Minimal analysis I:
A concise introduction to real analysis of a single variable

Tonatiuh Sánchez-Vizuet
The University of Arizona

Last update: May 12, 2025

Contents

Chapter 1

Preliminaries

We will start by laying out the notation and basic concepts and objects that we will be working with, along with the notion of cardinality, countability and uncountability. Along with all these definitions, we will prove tow results that will be of tremendous importance and utility throughout the course: the triangle inequality and the notion that there is no *real number* that is larger than zero but smaller in magnitude to any other real number.

We start by introducing the notation for different sets of numbers.

- The ***natural numbers*** will be denoted as

  $$\mathbb{N} := \{1, 2, 3, 4, \ldots\}.$$

  Note that we will not include 0 in this set. The union of the set of natural numbers and zero is sometimes denoted by

  $$\mathbb{N}_0 : \{0, 1, 2, 3, 4, \ldots\}.$$

- The ***integers*** (positive and negative naturals including zero) will be denoted by

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

- The ***rational numbers*** are all possible quotients between integers and naturals, and will be denoted by

  $$\mathbb{Q} := \{p/q : p \in \mathbb{Z} \text{ and } q \in \mathbb{N}\}.$$

  That the fact that the denominator is a natural number—and that we excluded zero from the set of naturals—prevents division by zero. We recall that any rational number can be expressed as a finite or periodic decimal expansion—where periodic means that there is a finite string of decimals that gets repeated indefinitely many times.

- Providing a precise definition for the ***real numbers*** is remarkably non-trivial and we will not try to do so at this time. We will come back and give a more precise axiomatic definition of $\mathbb{R}$ in Chapter 2—once we have introduced the concept of the limit of a sequence, the concept of the least upper bound, and the axiom of completeness. For the time being we will informally state that a real number is a (potentially infinite) decimal expansion and will denote the set of all such numbers by $\mathbb{R}$.

Remark 1.1. *Given that the subject matter of these notes is real analysis it is quite odd that we shy away from providing a rigorous definition for the real numbers. The reason for this is that providing such a definition would require us taking a fairly lengthy detour into algebra and set theory, distracting us from our main goal: analysis. The reader interested in the details can check the exposition by Spivak [8] or Abbott [1]. The book by Landau [3] is a particularly detailed and instructive construction that does not require an advanced background and can be used as an introduction proofs (for the student willing to put on the time).*

Remark 1.2. *Due the particular arithmetic behavior stemming from its property of "being larger than any natural number"* ***infinity is not a real number****. In advanced analysis it is common to bundle up the real numbers together with ∞ and $-\infty$ into an algebraic structure known as the* **extended real numbers** *denoted by $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$. However, for the purpose of these notes, any real number will be strictly finite in magnitude.*

The following two equivalent statements are known as the ***Archimedean property*** of the natural numbers, which will be very useful in our upcoming endeavors.

- For any real number $x \in \mathbb{R}$ there exists a natural number $n \in \mathbb{N}$ such that $x < n$.
- For any $\epsilon > 0$, there exists a natural number $n \in \mathbb{N}$ such that $1/n < \epsilon$.

In analysis we often want to show that two seemingly different mathematical objects are indeed equal. The most direct way of showing this is to show that the difference between these quantities is equal to zero. However, we will not always be able to explicitly compute this difference and instead we will have to content ourselves with *estimating* this difference by showing that it can not exceed a certain threshold and then forcing the threshold to become arbitrarily small. This technique is grounded in the following theorem, which will be tremendously useful, that tells us that if the magnitude of a real number remains below any arbitrarily small threshold, then it must be equal to zero. Following [5], we will refer to this result as the ϵ-**principle**.

