0}$ is increasing.

Proof. (5 $\iff$ 6) It's an equivalent useful restatement of characterization 5. □

vii) for every point $x_0 < x_1$ in I we have

$$f'_-(x_0) \le f'_+(x_0) \le \frac{f(x_1)-f(x_0)}{x_1-x_0} \le f'_-(x_1) \le f'_+(x_1)$$

Proof. (6 $\implies$ 7)[39] The central inequalities can be written in terms of limits:

$$\lim_{x\to x_0+} \frac{f(x)-f(x_0)}{x-x_0} \le \frac{f(x_1)-f(x_0)}{x_1-x_0} \le \lim_{x\to x_1-} \frac{f(x_1)-f(x)}{x_1-x}.$$

These limits exist because the functions $\varphi_0(x) = \frac{f(x)-f(x_0)}{x-x_0}$ and $\varphi_1(x) = \frac{f(x_1)-f(x)}{x_1-x}$ are, by assumption, increasing for every x_0, x_1 (Theorem 6.9).[40] Thus, the two inequalities can be written in terms of $\varphi_0(x)$ and $\varphi_1(x)$:

$$\lim_{x\to x_0+} \varphi_0(x) \le \varphi_0(x_1) \qquad \varphi_1(x_0) \le \lim_{x\to x_1-} \varphi_1(x)$$

Using $x_0 < x_1$ we can conclude that the first inequality must hold in a right neighbourhood of x_0+

[38] This can be seen as the connection between the concept of "convex function" and convexity as a property of sets. A function in an interval is convex if the set of points above it is convex.

[39] We intended this as a corollary of characterization 6. However, the converse could also be proven, for instance, as the following theorem: if f has right and left derivatives for every point in I, and the right (or left) derivative is increasing in x, then the function is convex.

[40] We haven't proven that these limits are actually finite. The idea is to use characterization 6 with suitable points $x_0 - h \in I$ and $x_1 + h \in I$, with h small enough. Only now we are sure $f'_+(x_0)$ and $f'_-(x_1)$ exist.

(since $\varphi_0(x)$ is increasing) and the second in a left neighbourhood of x_1, concluding the first part of the proof.

As for the other two inequalities, characterization 6 implies that, for $h > 0$ and $k > 0$ small enough, we have

$$\frac{f(x_0) - f(x_0 - h)}{h} = \varphi(x_0 - h) \leq \varphi(x_0 + k) = \frac{f(x_0 + k) - f(x_0)}{k}.$$

Passing the limit for $h \to 0+$ and the limit for $k \to 0+$ we obtain $f'_-(x_0) \leq f'_+(x_0)$. Similarly, $f'_-(x_1) \leq f'_+(x_1)$. ☐

Remark 10.9 (Geometric interpretation:). the left-hand side represents the slope of the tangent to f at x_0, the middle denotes the slope of the chord $[P_0, P_1]$ and the right-hand side symbolizes the slope of the tangent to f at x_1.

viii) The graph of f is above its tangents, i.e. $f(x) \geq f(x_0) + f'(x_0)(x - x_0)$ (if f is differentiable at $x_0 \in I$)

Proof. (7 $\iff$ 8) First, assuming the function is convex, we consider three scenarios: 1) When $x = x_t$, the inequality simplifies to $0 \geq 0$. 2) If $x > x_t$, the inequality transforms to $f'(x_t) \leq \frac{f(x)-f(x_t)}{x-x_t}$, which holds by point 6. 3) If $x < x_t$, it becomes $f'(x_t) \geq \frac{f(x_t)-f(x)}{x_t-x}$, which holds by point 6. The converse is analogous.

(8 $\implies$ 2) If $f(x) \geq f(x_t) + f'(x_t)(x - x_t)$, it follows that

$$f(x_0) \geq f(x_t) + f'(x_t)(x_0 - x_t) \quad \text{and} \quad f(x_1) \geq f(x_t) + f'(x_t)(x_1 - x_t), \quad x_0, x_1 \in I.$$

Consequently,

$$\begin{aligned}(1-t)f(x_0) + tf(x_1) &\geq (1-t)[f(x_t) + f'(x_t)(x_0 - x_t)] + t[f(x_t) + f'(x_t)(x_1 - x_t)] \\ &= f(x_t) + f'(x_t)[(1-t)(x_0 - x_t) + t(x_1 - x_t)] \\ &= f(x_t) + f'(x_t)[(1-t)x_0 + tx_1] - f'(x_t)[x_t] \\ &\overset{x_t=(1-t)x_0+tx_1}{=} f((1-t)x_0 + tx_1),\end{aligned}$$

where the last equality comes from setting $x_t = (1-t)x_0 + tx_1$, concluding the proof.

☐

ix) f' is increasing (if f is differentiable).

Proof. (7 $\implies$ 9) By characterization 7, if $x_0 < x_1$ we have $f'_-(x_0) \leq f'_+(x_0) \leq f'_-(x_1) \leq f'_+(x_1)$, therefore f' is increasing. ☐

x) $f''(x) \geq 0$ (if f is twice differentiable).

Proof. (9 $\iff$ 10) Clearly f' is increasing if and only if f'' is non-negative.

Alternative approach: (8 $\iff$ 10) Using Lagrange Remainders, for every $x \in I$ there exists ξ between x and x_0 such that $f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(\xi)(x - x_0)^2$.

1) If $f''(x) \geq 0$ for every $x \in I$, then $\frac{1}{2}f''(\xi)(x - x_0)^2 \geq 0 \implies f(x) \geq f(x_0) + f'(x_0)(x - x_0)$.

2) Conversely, if $f(x) \geq f(x_0) + f'(x_0)(x - x_0) = T_1(x)$, then $0 \leq f(x) - T_1(x) = \frac{1}{2}f''(x_0)(x - x_0)^2 + o((x - x_0)^2)$ as $x \to x_0$. Therefore $\frac{1}{2}f''(x_0) = \lim_{x \to x_0} \frac{f(x)-T_1(x)}{(x-x_0)^2} \geq 0$ for every $x_0 \in I$, thanks to the permanence of sign in the limit. ☐

10.3 Additional Considerations

i) *Convexity and concavity*: A function is said *concave* if and only if $-f$ is convex. Therefore, every inequality is inverted for concave functions[41].

ii) *Affine functions:* Affine functions may be defined as the only functions both convex and concave in an interval, i.e. for which any inequality is indeed an equality. In fact, they do satisfy $f((1-t)x_0 + tx_1) = (1-t)f(x_0) + tf(x_1)$, which is the equality case of characterization (2). In other words, affine functions commute with convex combinations.

[41] Informally, convex functions are *smiling*, while concave functions are *sad*.

iii) *Strict convexity*: If $x_0 \neq x_1 \in I$ and $t \in (0,1)$, then $\boldsymbol{f((1-t)x_0 + tx_1) < (1-t)f(x_0) + tf(x_1)}$. It's clear that strict convexity implies strict inequalities almost everywhere, but there is a notable exception: a strictly convex function may have the second derivative equal to zero ($f(x) = x^4$ at $x = 0$).

iv) *Convexity and Differentiability:* A convex function in an <u>open</u> interval I is always <u>continuous</u> but not necessarily differentiable (take for instance $f(x) = |x|$ in the interval $(-1,1)$). However, the left and right derivatives always exist and satisfy $f'_-(x) \leq f'_+(x)$. Therefore, the derivative of a convex function has at most jump discontinuities.

v) *Convexity Test on Intervals:* Here we sum up relationships between derivatives and convexity, assuming the function differentiable or twice differentiable on that interval:

   i) f is convex (resp. concave) on (a,b) if and only if $f'(x)$ is increasing (resp. decreasing).
   ii) f is strictly convex (resp. concave) on (a,b) if and only if $f'(x)$ is strictly increasing (resp. decreasing).
   iii) f is convex (resp. concave) on (a,b) if and only if $f''(x) \geq 0 \; \forall x \in (a,b)$ (resp. $f''(x) \leq 0$).
   iv) f is strictly convex (resp. concave) on (a,b) if $f''(x) > 0 \; \forall x \in (a,b)$ (resp. $f''(x) < 0$). <u>the converse is false!</u>

vi) *Convexity at a Point:* Convexity is a property of a function defined on an interval. Therefore, any time we say f being convex at $x = x_0$, we are actually saying that f is convex in a neighbourhood of x_0.

vii) *Convexity Test on Points:* A common mistake is to forget that the behaviour of the derivatives of the function at a single point is rarely sufficient to make inferences about the convexity. Therefore, we need some additional assumptions. Section **Counterexamples** shows what happens without these assumptions.

   - If f is of class C^1, then f is convex at x_0 if and only if f' is increasing in a neighbourhood of x_0.
   - If f is of class C^2 and $f''(x_0) > 0$, then f is strictly convex at x_0 and the graph of f is above its tangent at x_0 (<u>the converse is false!</u>).
   - If f is of class C^2 and f is convex at x_0, then $f''(x_0) \geq 0$ (<u>the converse is false!</u>).

viii) *Convexity and local Minima:* Let f be differentiable in a neighbourhood I of x_0 and twice differentiable at x_0. If $f'(x_0) = 0$ and $f''(x_0) > 0$, then x_0 is a local minimizer for f. If $f''(x_0) < 0$, then x_0 is a local maximizer for f. If $f''(x_0) = 0$, we can't conclude.[42]

ix) *Convexity and global Minima:* Let f be convex and twice differentiable in an interval I, that is, $f''(x) \geq 0$ in I. If $f'(x_0) = 0$, then x_0 is a global minimizer. If the function is strictly convex, that is, $f''(x) > 0$ in I, then x_0 is the unique minimizer of f. Maxima of f an be attained only at the boundary of I, unless f is constant.