Theorem 1.1 (The *ϵ-principle*). *Consider a real number $0 \leq x$ such that $0 \leq x < \epsilon$ for all $\epsilon > 0$. Then $x = 0$.*

Proof. We will proceed by contradiction and assume that there exists a real number $x \neq 0$ and such that $0 \leq x < \epsilon$ for all $\epsilon > 0$. This would imply that $1 < \epsilon/x$ for all $\epsilon > 0$. Therefore, letting $\epsilon = x$ it must hold that $1 < \epsilon/x = x/x = 1$, which is a contradiction. Therefore $x = 0$. □

The ***absolute value*** of a real number x is defined as

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

The absolute value can also be expressed in the following equivalent way.

Lemma 1.1. $|x| = \max\{x, -x\}$

Proof. We consider first $x \geq 0$. In this case we have that

$$|x| = x \geq 0 \geq -x$$

and so $|x| = \max\{x, -x\} = x$. Alternatively, if $x < 0$ it follows that

$$|x| = -x > 0 > x.$$

Hence $|x| = \max\{x, -x\} = -x$. $\square$

We will use the previous lemma to prove one of the most useful inequalities in mathematical analysis.

Theorem 1.2 (***Triangle inequality***)**.** *For any $x, y \in \mathbb{R}$ it follows*

$$|x + y| \leq |x| + |y|.$$

Proof. We analyze two cases. If $x + y \geq 0$, then

$$|x + y| = x + y \leq |x| + |y|,$$

where Lemma 1.1 ensures that $x \leq |x|$ and $y \leq |y|$. If $x + y < 0$, then

$$|x + y| = -x - y \leq |x| + |y|,$$

where we used Lemma 1.1 to guarantee that $-x \leq |x|$ and $-y \leq |y|$. $\square$

The following easy consequence of the triangle inequality is also of great importance. Its proof will be left as an exercise.

Corollary 1.3 (***Reverse triangle inequality***)**.** *For all $x, y \in \mathbb{R}$ it holds that*

$$|x| - |y| \leq |x - y|.$$

A ***function*** is an assignment rule that, to every element of a set D (called the ***domain***), assigns *at most* one element of another set C (called the ***codomain***). In mathematical analysis is common to introduce a function, its domain and codomain using the notation

$$f : A \longrightarrow B\,,$$

where the set to the left of the arrow is the domain and the one to the right is the codomain. For the most part of these notes, we will focus on cases where both the domain and codomain are subsets of the real numbers.

If $a \in A$ is an element of the domain, we say that a *is an argument of the function* f and about the corresponding element of the codomain $f(a) \in B$ we say that "$f(a)$ is the value of f at a". Note that f (without an argument) is a function, while $f(a)$ is an element of the codomain. We will typically denote sets with capital Latin letters and elements of the sets with lower case letters.

For a function $f : A \to B$ with domain A and codomain B, the subset of the codomain consisting of all the possible values of the function is known as the ***image* of A under f** or simply as the ***range* of f** and is defined as

$$f(A) := \{b \in B : f(a) = b \text{ for some } a \in A\} \subset B.$$

Conversely, a function $f : A \to B$, the set of all elements of the domain that get mapped to a certain subset S of the codomain is denoted as the ***inverse image* of S under f** and is defined as

$$f^{-1}(S) := \{a \in A : f(a) \in S\} \subset A.$$

Remark 1.3. *The, universally accepted, notation for the inverse image defined above is a little unfortunate, as it can be easily confused with the inverse function (that we will define below). The inverse image is a set and always exists (although it may be empty), while the inverse function is a function and is not always defined. It is usually easy to distinguish between the two of them by the context, but as a rule of thumb, if the argument of* f^{-1} *is capitalized, then the object in question is a set (and therefore is the inverse function) while if it is in lower case, then chances are that the inverse function is being discussed.*

The image—or range—of a function is not necessarily equal to the totality of the codomain. However, in the cases where equality happens, we say that the function is ***surjective*** or ***onto*** and note that $f : A \to B$ is surjective if and only if for every $b \in B$ these exists some $a \in A$ such that $f(a) = b$.