[Application: p-mean inequality, Holder Inequality, Cauchy Inequality, Young Inequality.]

10.4 Exercises

Example 1 Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous and convex function such that $f(x) \leq ax + b$ for every $x \geq 0$. Prove that the limit $l := \lim_{x\to+\infty} \frac{f(x)}{x}$ always exists and satisfies $l \leq a$. Show that if f is also differentiable, then $l = \lim_{x\to+\infty} f'(x)$.

Solution: By convexity, we know that the function $\varphi_{x_0}(x) = \frac{f(x)-f(x_0)}{x-x_0}$ is increasing for every $x \neq x_0 \in \mathbb{R}_0^+$. Choosing $x_0 = 0$ implies $\frac{f(x)-f(0)}{x}$ increasing for every $x > 0$. Therefore

$$l = \lim_{x\to+\infty} \frac{f(x)}{x} = \lim_{x\to+\infty} \frac{f(x)-f(0)}{x} + \frac{f(0)}{x} \overset{!}{=} \lim_{x\to+\infty} \frac{f(x)-f(0)}{x} + \lim_{x\to+\infty} \frac{f(0)}{x} = \lim_{x\to+\infty} \frac{f(x)-f(0)}{x},$$

where ! is justified since $\frac{f(x)-f(0)}{x}$ is increasing and therefore both limits exist. Notice that $f(x) \leq ax + b$ implies $\frac{f(x)-b}{x} \leq a$. Therefore:

$$l = \lim_{x\to+\infty} \frac{f(x)-f(0)}{x} = \lim_{x\to+\infty} \frac{f(x)-b}{x} + \frac{b-f(0)}{x} = \lim_{x\to+\infty} \frac{f(x)-b}{x} \leq a,$$

[42] These results come from a more general result on Taylor Expansions: in a neighbourhood of a critical point x_0, f behaves like the first non constant term of its Taylor Expansion.

by comparison test.
By convexity, if f is differentiable, then f' is increasing, and therefore $\lim_{x\to+\infty} f'(x)$ exists. Applying De l'Hôpital's Rule, we have $\lim_{x\to+\infty} f'(x) = \lim_{x\to+\infty} \frac{f(x)}{x} = l$.

Example 2: Let $f : \mathbb{R} \to \mathbb{R}$ be a concave differentiable function such that $f(0) = f(1) = f(2) = 0$. Prove that $f(x) = 0$ for every $x \in [0, 2]$.

Solution: Let $a \in (1, 2]$. Concavity implies that (draw a picture to convince yourself, it's characterization 5)

$$0 = \frac{f(2) - f(1)}{2 - 1} \le \frac{f(a) - f(1)}{a - 1} \le \frac{f(1) - f(0)}{1 - 0} = 0,$$

which implies $\frac{f(a)-f(1)}{a-1}$, that is, $f(a) = f(1) = 0$. Similarly, let $b \in [0, 1)$. Concavity implies that

$$0 = \frac{f(2) - f(1)}{2 - 1} \le \frac{f(1) - f(b)}{1 - b} \le \frac{f(1) - f(0)}{1 - 0} = 0,$$

that is, $f(b) = f(1) = 0$. Since a and b are arbitrary, we conclude that $f(x) = 0$ for every $x \in [0, 2]$.

Second Solution: By concavity, for any $t \in [0, 1]$,

$$(1 - t)f(0) + tf(1) \le f((1 - t) \cdot 0 + t \cdot 1) \quad \text{and} \quad tf(1) + (1 - t)f(2) \le f(t \cdot 1 + (1 - t) \cdot 2),$$

that is, $0 \le f(t)$ and $0 \le f(2 - t)$. On the other hand, by concavity,

$$0 \le \frac{1}{2}f(t) + \frac{1}{2}f(2 - t) \le f\left(\frac{t + 2 - t}{2}\right) = f(1) = 0.$$

Combining these two results, we conclude that, for any $t \in [0, 1]$, $f(t) = 0$ and $f(2 - t) = 0$, and thus $f(x) = 0$ for every $x \in [0, 2]$.

11 Integral Calculus

Given a **bounded**[43], non-negative function f on a **bounded** interval $[a, b]$, the area of the region $R = R(f) := \{(x, y) \in [a, b] \times \mathbb{R} : 0 \le y \le f(x)\}$ is the area between the function and the x-axis.
The rigorous definition of this area is based on a few straightforward assumptions: firstly, that the area of a rectangle is the product of the lengths of its sides; secondly, that the area of a disjoint union of regions equals the sum of the areas of each individual region; and lastly, that the area of a subregion cannot exceed that of the entire region. Building on this foundation, the aim is to give a proper definition of the area $\mathcal{A}(R)$.

11.1 (Darboux) Construction of the Riemann Integral

Definition 11.1 (Partition). A *partition* $\mathcal{P}$ of an interval $[a, b]$ is an ordered $N + 1$-uple of real numbers $a = x_0 < x_1 < \ldots < x_N = b$ for which we associate the intervals $I_n := [x_{n-1}, x_n]$ which, together, form a partition of $[a, b]$ up to a negligible set. We denote by $P(a, b)$ the collection of all partitions of $[a, b]$.

Remark 11.2. $\Delta x_n = x_n - x_{n-1}$ denotes the length of I_n. We often use a **uniform partition**, so that all the intervals I_n have the same length and $\Delta x_n = \frac{b-a}{N}$.

Definition 11.3 (Plurirectangle). Given a partition $\mathcal{P}$, the plurirectangle associated with it is the union of rectangles of the form $I_n \times J_n$, where J_n is an interval contained in $[0, \infty)$.[44]

Definition 11.4 (Lower and upper (Riemann) sums). Given a partition $\mathcal{P}$ of $[a, b]$, we define the lower and upper Riemann sums as

$$\mathcal{L}(f, \mathcal{P}) = \sum_{i=1}^{N} \left(\inf_{x \in I_n} f(x)\right) \cdot \Delta x_n = \sum_{i=1}^{N} l_n \Delta x_n \quad \text{and} \quad \mathcal{U}(f, \mathcal{P}) = \sum_{i=1}^{N} \left(\sup_{x \in I_n} f(x)\right) \cdot \Delta x_n = \sum_{i=1}^{N} u_n \Delta x_n$$

[43] $|f(x)| \le C$ for some $C \in \mathbb{R}$
[44] more formally: a plurirectangle is a step function, a linear combination of indicator functions of $I_0, \ldots, I_N$

Notice that, no matter the partition, **every lower** Riemann sum is less than or equal to **any upper** Riemann sum[45], that is,

$$\inf_{[a,b]} f \cdot (b-a) \le \mathcal{L}(f,\mathcal{P}_1) \le \mathcal{U}(f,\mathcal{P}_2) \le \sup_{[a,b]} f \cdot (b-a), \quad \text{for every } \mathcal{P}_1, \mathcal{P}_2 \in P(a,b)$$

Definition 11.5 (Riemann-Integrability). A **bounded** function in a closed bounded interval $[a,b]$ is Riemann-integrable if

$$\underbrace{\sup_{\mathcal{P}\in P(a,b)} \mathcal{L}(f,\mathcal{P})}_{\text{lower integral } I^-} = \underbrace{\inf_{\mathcal{P}\in P(a,b)} \mathcal{U}(f,\mathcal{P})}_{\text{upper integral } I^+} \overset{\text{def}}{=:} \underbrace{\int_a^b f(x)\mathrm{d}x}_{\text{(proper) integral of } f} .$$

Remark 11.6. Boundedness is used to assure that the lower and upper integrals are the supremum and infimum of nonempty sets. If the function is unbounded, it may occur that the set of lower / upper Riemann sums is empty.

Remark 11.7 (Criterion for integrability). Definition 11.5 is equivalent to the following *criterion for integrability:* f is integrable if and only if, for every $\varepsilon > 0$, there exists a partition $\mathcal{P}_\varepsilon$ such that

$$\mathcal{U}(f,\mathcal{P}_\varepsilon) - \mathcal{L}(f,\mathcal{P}_\varepsilon) \le \varepsilon.$$

Example 11.8. The **Dirichlet function** $f : [0,1] \to \mathbb{R}$ defined as the indicator of the rational numbers

$$f(x) = \mathbf{1}_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

is the usual, pathological example of non-integrability. No matter the partition, the upper integral I^+ is 1 and the lower integral I^- is 0.

11.2 Integrable Functions

Which families of functions are integrable? Restricting to bounded functions on compact intervals, a very general result known as *Lebesgue-Vitali* theorem states that a function is Riemann integrable if and only if it is continuous almost everywhere. Since we don't have the tools to prove this theorem, the next results will characterize *most* of the integrable functions.

Theorem 11.9 (Integrable functions). *A bounded function on a compact interval* $[a,b]$ *is Riemann integrable if it has a finite number of discontinuities or is a monotone function.*

Remark 11.10. in particular, continuous functions and Lipschitz functions are Riemann integrable.

Remark 11.11. The converse is false. For example,

$$\int_0^\pi f(x) = \int_0^\pi \sum_{i=1}^\infty \frac{\sin(nx)}{n^2} = \frac{7\zeta(3)}{4} \approx 1.05,$$

and $\sum_{i=1}^\infty \frac{\sin(nx)}{n^2}$ is a function with an infinite (but countable) number of discontinuities.