On the other hand, if a function maps distinct points in the domain into distinct points in the codomain, we say that the function is ***surjective*** or ***one to one*** and note that $f : A \to B$ is injective if and only if $f(a) = f(b)$ implies that $a = b$. Injectivity can sometimes also be expressed in terms of the contrapositive statement $f : A \to B$ is injective if and only if $a \neq b$ implies that $f(a) \neq f(b)$.

When a function $f : A \to B$ is both injective and surjective we say that it is ***bijective*** or ***invertible***. When this happens, there exists a function, known as the ***inverse function***, $f^{-1} : B \to A$ such that, if $f(a) = b$ it holds that

$$f^{-1}(f(a)) = a \qquad \text{and} \qquad f(f^{-1}(b)) = b.$$

Remark 1.4. *Note that, while the inverse function may not exists, the inverse image is always defined—although it may be empty. For instance, if we consider the function* $f(x) = \sqrt{x}$*, then*

$$f^{-1}(\{x < 0\}) = \varnothing.$$

The ***cardinality*** of a set A refers to the number of elements contained in the set. The cardinality of a set A is sometimes denoted symbolically by either

$$|A| = x \qquad \text{or} \qquad \#A = x$$

where x is the number of elements contained in A.

We say that a set is ***countable*** if it is either finite, or if there exists a one-to-one cvorrespondence between the natural numbers and the elements of the set. If a countable set is infinitye we say that its cardinality is ***aleph-zero*** and write

$$|A| = \aleph_0.$$

We point out that, despite the definition of a countable set applies to finite sets, in analysis it is very common to use the adjective *countable* to refer almost exlusively to infinite sets. We say that a set A is ***uncountable*** if it is not finite but there is no one-to-one correspondence between its elements and the natural numbers. In this case we write

$$|A| = 2^{\aleph_0}.$$

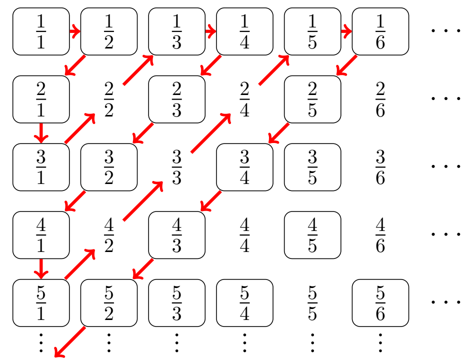

Figure 1.1: Cantor's ordering of the positive rationals. Starting at the upper left corner, the rationals are ordered following the direction of the arrows. To avoid repetitions, only the numbers in the rectangles (representing the first instance of any particular rational) are counted.

1.1 Exercises

1. Using the triangle inequality $|x + y| \le |x| + |y|$, prove that the following statements are true for all $x, y, z \in \mathbb{R}$.

   (a) $|x - y| \le |x - z| + |z - y|$ (In more general settings, this inequality is sometimes also called *the triangle inequality*).

   (b) $|x| - |y| \le |x - y| \le |x| + |y|$ (The inequality on the left is sometimes called the *reverse triangle inequality*).

2. Prove the following companion to the ϵ-principle: If $a, b \in \mathbb{R}$ are such that $a \le b + \epsilon$ for every $\epsilon > 0$, then $a \le b$. (Hint: argue by contradition)

3. Consider a function $f : A \to B$ and let E, F be subsets of A. Prove that:

   (a) $f(E \cup F) = f(E) \cup f(F)$.

   (b) $f(E \cap F) \subset f(E) \cap f(F)$.

4. Consider a function $f : A \to B$ and a (possibly uncountable) collection of sets $V_i \subset B$. Prove that:

   (a) $f^{-1}(V^c) = \left(f^{-1}\left(V_i\right)\right)^c$.

   (b) $\cup_i f^{-1}(V_i) = f^{-1}\left(\cup_i V_i\right)$.

   (c) $\cap_i f^{-1}(V_i) = f^{-1}\left(\cap_i V_i\right)$.

5. Let $f : A \to B$

   (a) Show that if f is injective and $E \subset A$ then $f^{-1}(f(E)) = E$. Provide an example that shows that the equality may not hold if f is not injective.

   (b) Show that if f is surjective and $F \subset A$ then $f(f^{-1}(F)) = F$. Provide an example that shows that the equality may not hold if f is not surjective.

6. Countable and uncountable sets

   (a) Prove that the intersection of two countable sets is countable. (Hint: There is a very short answer).

(b) Given two sets A and B, we define their Cartesian product as

$$A \times B := \{(a, b) : a \in A \text{ and } b \in B\}.$$

Prove that if both A and B are countable, then $A \times B$ is countable as well. [Hint: Attempt a similar strategy as the one used to prove the countability of the rationals. Figure 1.1 might prove useful].

Chapter 2

Completeness of the real line

Definition 2.1. Let $E \subset \mathbb{R}$ be a non empty set. A number $x \in \mathbb{R}$ is called:

- An ***upper bound*** of E if $x \geq e$ for all $e \in E$.
- A ***lower bound*** of E if $x \leq e$ for all $e \in E$.

A set is said to be ***bounded from above*** if it has at least one upper bound, ***bounded from below*** if it has at least one lower bound and simply ***bounded*** if it has both upper and lower bounds.

The axiom of the infimum/supremum. The following properties will be taken as an axiom of the real number system:

- If $E \subset \mathbb{R}$ is non empty and bounded from above, then there exists $S \in \mathbb{R}$ such that if U is an upper bound of E, then $S \leq U$. This number is called the ***supremum*** or the ***least upper bound*** of E, and is denoted as $\sup(E)$.
- If $E \subset \mathbb{R}$ is non empty and bounded from below, then there exists $I \in \mathbb{R}$ such that if L is a lower bound of E, then $I \geq L$. This number is called the ***infimum*** or the ***greatest lower bound*** of E, and is denoted as $\inf(E)$.

Theorem 2.1. *Let $E \subset \mathbb{R}$ be a non empty set bounded from above and S be an upper bound of E. Then $S = \sup(E)$ if and only if for all $\epsilon > 0$ there exists $e \in E$ such that $S - \epsilon < e$.*

Proof. $\boxed{\Rightarrow}$ Let $S = \sup(E)$ and assume for contradition that there exists $\widetilde{\epsilon} > 0$ such that for all $e \in E$ $S - \widetilde{\epsilon} \geq e$. Since $\widetilde{\epsilon} > 0$ it follows that

$$S > S - \widetilde{\epsilon} \geq e.$$

Therefore, $S - \widetilde{\epsilon}$ is an upper bound of E that is strictly smaller than S, which contradicts the assumption that $S = \sup(E)$.

$\boxed{\Leftarrow}$ Assume for contradiction that $S \neq \sup(E)$. Then $S - \sup(E) > 0$ and therefore, from the hypothesis, there exists $e \in E$ such that

$$e > S - (S - \sup(E)) = \sup(E).$$

However, this would imply that $\sup(E)$ *is not* an upper bound of E, which is a contradiction. □

The proof of the following result is analogous to the previous one.

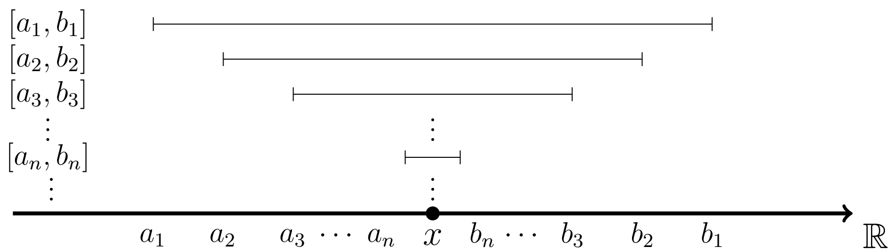

Figure 2.1: Nested interval theorem: The uncountable intesection of closed nested intervals whose length dereases to zero contains a single point.