[exercise: try to show that $\int_0^1 x\mathrm{d}x = \frac{1}{2}$ using the definition of Riemann integral.]

Proposition 11.12 (Continuity and Integrability). *continuous functions on a compact interval* $[a,b]$ *are (Riemann) integrable.*

Proof.
Since $[a,b]$ is closed and bounded, f is uniformly continuous (Heine's Theorem), so that (5.1) holds.
Now consider a uniform partition $\mathcal{P}_\varepsilon$ of size $\Delta x_n = (b-a)/N$, selecting N large enough so that $\Delta x_n = (b-a)/N < \delta$.
In any compact interval I_n we can apply Weierstrass Theorem, finding two points m_n and M_n such

[45] This property is hard to prove but intuitive. graphically, the upper Riemann sum is the area of the smallest plurirectangle associated with $\mathcal{P}$ **containing** R, and the lower Riemann sum is the area of the largest plurirectangle associated with $\mathcal{P}$ **contained in** R.

that

$$l_n = \inf_{I_n} f = \min_{I_n} f = f(m_n) \quad \text{and} \quad u_n = \sup_{I_n} f = \max_{I_n} f = f(M_n).$$

Since

$$|M_n - m_n| \leq \Delta x_n = (b-a)/N < \delta,$$

then

$$|u_n - l_n| = |f(M_n) - f(m_n)| \leq \varepsilon,$$

by uniform continuity. Therefore

$$\mathcal{U}(f, \mathcal{P}_\varepsilon) - \mathcal{L}(f, \mathcal{P}_\varepsilon) = \sum (u_n - l_n)\Delta x_n = \frac{b-a}{N} \sum u_n - l_n \leq \frac{b-a}{N} \cdot \sum \varepsilon = (b-a)\varepsilon.$$

Setting $\varepsilon' = \varepsilon/(b-a)$ we conclude. $\square$

Proposition 11.13 (Lipschitz and Integrability). *Lipschitz functions on a compact interval* $[a, b]$ *are (Riemann) integrable.*

Proof.
Since Lipschitz functions are continuous, this proposition follows from the previous. However, we propose a direct proof: we apply Weierstrass Theorem as in the previous proposition, finding points m_n and M_n in every I_n such that $u_n = f(M_n)$ and $l_n = f(m_n)$. Therefore

$$u_n - l_n = f(M_n) - f(m_n) \overset{1}{\leq} L|M_n - m_n| \overset{2}{\leq} L\Delta x_n,$$

where the first inequality is the Lipschitz property and the second is just the fact that both M_n and m_n belong to the interval $[x_{n-1}, x_n]$. Therefore, choosing a uniform partition,

$$\mathcal{U}(f, \mathcal{P}) - \mathcal{L}(f, \mathcal{P}) = \Delta x_n \sum_{n=1}^{N} u_n - l_n \leq \Delta x_n \sum_{n=1}^{N} L\Delta x_n = (\Delta x_n)^2 LN = \left(\frac{b-a}{N}\right)^2 LN = \frac{(b-a)^2 L}{N},$$

which is arbitrarily small as N is sufficiently large. $\square$

Proposition 11.14 (Monotonicity and Integrability). *Monotone functions on a compact interval* $[a, b]$ *are (Riemann) integrable.*

Proof.
Suppose f is increasing. Fix $\varepsilon > 0$ and for $N \in \mathbb{N}$ consider the **uniform** partition so that $\Delta x_n = \frac{b-a}{N}$. Since f is increasing, we have $u_n = f(x_n)$ and $l_n = f(x_{n-1})$. Therefore

$$\begin{aligned}
\boldsymbol{\mathcal{U}(f, \mathcal{P}_\varepsilon) - \mathcal{L}(f, \mathcal{P}_\varepsilon)} &= \sum u_n \Delta x_n - \sum l_n \Delta_n \\
&= \Delta x_n \sum (u_n - l_n) = \boldsymbol{\Delta x_n \sum (f(x_n) - f(x_{n-1}))} \\
&= \Delta x_n (\cancel{f(x_1)} - f(x_0) + \cancel{f(x_2)} - \cancel{f(x_1)} + \cancel{\dots} + f(x_n) - \cancel{f(x_{N-1})}) \\
&= \boldsymbol{\frac{b-a}{N}(f(x_n) - f(x_0))} = \frac{b-a}{N}(f(b) - f(a)),
\end{aligned}$$

which is arbitrarely small as N is sufficiently large. $\square$

We conclude this section with a list of properties on definite integrals, which are almost identical also for indefinite integrals.

i) *Linearity:* If f, g are integrable, then

$$\int_a^b (\alpha f(x) + \beta g(x))\mathrm{d}x = \alpha \int_a^b f(x)\mathrm{d}x + \beta \int_a^b g(x)\mathrm{d}x;$$

ii) *Additivity:* f is integrable in $[a, b]$ if and only if for every $c, d, e \in [a, b]$:

$$\int_c^e f(x)\mathrm{d}x = \int_c^d f(x)\mathrm{d}x + \int_d^e f(x)\mathrm{d}x$$

iii) *Monotonicity:* If $f \geq g$ then

$$\int_a^b f(x)\mathrm{d}x \geq \int_a^b g(x)\mathrm{d}x$$

iv) *Triangle Inequality:* If f is integrable then $|f|$ is integrable and

$$\left| \int_a^b f(x)\mathrm{d}x \right| \leq \int_a^b |f(x)|\mathrm{d}x$$

v) *Symmetry:*

$$\int_{-a}^{-b} f(x)\mathrm{d}x = \int_a^b f(-x)\mathrm{d}x$$

vi) *Parity:*

$$\int_{-a}^{a} f(x)\mathrm{d}x = \begin{cases} 0 & \text{if } f \text{ is odd} \\ 2\int_0^a f(x)\mathrm{d}x & \text{if } f \text{ is even} \end{cases}$$

Remark 11.15 (Oriented Intervals). In order to extend the previous properties and definitions for functions that take negative values and intervals $[a, b]$ where $a > b$ it is convenient to set

$$\int_b^a f(x)\mathrm{d}x := -\int_a^b f(x)\mathrm{d}x,$$

hence $\int_c^c f(x)\mathrm{d}x = 0$. Moreover, the area under the curve f is signed: by linearity, if $f < 0$ in $[a, b]$, then $g := -f > 0$, hence

$$\int_a^b f(x)\mathrm{d}x = \int_a^b -g(x)\mathrm{d}x = -\int_a^b g(x)\mathrm{d}x.$$

11.3 Cauchy-Riemann Integral

We now present a different yet equivalent[46] construction for the Riemann Integral, and two additional criteria for integrability. For any partition $\mathcal{P} \in P(a, b)$ we consider a set of N points $\boldsymbol{S_{\mathcal{P}} = \{\xi_1, \ldots, \xi_N\}}$, where $\xi \in I_n$ for each n. The Cauchy-Riemann Integral sum is defined as

$$\mathcal{R}(f, \mathcal{P}, S_{\mathcal{P}}) := \sum_{n=1}^{N} f(\xi_n)\Delta x_n$$

Unlike the previous approach of considering the supremum u_n and the infimum l_n in each interval, we now consider any representing value $f(\xi_n)$, which, clearly, lies between u_n and l_n. Therefore, regardless of the partition and choice of $S_{\mathcal{P}}$, **the Cauchy-Riemann sum always falls between the upper Riemann sum and the lower Riemann sum**. More precisely,

$$\inf_{S_{\mathcal{P}}} \mathcal{R}(f, \mathcal{P}, S_{\mathcal{P}}) = \mathcal{L}(f, \mathcal{P}) \quad \text{and} \quad \sup_{S_{\mathcal{P}}} \mathcal{R}(f, \mathcal{P}, S_{\mathcal{P}}) = \mathcal{U}(f, \mathcal{P}).$$

Lemma 11.16 (Criterion for integrability (2)). *A bounded function $f : [a, b] \to \mathbb{R}$ is integrable if and only if for every $\varepsilon > 0$ there exists a partition $\mathcal{P}_\varepsilon$ such that, for any choice of $S'_{\mathcal{P}}$ and $S''_{\mathcal{P}}$, it holds that*

$$|\mathcal{R}(f, \mathcal{P}_\varepsilon, S'_{\mathcal{P}}) - \mathcal{R}(f, \mathcal{P}_\varepsilon, S''_{\mathcal{P}})| \leq \varepsilon$$

Lemma 11.17 (Integral as a limit). *A bounded function $f : [a, b] \to \mathbb{R}$ is integrable if and only if there exists the limit*

$$\lim_{|\mathcal{P}| \to 0} \mathcal{R}(f, \mathcal{P}, S_{\mathcal{P}}).$$

[46] The equivalence of the two constructions is a strong theorem which is given without a proof.

In such case, it coincides with the usual definition of Riemann integral of f in $[a, b]$.

Remark 11.18. $|\mathcal{P}|$ is the size of the partition, formally defined as the largest length Δx_n of the intervals I_n

Remark 11.19. Remark 2: the limit notation indicates that $\mathcal{R}(f, \mathcal{P}, S_{\mathcal{P}})$ converges to a fixed value I (i.e. $|\mathcal{R}(f, \mathcal{P}, S_{\mathcal{P}}) - I| < \varepsilon$) as the size of the partition approaches 0 (i.e. $|\mathcal{P}| < \delta$), *regardless* of the choice of S.