Theorem 2.2. *Let $E \subset \mathbb{R}$ be a non empty set bounded from below and I be a lower bound of E. Then $I = \inf(E)$ if and only if for all $\epsilon > 0$ there exists $e \in E$ such that $I + \epsilon > e$.*

The supremum and infimum have several important properties associated with the sum and multiplication of sets. We wil prove the following one pertaining multiplications and the rest will be left as exercises.

Lemma 2.1. *If $A \subset \mathbb{R}$ is non-empty and bounded and $c \in \mathbb{R}$, then*

1. *If $c > 0$ then*
$$i)\ \inf(cA) = c\inf(A) \qquad \textit{and} \qquad ii)\ \sup(cA) = c\sup(A).$$

2. *If $c < 0$ then*
$$iii)\ \inf(cA) = c\sup(A) \qquad \textit{and} \qquad iv)\ \sup(cA) = c\inf(A).$$

Proof. We will prove the identities $i)$ and $iv)$, the remaining two can be proven in an analogous fashion. For $i)$ we have that for all $a \in A$ and $c \geq 0$ it follows that

$$\inf(cA) \leq ca \quad \Rightarrow \quad \frac{\inf(cA)}{c} \leq a \quad \Rightarrow \quad \frac{\inf(cA)}{c} \leq \inf(A) \quad \Rightarrow \quad \inf(cA) \leq c\inf(A).$$

Conversely, for all $a \in A$ and $c > 0$ we have

$$\inf(A) \leq a \quad \Rightarrow \quad c\inf(A) \leq ca \quad \Rightarrow \quad c\inf(A) \leq \inf(cA)\,,$$

which proves $i)$. For $iv)$, since $c < 0$, we have that for all $a \in A$

$$\sup(cA) \geq ca \quad \Rightarrow \quad \frac{\sup(cA)}{c} \leq a \quad \Rightarrow \quad \frac{\sup(cA)}{c} \leq \inf(A) \quad \Rightarrow \quad \sup(cA) \geq c\inf(A).$$

Conversely
$$\inf A \leq a \quad \Rightarrow \quad c\inf(A) \geq ca \quad \Rightarrow \quad c\inf(A) \geq \sup(cA).$$

Which proves $iv)$. $\square$

The following theorem is a very important tool in analysis whose proof relies heavily on the properties of the infimum and the supremum.

Theorem 2.3 (***Nested interval theorem***)**.** *Let $I_n := [a_n, b_n]$ be a family of closed and nested intervals*

$$I_1 \supset I_2 \supset \ldots \supset I_n \supset I_{n+1} \supset \ldots$$

such that $|I_n| := |b_n - a_n| \to 0$ as $n \to \infty$. Then, the intersection $\cap_{n=1}^{\infty} I_n$ contains exactly one point.

Proof. The proof has two steps: First we will show that the intersection of all the nested intervals is a closed interval, then we will show that the endpoints of this interval are equal.

Step 1. The fact that the intervals are closed and nested implies the following three inequalities

$$i)\ a_n \le a_{n+1}\ \forall\, n \in \mathbb{N} \qquad ii)\ b_{n+1} \le b_n\ \forall\, n \in \mathbb{N} \qquad iii)\ a_n \le b_m\ \forall\, n, m \in \mathbb{N}.$$