Remark 11.20. The usefulness of this proposition lies in the fact that to compute the limit, it is sufficient to select a partition $\mathcal{P}_k$ that approaches 0 as k increases, along with a set of points $S_{\mathcal{P}_k}$ for which the Cauchy-Riemann sum is easy to compute, and then compute the limit as $k \to \infty$. **[Integrate x^2 with this approach.]**

11.4 Primitives

In this chapter we will understand why integration and differentiation are inverses of one another, stating the two versions of the Fundamental Theorem of Calculus.

Definition 11.21 (Primitive or Antiderivative). Given $f : [a, b] \to \mathbb{R}$, where $[a, b]$ is a closed an bounded interval, a *primitive* of f in $[a, b]$ is any function F such that $F'(x) = f(x)$ for every $x \in (a, b)$.

Remark 11.22. By definition, primitive functions of f in $[a, b]$ must be *differentiable* in (a, b). No explicit requirement is stated in $[a, b]$. However, for most applications we also assume continuity in $[a, b]$. Moreover, a simple application of Theorem 8.29 implies that all primitives of f in an *interval* $[a, b]$ differ by a constant.

Definition 11.23 (Indefinite integral). Given $f : \mathbb{R} \to \mathbb{R}$, its *indefinite integral* is the collection of all the primitives of a function in a given interval, that is,

$$\underbrace{\int f(x)\mathrm{d}x}_{\text{family of primitives}} \quad := \quad \underbrace{F(x)}_{\text{a primitive}} \quad + \quad \underbrace{c}_{\text{constant}} \quad .$$

By remark 11.22, the indefinite integral is a class of primitives of f in a given interval that differ by a constant $c \in \mathbb{R}$. It does not coincide with the set of all primitives of f, as the following example shows:

Example 11.24. We usually refer to primitives and antiderivatives in an interval. Instead, if we consider the function $f(x) = \frac{1}{x}$, which is not defined in an interval, the general primitive has the form

$$F(x) = \begin{cases} \log(-x) + c_1 & \text{if } x < 0 \\ \log(x) + c_2 & \text{if } x > 0 \end{cases}.$$

In such case, all the primitives do not differ by a constant, and the indefinite integral does not coincide with the family of all primitives of f.

Definition 11.25 (Integral Function). Let f be a function defined on an interval I that is Riemann integrable in every bounded subinterval $[a, b] \subset I$. Fix $c \in I$. The integral function of f at I is the function

$$F(x) := \int_c^x f(t)\mathrm{d}t.$$

Remark 11.26. The integral function is Lipschitz continuous on every bounded interval. If f is also continuous, its integral function is differentiable and of class C^1, by Theorem 11.30. The integral function of f exists even if it may not be the antiderivative of f.**[Example]**

11.5 Theorems on Integrable functions in an interval

Theorem 11.27 (Fundamental Theorem of Calculus 1 (FTC 1)). *Let $f : [a, b] \to \mathbb{R}$ be integrable and F be a primitive for f in $[a, b]$, except at most for a finite number of points*[47]*. Then*

$$\int_a^b f(x)\mathrm{d}x = F(b) - F(a) = F(x)\Big|_{x=a}^{x=b} \tag{11.1}$$

[47] Meaning that F is continuous in $[a, b]$ and it's differentiable in $[a, b]$, with $F' = f$, except for a finite number of points, for which F may be not differentiable, but it must be there continuous.

Proof.
Given an arbitrary partition $\mathcal{P} = \{x_0, \dots, x_N\}$ of $[a, b]$ which contains every point where F is not differentiable, then we can apply Theorem 8.26 on the function F in each interval $I_n = [x_{n-1}, x_n]$ finding points $\xi_n \in (x_{n-1}, x_n)$ such that

$$F'(\xi_n) = \frac{F(x_n) - F(x_{n-1})}{x_n - x_{n-1}} \implies F'(\xi_n)\Delta x_n = F(x_n) - F(x_{n-1}).$$

Choosing $S_{\mathcal{P}} := (\xi_1, \dots, \xi_n)$ we get

$$\begin{aligned}\int_a^b f(x)\mathrm{d}x = \mathcal{R}(f, \mathcal{P}, S_{\mathcal{P}}) = \sum_{n=1}^{N} f(\xi_n)\Delta x_n = \sum_{n=1}^{N} F'(\xi_n)\Delta x_n = \sum_{n=1}^{N} F(x_n) - F(x_{n-1}) \\ = \cancel{F(x_1)} - F(x_0) + \cancel{F(x_2)} - \cancel{F(x_1)} + \cancel{\dots} + F(x_N) - \cancel{F(x_{N-1})} = F(b) - F(a).\end{aligned}$$

□

Theorem 11.28 (Integral mean value theorem). *If f is continuous in $[a, b]$, then there exists $\xi \in (a, b)$ such that $\int_a^b f(x)\mathrm{d}x = (b - a)f(\xi)$.*

Proof.
Weierstrass' theorem guarantees the existence of a minimum m and a maximum M in $[a, b]$. By monotonicity of the integral function,

$$m(b-a) = \int_a^b m\mathrm{d}x = \int_a^b \inf_{[a,b]} f\mathrm{d}x \leq \int_a^b f(x)\mathrm{d}x \leq \int_a^b \sup_{[a,b]} f\mathrm{d}x = \int_a^b M\mathrm{d}x = M(b-a).$$

In particular,

$$m \leq \frac{1}{b-a}\int_a^b f(x)\mathrm{d}x \leq M.$$

By Theorem 7.1, there exists a point $\xi \in (a, b)$ such that $f(\xi) = \frac{1}{b-a}\int_a^b f(x)\mathrm{d}x$. □

Remark 11.29 (Intermediate Value Theorems). Let us recap the three "intermediate value" theorems on a given function $f : [a, b] \to \mathbb{R}$ continuous:

i) *Bolzano:* if $y^\star \in [\inf_{[a,b]} f, \sup_{[a,b]} f]$, there exists $\xi \in [a, b]$ such that

$$f(\xi) = y^\star;$$

ii) *Lagrange:* if f is differentiable in (a, b), there exists $\xi \in (a, b)$ such that

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

iii) *Integral:* there exists $\xi \in (a, b)$ such that

$$f(\xi) = \frac{\int_a^b f(x)\mathrm{d}x}{b - a}$$

Theorem 11.30 (Fundamental Theorem of Calculus 2 (FTC 2)). *If $f : I = [a, b] \to \mathbb{R}$ is integrable on I, and f is **continuous**[48] in an interior point x_0, then the integral function $F(x) := \int_c^x f(t)\mathrm{d}t$, with $c \in I$, is **differentiable** at x_0 and*

$$F'(x_0) = f(x_0).$$

In particular, if f is continuous in I, then the integral function is a primitive of f in I (and any other primitive is up to a constant).[49]

[48] pay attention to the difference! The first FTC only requires integrability and the existence of an (almost) primitive of f. Instead, the second FTC requires the continuity of f.
[49] FTC 2 gives a sufficient condition to have a primitive. There's no easy necessary condition to assure the existence of a

Proof.
Consider the right derivative:

$$F'_+(x_0) = \lim_{h\to 0^+} \frac{F(x_0+h)-F(x_0)}{h} = \lim_{h\to 0^+} \frac{1}{h}\left(\int_c^{x_0+h} f(t)dt - \int_c^{x_0} f(t)dt\right) = \lim_{h\to 0+} \frac{1}{h}\cdot\int_{x_0}^{x_0+h} f(t)\mathrm{d}t.$$

If f is continuous in I, we can apply Theorem 11.28 on the interval $[x_0, x_0+h]$, finding $\varepsilon_h \in (x_0, x_0+h)$ such that

$$f(\varepsilon_h) = \frac{1}{h}\int_{x_0}^{x_0+h} f(t)\mathrm{d}t.$$

As $h \to 0^+$, ε_h is a sequence of points converging to x_0^+. Hence, by transference principle,

$$F'_+(x_0) = \lim_{h\to 0^+} \frac{1}{h}\int_{x_0}^{x_0+h} f(t)\mathrm{d}t = \lim_{\varepsilon_h \to x_0^+} f(\varepsilon_h) = f(x_0).$$

Similarly, one proves $F'_-(x_0) = f(x_0)$, concluding that $F'(x_0) = f(x_0)$. □

Remark 11.31. If f is continuous only at x_0, the argument is trickier. We want to bound the difference $|F'_+(x_0) - f(x_0)|$ using continuity. An inequality of the form $|F'_+(x_0) - f(x_0)| \leq \varepsilon$, for any $\varepsilon > 0$, would conclude the proof. Indeed,

$$|F'_+(x_0) - f(x_0)| = \left|\frac{1}{h}\int_{x_0}^{x_0+h} f(t)\mathrm{d}t - f(x_0)\right| = \left|\frac{1}{h}\int_{x_0}^{x_0+h}(f(t)\mathrm{d}t - \frac{1}{h}\int_{x_0}^{x_0+h} f(x_0))\mathrm{d}t\right| = \frac{1}{h}\left|\int_{x_0}^{x_0+h} f(t) - f(x_0)\mathrm{d}t\right|$$
$$\overset{1}{\leq} \frac{1}{h}\int_{x_0}^{x_0+h} |f(t) - f(x_0)|\mathrm{d}t \overset{2}{\leq} \frac{1}{h}\int_{x_0}^{x_0+h} \varepsilon\mathrm{d}t = \frac{h\varepsilon}{h} = \varepsilon,$$

where 1 is the triangle inequality and 2 holds by continuity of f at x_0, with δ-close to x_0.

Theorem 11.32 (Integration by parts). [50] *If F, G are differentiable on an interval I, $f = F'$ and $g = G'$, then if fG has a primitive in I (for example, f is continuous), Fg has a primitive in I and*

$$\int F(x)g(x)\mathrm{d}x = F(x)G(x) - \int f(x)G(x)\mathrm{d}x \tag{11.2}$$

Proof.
By the product rule,

$$(FG)' = F'G + FG' = fG + Fg.$$

If fG has a primitive H on I, i.e., $H' = fG$, then

$$Fg = (FG)' - fG = (FG)' - H' = (FG - H)'.$$

Hence Fg has an antiderivative $FG - H$ and, integrating on both sides we recover (11.2). □

Theorem 11.33 (Integration by substitution). *If $\varphi : \mathcal{Y} \to \mathcal{X}$ is differentiable, $\mathcal{X}$ and $\mathcal{Y}$ are <u>intervals</u>, f has an antiderivative and $x = \varphi(y)$, then*

$$\int_a^b f(x)\mathrm{d}x = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(y))\varphi'(y)\mathrm{d}y.$$

primitive, and thus to disprove that a function can have it. However, we can intuitively argue that $f(x) = \mathrm{sign}(x)$ does not have a primitive in $[-1, 1]$, since it would be $|x| + c$, which however is not differentiable at $x = 0$.

[50] Look what happens if we integrate by parts $\tan x$:

$$\int \tan(x)\mathrm{d}x = \int \sin x \cdot \frac{1}{\cos x}\mathrm{d}x = \frac{-\cos x}{\cos x} - \int (-\cancel{\cos x})\left(-\frac{1}{\cos^{\cancel{2}}(x)}\right)(-\sin x)\mathrm{d}x = -1 + \int \tan(x)\mathrm{d}x.$$

This is a completely legitimate result, since $\int f(x)\mathrm{d}x$ denotes a set of primitives. Therefore, it can legitimally differ from itself by a constant, since the above equality is an equality between sets.

Proof.
Let F be a primitive of f at $\mathcal{X}$ and $G := F \circ \varphi$ on $\mathcal{Y}$. By the *chain rule*,

$$G'(y) = F'(\varphi(y)) \cdot \varphi'(y) = f(\varphi(y))\varphi'(y).$$

Hence, by the Fundamental Theorem of Calculus,

$$\int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f(\varphi(y))\varphi'(y)\mathrm{d}y = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} G'(y)\mathrm{d}y = G(\varphi^{-1}(b)) - G(\varphi^{-1}(a)) = F(b) - F(a) = \int_a^b f(x)\mathrm{d}x.$$

$\square$

11.6 Exercises

Example 1: trivial integrals

$$\int \tan^2(x)\,dx = \int (\tan^2(x) + 1 - 1)\,dx = \tan(x) - x + c$$

$$\int \tan(x)\,dx = \int \frac{\sin x}{\cos x}\,dx = -\log(|\cos x|) + c$$

$$\int \frac{\sin x - \cos x}{\sin x + \cos x}\,dx = -\log(|\sin x + \cos x|) + c$$

$$\int \sin^2(x)\,dx = \int \left(\frac{1}{2} - \frac{\cos(2x)}{2}\right)\,dx = \frac{1}{2}x - \frac{1}{4}\sin(2x) + c$$

$$\int \frac{e^x + e^{-x}}{e^x - e^{-x}}\,dx = \log(e^x - e^{-x}) + c$$

$$\int e^{x+e^x}\,dx = e^{e^x} + c$$

$$\int \sin(x)\cos(x)\,dx = \frac{1}{2}\sin^2(x) + c$$

$$\int \frac{1}{x(1+\log^2(x))}\,dx = \arctan(\log x) + c$$

$$\int \frac{\sin(\log x)}{x}\,dx = -\cos(\log x) + c$$

$$\int \frac{x}{\sqrt{1-x^4}}\,dx = \frac{1}{2}\arcsin(x^2) + c$$

Example 2: Integration by parts this shows how to integrate $\int P(x)e^{\alpha x}\mathrm{d}x$, for any polynomial P.

$$\begin{aligned}\int x^2 e^{7x}\mathrm{d}x &= x^2\left(\int e^{7x}\mathrm{d}x\right) - \int 2x\left(\int e^{7x}\mathrm{d}x\right)\mathrm{d}x = \frac{1}{7}x^2e^{7x} - \frac{2}{7}\int xe^{7x}\mathrm{d}x \\ &= \frac{x^2}{7}e^{7x} - \frac{2}{7}\left(\frac{1}{7}xe^{7x} - \int \frac{1}{7}e^{7x}\mathrm{d}x\right) = e^{7x}\left(\frac{x^2}{7} - \frac{2x}{49} + \frac{2}{343}\right) + c\end{aligned}$$

Example 3: Integration by parts this shows how to integrate $\int P(x)\sin(\alpha x)\mathrm{d}x$ and $\int P(x)\cos(\alpha x)\mathrm{d}x$, for any polynomial P.

$$\begin{aligned}\int x^2\cos(3x)\mathrm{d}x &= x^2 \cdot \left(\int \cos(3x)\mathrm{d}x\right) - \int 2x \cdot \left(\int \cos(3x)\mathrm{d}x\right)\mathrm{d}x \\ &= \frac{x^2}{3}\cdot\sin(3x) - \frac{2}{3}\int x\cdot\sin(3x)\mathrm{d}x = \frac{x^2}{3}\cdot\sin(3x) - \frac{2}{3}\left[x\left(\int\sin(3x)\mathrm{d}x\right) - \int\left(\int\sin(3x)\mathrm{d}x\right)\mathrm{d}x\right] \\ &= \frac{x^2}{3}\cdot\sin(3x) - \frac{2}{3}\left[-\frac{x}{3}\cdot\cos(3x) + \frac{1}{3}\int\cos(3x)\mathrm{d}x\right] = \frac{x^2}{3}\cdot\sin(3x) + \frac{2x}{9}\cdot\cos(3x) - \frac{2}{27}\sin(3x) + c\end{aligned}$$

Example 4: the great return this is how to integrate $\int e^{\alpha x}\sin(\beta x)\mathrm{d}x$ and $\int e^{\alpha x}\cos(\beta x)\mathrm{d}x$. For example, to integrate $e^{2x}\cos x$, let us call $I := \int e^{2x}\cos x\mathrm{d}x$. Let's integrate I by applying integration by parts twice:

$$\begin{aligned}I &= \int e^{2x}\cos x\mathrm{d}x = e^{2x}\sin x - \int 2e^{2x}\sin x\mathrm{d}x = e^{2x}\sin x - 2\left[e^{2x}(-\cos x) - \int 2e^{2x}(-\cos x)\mathrm{d}x\right] \\ &= e^{2x}(\sin x + 2\cos x) - 4\int e^{2x}\cos x\mathrm{d}x = e^{2x}(\sin x + 2\cos x) - 4I \\ &\implies 5I = e^{2x}(\sin x + 2\cos x) \implies I = \frac{e^{2x}}{5}\cdot(\sin x + 2\cos x) + c\end{aligned}$$

Example 5: the hidden 1 this shows how to integrate by parts $\log x$ and $\arctan x$:

$$\int \log x \mathrm{d}x = \int 1 \cdot \log x \mathrm{d}x = x \log x - \int x \cdot \frac{1}{x} \mathrm{d}x = x \log x - \int 1 \mathrm{d}x = x \log x - x + c$$

$$\int \arctan x \mathrm{d}x = x \arctan x - \int x \cdot \frac{1}{1+x^2} \mathrm{d}x = x \arctan x - \frac{1}{2} \int \frac{2x}{1+x^2} \mathrm{d}x = x \arctan x - \frac{1}{2} \log(1+x^2) + c$$

Example 6: rational functions this shows how to integrate (easy) rational functions:

$$\int \frac{x}{(x^2+1)^3} \mathrm{d}x = \frac{1}{2} \int \frac{2x}{(x^2+1)^3} = \frac{1}{2} \int y^{-3} \mathrm{d}y = \frac{1}{2} \left(-\frac{1}{2y^2} \right) + c = -\frac{1}{4(x^2+1)^2} + c$$

$$\begin{aligned} \int \frac{3x+2}{x^2+2x+2} \mathrm{d}x &= \int \frac{3x+3-2}{x^2+2x+2} \mathrm{d}x = \frac{3}{2} \int \frac{2x+2}{x^2+2x+2} \mathrm{d}x - \int \frac{1}{x^2+2x+2} \mathrm{d}x \\ &= \frac{3}{2} \int \frac{2x+2}{x^2+2x+2} \mathrm{d}x - \int \frac{1}{(x+1)^2+1} \mathrm{d}x = \frac{3}{2} \log(x^2+2x+2) - \arctan(x+1) + c \end{aligned}$$

$$\begin{aligned} \int \frac{x^4-2x^2+4x+1}{x^3-x^2-x+1} \mathrm{d}x &= \int \frac{(x^4-x^3-x^2+x)+(x^3-x^2+3x+1)}{x^3-x^2-x+1} \mathrm{d}x = \int x + \frac{(x^3-x^2-x+1)+4x}{x^3-x^2-x+1} \mathrm{d}x \\ &= \int x + 1 + \frac{4x}{(x-1)^2(x+1)} \mathrm{d}x = \frac{x^2}{2} + x + \int \frac{A}{x+1} + \frac{B}{x-1} + \frac{C}{(x-1)^2} \mathrm{d}x \\ &= \frac{x^2}{2} + x + \int -\frac{1}{x+1} + \frac{1}{x-1} + \frac{2}{(x-1)^2} \mathrm{d}x = \frac{x^2}{2} + x + \log\left|\frac{x-1}{x+1}\right| - \frac{2}{x-1} + c \end{aligned}$$