From $iii)$ we can conclude that $a_1 \le b_m\ \forall\, m \in \mathbb{N}$ and that $a_n \le b_1\ \forall\, n \in \mathbb{N}$. Therefore, the sequence $\{a_n\}$ is bounded from above and the sequence $\{b_m\}$ is bounded from below. Let's define then

$$a := \sup(\{a_n\}) \qquad \text{and} \qquad b := \inf(\{b_m\})\,.$$

Once again, from $iii)$ we observe that *every* b_m is an upper bound for the sequence $\{a_n\}$ and therefore

$$a_n \le a \le b_m \ \ \forall\, n, m \in \mathbb{N},$$

where the first inequality follows from the fact that a is an upper bound for $\{a_n\}$, and the second inequality follows from observing that every b_m is an upper bound of $\{a_n\}$ and a is the least upper bound. The sequence of inequalities above then shows that a is a lower bound for the sequence $\{b_m\}$ and therefore, since b is the greatest lower bound of $\{b_m\}$, we must have

$$a_n \le a \le b \le b_m \ \ \forall\, n, m \in \mathbb{N}.$$

If we now let $m = n$ in the expression above, we see that the interval $[a, b] \subset [a_n, b_n]$ for every n. Therefore $[a, b] \subset \cap_{n=1}^{\infty} I_n$. We will now show that the reverse inclusion also holds. Take $x \in \cap_{n=1}^{\infty} I_n$, then $a_n \le x \le b_n$ for all n. This implies that: a) x is an upper bound for $\{a_n\}$ and therefore $a \le x$, and b) x is a lower bound for $\{b_n\}$ and thus $x \le b$. These two inequalities together imply that $x \in [a, b]$ and therefore $\cap_{n=1}^{\infty} I_n \subset [a, b]$. We have then proven that $\cap_{n=1}^{\infty} I_n = [a, b]$, which completes the first step.

Step 2. We will now show that $a = b$. Take an arbitrary $\epsilon > 0$ and observe that:

1. Since we have assumed that $|b_n - a_n| \to 0$, there exists $N_1 \in \mathbb{N}$ such that, for all $n \ge N_1$, $|b_n - a_n| < \epsilon/3$.
2. Since $a = \sup(\{a_n\})$, the Theorem 2.1 and the inequality $i)$ above imply that there exists N_2 such that for all $n \ge N_2$, $0 \le a - a_n < \epsilon/3$.
3. Since $b = \inf(\{b_n\})$, the Theorem 2.2 and the inequality $ii)$ above imply that there exists N_3 such that for all $n \ge N_3$, $0 \le b_n - b < \epsilon/3$.

Therefore, taking $N = \max\{N_1, N_2, N_3\}$ it follows that for all $n \ge N$

$$|b - a| \le |b - b_n| + |b_n - a_n| + |a_n - a| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.$$

Since the left hand side of the inequality above is a fixed number and ϵ is arbitrary, it follows that $|b - a| < \epsilon$ for any positive ϵ and therefore $a = b$. Hence, the intersection of all the closed and nested intervals contains a single point. □

Remark: If we had taken *open* intervals of the form $I_n := (a_n, b_n)$ instead of closed, then step one of the proof would have led us to conclude that $\cap_{n=1}^{\infty} I_n = (a, b)$, and step two would have led to $a = b$. Since the open interval (a, a) is empty, it would follow that the intersection of all the nested intervals is empty.

The following result is a big theorem in analysis. Our proof will make use of the previous result and showcases a technique known as ***bisection***.

Theorem 2.4 (***Bolzano-Weierstrass theorem***)**.** *Let $\{x_n\} \subset \mathbb{R}$ be a bounded sequence. Then there exists a subsequence $\{x_{n_k}\} \subset \{x_n\}$ that converges to a point $x \in \mathbb{R}$.*

Proof. We will build a convergent subsequence by picking terms in $\{x_n\}$ contained in a family of closed and nested intervals of shrinking size. Using the nested interval theorem we will argue that there exists a point contained in the intersection of all the intervals and we will the prove that this point is in fact the limit of the subsequence.