Example 7: some simple substitutions

$$\int \frac{e^{2x}}{1+e^x} \mathrm{d}x \overset{y=e^x}{=} \int \frac{y}{1+y} \mathrm{d}y = \int \frac{y+1-1}{1+y} \mathrm{d}y = \int 1 \mathrm{d}y - \int \frac{1}{1+y} \mathrm{d}y = y - \log|1+y| + c = e^x - \log(1+e^x) + c$$

$$\begin{aligned} \int \frac{1}{x+\sqrt{x-3}} \mathrm{d}x &\overset{y=\sqrt{x-3}}{=} \int \frac{1}{y^2+3+y} \cdot 2y \mathrm{d}y = \int \frac{2y+1-1}{y^2+y+3} \mathrm{d}y = \int \frac{2y+1}{y^2+y+3} \mathrm{d}y - \int \frac{1}{\left(y+\frac{1}{2}\right)^2 + 3 - \frac{1}{4}} \mathrm{d}y \\ &= \log(y^2+y+3) - \int \frac{1}{\left(y+\frac{1}{2}\right)^2 + \frac{11}{4}} \mathrm{d}y = \log(y^2+y+3) - \frac{4}{11} \int \frac{1}{\left(\frac{y+1/2}{\sqrt{11}/2}\right)^2 + 1} \mathrm{d}y \\ &= \log(y^2+y+3) - \frac{2}{\sqrt{11}} \int \frac{2/\sqrt{11}}{\left(\frac{2y+1}{\sqrt{11}}\right)^2 + 1} \mathrm{d}y = \log(y^2+y+3) - \frac{2}{\sqrt{11}} \arctan\left(\frac{2y+1}{\sqrt{11}}\right) + c \\ &= \log(x+\sqrt{x-3}) - \frac{2}{\sqrt{11}} \arctan\left(\frac{2\sqrt{x-3}+1}{\sqrt{11}}\right) + c \end{aligned}$$

$$\int x \sqrt[3]{x+2} \mathrm{d}x \overset{y=x+2}{=} \int (y-2) \sqrt[3]{y} \mathrm{d}y = \int y^{4/3} - 2y^{1/3} \mathrm{d}y = \frac{3}{7}(x+2)^{7/3} - \frac{3}{2}(x+2)^{4/3} + c$$

Example 8: oddness If $f(x) = -f(-x)$, the function is odd and in such case $\int_{-a}^{a} f(x) \mathrm{d}x = 0$:

$$\int_{-1}^{1} \sin\left(\log(1+x^6)\cos(-x)x^3\right) + e^x \mathrm{d}x = \int_{-1}^{1} e^x \mathrm{d}x = e - e^{-1}.$$

11.7 Improper Integrals

Improper integrals occur when we integrate over unbounded intervals (for example, $\int_2^{+\infty} \frac{1}{x^2} \mathrm{d}x$), we integrate unbounded functions (for example, $\int_0^2 \frac{1}{x} \mathrm{d}x$), or both (for example, $\int_0^{+\infty} \frac{1}{x^3+x} \mathrm{d}x$). Such integrals are defined via limits and, therefore, they can be convergent, divergent to $\pm\infty$, or divergent (indeterminate, like $\int_0^{\infty} \cos(x) \mathrm{d}x$). Notice that $\int_0^2 \sin\left(\frac{1}{x}\right)$ is a proper integral, since the function is bounded, continuous and integrated over a bounded domain.

Definition 11.34 (Improper integrals: unbounded intervals). Suppose that $f : [a, +\infty) \to \mathbb{R}$ is integrable in each bounded subinterval of $[a, +\infty)$ (ex. f continuous). Then

$$\int_a^{+\infty} f(x)\mathrm{d}x := \lim_{b\to+\infty} \int_a^b f(x)\mathrm{d}x.$$

Remark 11.35. Similarly, one computes

$$\int_{-\infty}^a f(x)\mathrm{d}x := \lim_{b\to-\infty} \int_b^a f(x)\mathrm{d}x$$

and

$$\int_{-\infty}^{+\infty} f(x)\mathrm{d}x := \int_{-\infty}^a f(x)\mathrm{d}x + \int_a^{+\infty} f(x)\mathrm{d}x.$$

Example 11.36. Suppose we want to integrate the function x over the whole real line. One might be tempted to argue that the integral is zero because the integrand is odd, writing

$$\lim_{t\to+\infty} \int_{-t}^t x\,dx.$$

However, if we shift the function to the right by one unit, obtaining $x-1$, the total area is only translated and thus unchanged, yet in this case the same reasoning would suggest that the integral equals $-\infty$, a clear paradox. The standard definition of improper integral resolves this issue by requiring a split into two separate limits:

$$\int_{-\infty}^{+\infty} x\,dx := \lim_{p\to-\infty} \int_p^0 x\,dx + \lim_{q\to+\infty} \int_0^q x\,dx = -\infty + \infty,$$

hence the integral simply does not converge.

Definition 11.37 (Improper integrals: unbounded functions). Suppose that $f : [a, b) \to \mathbb{R}$ is continuous. Then

$$\int_a^b f(x)\mathrm{d}x := \lim_{c\to b-} \int_a^c f(x)\mathrm{d}x.$$

We conclude this section by stating some properties which are dual to the corresponding results on the convergence of series. Similar results also hold for the case of unbounded functions defined on a bounded domain.

i) *Zero test:* If $\lim_{x\to+\infty} f(x)$ exists and $\int_a^{+\infty} f(x)\mathrm{d}x$ converges, then $\lim_{x\to+\infty} f(x) = 0$.

ii) *Comparison:* If $0 \le f \le g$ in $[a, +\infty)$ then $0 \le \int_a^{+\infty} f(x)\mathrm{d}x \le \int_a^{+\infty} g(x)\mathrm{d}x$.

iii) *Asymptotic Comparison:* if $f, g \ge 0$ and $f \sim g$ as $x \to +\infty$, then $\int_a^{+\infty} f(x)\mathrm{d}x$ converges if and only if $\int_a^{+\infty} g(x)\mathrm{d}x$ does.

iv) *Absolute Convergence:* If $\int_a^{+\infty} |f(x)|\mathrm{d}x$ converges, then $\int_a^{+\infty} f(x)\mathrm{d}x$ does.

Example 11.38 (Graphical Approaches). Sometimes, graphical arguments may help. Let us study the convergence of

$$\int_{\pi/2}^{+\infty} f(x)\mathrm{d}x = \int_{\pi/2}^{+\infty} \left|\frac{\cos x}{x}\right| \mathrm{d}x.$$

Note that $\left|\frac{\cos x}{x}\right| \le \frac{1}{x}$, and in particular the equality case is attained if $\cos x = \pm 1$, that is $x = n\pi$ and $f(n\pi) = \frac{1}{n\pi}$. If we try to graph this function, it will be made of many concave hills that touch the x axis at $n = \frac{\pi}{2} + n\pi$. We can lower bound each area in $[n\pi, (n+1)\pi]$ by a triangle with base π and height $\frac{1}{n\pi}$. Therefore, the area we are computing by the integral can be lower bounded by the sum of these triangles, which is $\sum_{n=1}^{\infty} \frac{\pi}{2} \cdot \frac{1}{n\pi} = \sum_{n=1}^{\infty} \frac{1}{2n} \to \infty$. Therefore, the integral diverges. **[Add picture]**

Theorem 11.39 (Series and improper integrals). *If $f \ge 0$ in $[1, +\infty)$ and f is **decreasing**, then*

$$\sum_{n=2}^{\infty} f(n) \le \int_1^{+\infty} f(x)\mathrm{d}x \le \sum_{n=1}^{+\infty} f(n).$$

Proof.
Observe that for every integer n, $\int_n^{n+1} f(x)\mathrm{d}x \leq \int_n^{n+1} f(n)\mathrm{d}x = f(n)$. Consequently,

$$\sum_{n=1}^{N} \int_n^{n+1} f(x)\mathrm{d}x \leq \sum_{n=1}^{N} f(n) \implies \int_1^{N+1} f(x)\mathrm{d}x \leq \sum_{n=1}^{N} f(n).$$

Taking the limit as $N \to +\infty$ yields the second inequality. Conversely, note that for every integer n, $f(n) = \int_{n-1}^{n} f(n)\mathrm{d}x \leq \int_{n-1}^{n} f(x)\mathrm{d}x$. Therefore,

$$\sum_{n=2}^{N} f(n) \leq \sum_{n=2}^{N} \int_{n-1}^{n} f(x)\mathrm{d}x = \int_1^N f(x)\mathrm{d}x$$

Passing the limit as $N \to +\infty$ results in the first inequality. $\square$

Remark 11.40 (Convergence Rates). The vast majority of improper integrals can be studied by asymptotic or pointwise comparison with integrals of the form $\int_0^1 x^\alpha \mathrm{d}x$ or $\int_1^\infty x^\alpha \mathrm{d}x$:

i) When integrating from 1 to ∞, the bigger the exponent, the easier the convergence:

$$\int_1^{+\infty} \frac{1}{x^\alpha}\mathrm{d}x \quad \text{converges if and only if } \alpha > 1$$

ii) When integrating from 0 to 1, the smaller the exponent, the easier the convergence:

$$\int_0^1 \frac{1}{x^\alpha}\mathrm{d}x \quad \text{converges if and only if } \alpha < 1.$$

As we can see, none of the two integrals converges if $\alpha = 1$.

Primitives

$x^\alpha, \alpha \neq -1$	$\frac{1}{\alpha+1}x^{\alpha+1}$	$\frac{1}{x}$	$\ln x$	$D(\alpha f + \beta g) = \alpha Df + \beta Dg$
e^x	e^x	$a^x \; 0 < a \neq 1$	$\frac{a^x}{\ln a}$	$D(fg) = Df \cdot g + f \cdot Dg$
$\cos x$	$\sin x$	$\sin x$	$-\cos x$	$D(f/g) = \frac{Df \cdot g - f \cdot Dg}{g^2}$
$\cosh x$	$\sinh x$	$\sinh x$	$\cosh x$	$D(1/g) = \frac{-Dg}{g^2}$
$\frac{1}{\cos^2 x} = 1 + \tan^2 x$	$\tan x$	$\frac{1}{\cosh^2 x}$	$\tanh x$	$D(g \circ f) = D(g(f)) \cdot Df$
$\frac{1}{1+x^2}$	$\arctan x$	$\frac{1}{1-x^2}$	$\tanh^{-1} x$	$D(f^{-1}(a)) = \frac{1}{Df(b)}$ with $f(b) = a$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$	$\frac{1}{\sqrt{1+x^2}}$	$\sinh^{-1} x$	
$\frac{1}{\sqrt{x^2-1}}$	$\cosh^{-1} x$			

11.8 Exercises

Example 1: improper integral

$$\int_0^{+\infty} \frac{1}{x^2+x+\sqrt{x}}\mathrm{d}x = \int_0^1 \frac{1}{x^2+x+\sqrt{x}}\mathrm{d}x + \int_1^{+\infty} \frac{1}{x^2+x+\sqrt{x}}\mathrm{d}x.$$

Notice that

$$\frac{1}{x^2+x+\sqrt{x}} \sim \frac{1}{\sqrt{x}} \text{ as } x \to 0 \qquad \text{and} \qquad \frac{1}{x^2+x+\sqrt{x}} \sim \frac{1}{x^2} \text{ as } x \to +\infty.$$

Therefore, by asymptotic comparison with the proper fraction both integrals converge, and so does the original one.
Example 2: integral function Find the natural domain of:

$$F(x) = \int_1^x \frac{e^t - 1}{\sqrt[3]{t^2(t-2)}}\mathrm{d}t$$

Solution: the integrand function is defined in $\mathbb{R} \setminus \{0, 2\}$. When $x \to 0$, we have

$$\lim_{x\to 0} \frac{e^x - 1}{\sqrt[3]{x^2(x-2)}} = \lim_{x\to 0} \frac{x}{-\sqrt[3]{2}x^{2/3}} = \lim_{x\to 0} -\sqrt[3]{\frac{x}{2}} = 0.$$

Therefore, $F(0)$ is well defined (it's not even an improper integral, since f is bounded). When $x \to 2$, we have

$$\frac{e^x - 1}{\sqrt[3]{x^2(x-2)}} \sim \frac{e^2 - 1}{4} \cdot \frac{1}{(x-2)^{1/3}}.$$

Even if f is not there defined, $F(2)$ exists as an improper integral, since $\int_1^2 \frac{1}{(x-2)^{1/3}}\mathrm{d}x = \frac{3}{2}(x-2)^{2/3}$ an, by asymptotic comparison, $\int_1^2 f(x)\mathrm{d}x$ also converges. Therefore, $dom(F) = \mathbb{R}$.
This example shows that the domain of the integral function is always an interval containing the center c. It's usually a subset of $dom(f)$, except for some accumulation points of $dom(f)$ for which f may be not defined, but its improper intergral may be.

Example 3: series and integral comparison Prove that

$$\sum_{n=0}^{\infty} \frac{1}{1+n^2} < 1 + \frac{\pi}{2}.$$

Solution: Let us define $f(x) := \frac{1}{1+x^2}$. Notice that the function is positive and strictly decreasing for $x \in [0, \infty)$. Therefore, we can apply the comparison between limits and integrals:

$$\sum_0^{\infty} \frac{1}{1+n^2} = 1 + \sum_{n=1}^{\infty} \frac{1}{1+n^2} < 1 + \int_0^{+\infty} \frac{1}{1+x^2}\mathrm{d}x = 1 + [\arctan x]_0^{+\infty} = 1 + \frac{\pi}{2}.$$

Example 4: Riemann Integral compute

$$\lim_{n\to+\infty} \frac{1}{n^2} \sum_{k=1}^{n} k(\log k - \log n).$$

Solution: Our claim is that the latter limit actually represents a Cauchy-Riemann integral and thus we need to get to that form. It seems reasonable to consider a uniform partition of step size $\Delta x_n = 1/n$ and $\xi_k = k/n$ so that

$$\begin{aligned}\lim_{n\to+\infty} \frac{1}{n^2} \sum_{k=1}^{n} k(\log k - \log n) &= \lim_{n\to+\infty} \frac{1}{n} \sum_{k=1}^{n} \frac{k}{n} \log\left(\frac{k}{n}\right) = \lim_{n\to+\infty} \sum_{k=1}^{n} \xi_k \log(\xi_k)\Delta_n = \int_0^1 x \log x \mathrm{d}x \\ &= \left[\frac{x^2}{2}\log x\right]_0^1 - \int_0^1 \frac{x^2}{2}\frac{1}{x}\mathrm{d}x = \left[\frac{x^2}{2}\log x - \frac{x^2}{4}\right]_0^1 = 0 - \frac{1}{4} - \lim_{x\to 0} \frac{x^2}{2}\log x = -\frac{1}{4}.\end{aligned}$$

12 Counterexamples

This section is specifically intended to help you in the multiple-choice questions, where seemingly reasonable deductions fail due to some intricated counterexamples.

i) A sequence for which $\lim_{n\to+\infty} a_{n+p} - a_n = 0$ for every positive integer p can diverge to $+\infty$. An example is the n-th partial sum of the harmonic series $S_n = S_{n-1} + \frac{1}{n}$ and $S_1 = 1$. Apparently, this violates the Cauchy criterion of convergence. Actually, in the Cauchy criterion we are able to find an N large enough for which $|a_{n+p} - a_n| < \varepsilon$ for every p, while here N depends on p.

ii) $a_n := \sin(\log(n))$ is an example of a sequence for which $\lim_{n\to+\infty} a_{n+1} - a_n = 0$ but a_n does not converge.

iii) The function

$$f(x) := \begin{cases} 1/n & \text{if } x \text{ is rational, } x = m/n \text{ in lowest terms, } n > 0; \\ 0 & \text{if } x \text{ is irrational.} \end{cases}$$

has many remarkable properties. It a function continuous at every irrational point and discontinuous at every rational point.[51] Such function is Riemann-integrable, and still has a dense set of points of discontinuity. Therefore, we can define $g(x) = \int_0^x f(t)\mathrm{d}t$ and $g'(x) \neq f(x)$ almost everywhere (it seems a violation of FTC 2, but actually f is not continuous). The same function, but with n instead of $1/n$, is a function everywhere finite but nowhere locally bounded.

iv) Let f be a continuous function such that $f(1) = 2$ and $(a_n)_{n\in\mathbb{N}}$ a sequence such that $f(a_n) \to 2$. Then a_n does not need to converge to 1. A counterexample is the function $f(x) = 2x^2$ and the sequence $a_n := (-1)^n$. What fails here is the injectivity: if we require f to be injective, then $f(a_n) \to 2$ implies $a_n \to 1$, due to the continuity of the inverse function.

v) The function

$$f(x) := \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is an example of a function which is continuous and differentiable at $x = 0$, but the limit of the derivative as $x \to 0$ does not exist. The same function, together with $g(x) = x$, can be employed to prove that L'Hôpital's rule is not invertible, since $\lim\limits_{x\to 0+} \frac{f(x)}{g(x)} = 0$ but $\lim\limits_{x\to 0+} \frac{f'(x)}{g'(x)}$ does not exist. What fails here is the continuity of the derivative, since the function is not C^1.

vi) $f(x) = 1/x$ is an example of a continuous function whose inverse is not monotone. What fails here it's the monotonicity of f: if f is monotone and injective (and therefore strictly monotone), its inverse is also strictly monotone and injective.

vii) $f : (-\infty, -1] \cup (1, +\infty) \to \mathbb{R}$ such that $f(x) = x^3 - x$ is an example of a continuous function with discontinuous inverse. What fails here is that the domain of f is not an interval.

viii) $f(x) = -1/x$ is an example of a function such that $f'(x) > 0$ for every $x \in dom(f)$ but it's not increasing. The theorem about the derivative of monotonic functions fails in this case because it applies to intervals.

ix) $f(x) = x^3$ is an example of a strictly increasing function on the open interval $(-1, 1)$ such that $f'(x) > 0$ is false.

x) $f(x) = x^4$ is an example of a strictly convex function on the open interval $(-1, 1)$ such that $f''(x) > 0$ is false.

xi) $f(x) = -x^4$ is an example of a function such that $f''(0) \geq 0$ but f is concave. Even better, $g(x) = x^3$ satisfies $g''(0) \geq 0$, but has an inflection point at $x = 0$.

xii) Let f be a function everywhere differentiable. Suppose $f'(0) > 0$. Then the function is not necessarily increasing in a neighbourhood of 0.

A counterexample is the function

$$f(x) := \begin{cases} x + 2x^2 \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases},$$

which is never increasing at 0. What fails here is the continuity of the derivative: for function with continuous derivative, $f'(0) > 0$ is sufficient to conclude monotonicity at 0. The same function has another counterintuitive property: $f(x) < 0$ if $x < 0$, $f(x) > 0$ if $f(x) > 0$, $f(0) = 0$, but the function is neither increasing nor decreasing in a neighbourhood of 0.

xiii) Let f be a function everywhere differentiable and suppose 0 is a global minima for f. Then the derivative of the function does not necessarily change sign in a neighbourhood of 0.

A counterexample is the function

$$f(x) := \begin{cases} 2x^4 + 2x^4 \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases},$$

[51] There do not exist a function continuous at every rational point and discontinuous at every irrational point.

for which 0 is a critical point (as required by Fermat's theorem), but there is no neighbourhood of 0 where the derivative is positive or negative: it continues to oscillate.

xiv) Let f be a function everywhere differentiable. Suppose $f''(0) > 0$. Then the function is not necessarily convex in a neighbourhood of 0.

A counterexample is the integral function of the previous function, $F(x) = \int_0^x f(x)\mathrm{d}x$. What fails here is the continuity of the second derivative: for functions of class C^2, $f''(0) > 0$ is sufficient to conclude strictly convexity.

xv) There are functions differentiable, monotonic, such that $f(x) \to 0$ as $x \to +\infty$ but $f'(x) \not\to 0$ as $x \to +\infty$. However, the construction is too complicated.

xvi) Let f be a function differentiable in $[0,1]$. Then its derivative is not necessarily bounded in $[0,1]$.

A counterexample is the function

$$f(x) := \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right) & \text{if } 0 < x \leq 1 \\ 0 & \text{if } x = 0 \end{cases},$$

for which $f'(0) = 0$ and $f'(x) = 2x\sin(1/x^2) - \frac{2}{x}\cos(1/x^2)$ if $x > 0$. What fails here is the continuity of the second derivative.

xvii) A continuous nonnegative function such that $\int_0^{+\infty}$ converges, does not need to converge to zero, i.e. the zero test fails without the condition that $\lim_{x\to+\infty} f(x)$ exists. To see why, imagine a function which forms isosceles triangles of height 2 and base $1/n^2$ centered in each natural number, and which is zero everywhere else. The integral of this function is the sum of the areas of the triangles, which is $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi}{6}$. You can even require f to be positive by just adding $1/x^2$ to the integrand.

A Appendix

A.1 Recap on real valued functions and topological notions

Real valued functions are functions of the form $f : A \subseteq \mathbb{R} \to B \subseteq \mathbb{R}$, assigning a unique real value to each element of its domain A. The image of f, sometimes called range of f, or image of A through f, is the set $f(A) := \{y \in B | \exists x \in A : f(x) = y\}$.

i) f is even if $f(-x) = f(x)$ for every $x \in A$ (and if $x \in A$, then $-x \in A$). f is odd if $f(-x) = -f(x)$ for every $x \in A$ (and if $x \in A$, then $-x \in A$).

ii) f is injective if, for every $a \neq b \in A$, we have $f(a) \neq f(b)$, i.e. $f(a) = f(b) \implies a = b$.

iii) f is surjective if, for every $y \in B$, there exists $x \in A$ such that $f(x) = y$, i.e. $f(A) = B$.

iv) f is bijective if it's injective and surjective.

v) f is monotone increasing (resp. decreasing) if $a < b \in A \implies f(a) \leq f(b)$ (resp. $f(a) \geq f(b)$). f is strictly increasing if $a < b \in A \implies f(a) < f(b)$.

Definition A.1 (Extended real line). The symbol ∞ is used to indicate both $\pm\infty$ together. They do not belong to $\mathbb{R}$, but the set $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\} = [-\infty, +\infty]$ is called *extended* real line and in such case $-\infty < x < +\infty$ for every $x \in \mathbb{R}$. We can extend the usual operations to $\overline{\mathbb{R}}$ almost always, for example $x + \infty = +\infty$, $(+\infty)\cdot(-\infty) = -\infty$, $-3\cdot(-\infty) = +\infty$, $\frac{3}{-\infty} = 0$. The only indeterminate forms are $+\infty - \infty$, $0 \cdot (\pm\infty)$, $\frac{\pm\infty}{\pm\infty}$, $\frac{0}{0}$. Moreover, we may add also 1^∞, 0^0 and ∞^0.

Definition A.2 (Open Set). Avoiding the formal topological definition (which will be presented in the Analysis 2 course), a set is closed if it contains the boundary points, and open if not. For example, $(a,b) = \{x \in \mathbb{R} : a < x < b\}$ is open, $[a,b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ is closed, $(a,b] = \{x \in \mathbb{R} : a < x \leq b\}$ is neither open nor closed, and the singleton $\{a\}$ is a closed (degenerate) interval. As a general remark, remember that the majority of theorems in analysis 1 behave better with open sets, since strange behaviours may appear at the boundaries. For this reason, in many theorems we will discard boundary points in two possible ways: either we will consider an open interval I, or we will consider only the points in the **interior** of I, being any interval.

Definition A.3 (Locality). Some properties are local. For example, if two functions coincide in a neighbourhood of x_0, their derivatives coincide as well. Similarly for the left and right derivatives. This is useful for functions defined piecewise. For example, $f(x) = x$ if $x < 0$ and $f(x) = \sin x$ if $x \geq 0$.

Definition A.4 (Omega function). In some proofs we will employ a function omega, which comes from the equivalence of these conditions:

i) $g : D \subseteq \mathbb{R} \to \mathbb{R}$ defined in a neighbourhood of 0 is little o of $h \in \mathbb{R}$, written $g(h) = o(h)$

ii) $g(0) = 0$ and $\lim_{h\to 0} \frac{g(h)}{h} = 0$.

iii) There exists a function $\omega : D \subseteq \mathbb{R} \to \mathbb{R}$ such that $\omega(0) = 0$, ω is continuous at 0 and such that $g(h) = h\omega(h)$ on D.

A.2 Types of discontinuities

i) **Removable** discontinuity: This occurs if $\lim_{x\to x_0+} f(x)$ and $\lim_{x\to x_0-} f(x)$ exist, are finite, and equal each other, but they are not equal to $f(x_0)$.[52] For example, $f(x) = \frac{\sin x}{x}$ for $x \neq 0$ and $f(0) = 10$.

ii) **Jump** discontinuity: This occurs if $\lim_{x\to x_0+} f(x)$ and $\lim_{x\to x_0-} f(x)$ exist and are finite, but do not equal each other. For example, $f(x) = \frac{x}{|x|}$ for $x \neq 0$ and $f(0) = 10$.

iii) **Essential** discontinuity: This occurs if at least one of $\lim_{x\to x_0+} f(x)$ or $\lim_{x\to x_0-} f(x)$ does not exist or is infinite. For example,

$$f(x) = \begin{cases} \frac{1}{x} & \text{if } x < 0, \\ 10 & \text{if } x = 0, \\ \sin\left(\frac{1}{x}\right) & \text{if } x > 0 \end{cases} \qquad \text{or} \qquad g(x) = \begin{cases} x & \text{if } x \leq 0, \\ \frac{1}{x} & \text{if } x > 0 \end{cases}.$$

A.3 Singularity points

There is no general consensus on what exactly constitutes a singularity. Some authors define a singular point as an accumulation point x_0 of $\text{dom}(f)$ which is not included in the domain. In this case, depending on the limit as $x \to x_0$, we may encounter removable singularities, jump singularities, or essential singularities. As a rule of thumb, if you remove the point $x_0 = 0$ from the domain of the functions used as examples of discontinuities, you obtain the corresponding singularities. For example, $f(x) = \frac{1}{x}$ has a singularity at $x = 0$, even though the function is continuous at $x = 0$. Note that discussing differentiability at singular points is meaningless. Statements like "the function $f(x) = \frac{x^2}{|x|}$ has a sharp corner at $x = 0$" are incorrect, because the behaviour of the limits of the derivative has nothing to do with the left and right derivatives, which require $f(x_0)$ to be defined.

A.4 Points of non-differentiability

i) **Sharp corner**: This occurs when $f'_-(x_0)$ and $f'_+(x_0)$ exist, are finite, but do not equal each other, or when one is finite and the other is not. For example,

$$f(x) = |x| \qquad \text{or} \qquad f(x) = \begin{cases} x & \text{if } x \leq 0, \\ \sqrt{x} & \text{if } x > 0 \end{cases}.$$

ii) **Cusp**: This occurs if $f'_-(x_0)$ and $f'_+(x_0)$ are both infinite and of opposite signs. For example, $f(x) = \sqrt[3]{|x|}$.

iii) **Inflection point with vertical tangent**: This occurs if $f'_-(x_0) = f'_+(x_0)$ are both infinite and of the same sign. For example, $f(x) = \sqrt[3]{x}$.

iv) The last case occurs when at least one of $f'_-(x_0)$ or $f'_+(x_0)$ does not exist. For example, $f(x) = x\sin(1/x)$ for $x \neq 0$ and $f(0) = 0$.

[52] Here, x_0 is considered an interior point of the domain of f. If x_0 lies on the boundary of $\text{dom}(f)$, the definitions may vary slightly, but the concept is essentially the same.