Since the sequence $\{x_n\}$ is bounded, there exists M such that $-M \leq x_n \leq M$ for all n. Define $I_1 := [-M, M]$, pick any term in the sequence that is contained in I_1, and label it as $x_{n_1} = x$. Then divide the interval I_1 into two closed subintervals of length M. Since the sequence is infinite, at least one of these two subintervals must contain infinitely many terms of the sequence $\{x_n\}$, for if both halves contain only finitely many terms, then the sequence would have to be finite. Let I_2 be one of the halves that contains infinitely many terms; pick any of such terms and call it x_{n_2}. Since the sequence has infinitely many terms, we can continue the process in the same fashion: dividing the previous interval in two closed subintervals, picking one element of the sequence contained on the subdivision, choosing one of the two halves that contains infinitely many terms and repeating.

After k such steps we will have constructed:

1. A sequence of closed and nested intervals I_k with length $|I_k| = M/2^{k-2}$, and
2. A sequence of k points $x_{n_k} \in \{x_n\}$ each of them contained inside the closed interval I_k.

Since the intervals are closed, nested and their length converges to zero, the nested interval theorem guarantees that the intersection $\cap_{k=1}^{\infty} I_k$ contains exactly one point, call it x.

We now note that both x and x_{n_k} are contained in I_k, and therefore $|x - x_{n_k}| \leq M/2^{n-2}$. Hence, for any given $\epsilon > 0$ it is enough to take $N > \log_2(M/\epsilon) + 2$ to guarantee that for all $n_k \geq N$ we will have $|x_{n_k} - x| < \epsilon$. This proves that $x_{n_k} \to x$. We have thus extracted from $\{x_n\}$ a convergent subsequence $\{x_{n_k}\}$. □

2.1 Exercises

1. Let $A \subset B \subset \mathbb{R}$ be non-empty and bounded. Prove that
$$\inf(B) \leq \inf(A) \leq \sup(A) \leq \sup(B).$$
2. Prove that if sequence of real numbers is both increasing and bounded from above then it is convergent. (The same result holds for decreasing sequences that are bounded from below).
3. Prove that a set $S \in \mathbb{R}$ is closed and bounded if and only if every sequence of points in S has a subsequence that converges to a point in S.
4. Prove that every Cauchy sequence of real numbers converges.
5. Prove that if an non empty closed set $S \subset \mathbb{R}$ is bounded above, then it has a largest element.

Chapter 3

Differentiation

Definition 3.1 (***Differentiability***)**.** We say that the function $f : \mathbb{R} \to \mathbb{R}$ is ***differentiable*** at x and define

$$f'(x) := \lim_{t \to x} \frac{f(t) - f(x)}{t - x}, \tag{3.1}$$

whenever the limit on the right hand side exists. We say that f is differentiable on $A \subset \mathbb{R}$ when f is differentiable for every $x \in A$.

The quotient appearing in the definition of the derivative, is often known as the ***Newton quotient*** and, whenever it is defined, gives the slope of the straight line passing through the points $(x, f(x))$ and $(t, f(t))$. Note that by defining $h := t - x$ the Newton quotient can be expressed in the equivalent form

$$\frac{f(t) - f(x)}{t - x} = \frac{f(x + h) - f(x)}{h}.$$

The following well known properties of the derivative follow directly (perhaps after some algebraic manipulation) from the definition of the derivative. Properties 2) and 3) together are referred to as the ***linearity of the derivative***.

Theorem 3.1. *Let $f, g : \mathbb{R} \to \mathbb{R}$ be differentiable at a point x and $c \in \mathbb{R}$ be a constant. Then*

1. *If f is constant, then $f'(x) = 0$.*
2. $(f + g)'(x) = f'(x) + g'(x)$.
3. $(cf)'(x) = cf'(x)$.
4. $(f \cdot g)'(x) = f'(x) \cdot g(x) + g'(x) \cdot f(x)$.
5. *If $g(x) \neq 0$ then* $(f/g)'(x) = \dfrac{f'(x) \cdot g(x) - g'(x) \cdot f(x)}{g^2(x)}$.

The property of differentiability is much stronger than that of continuity. Not every function that is continuous at a point is differentiable at that same point. However, as we shall now prove, if a function is differentiable at a point, it must necessarily be continuous.

Proposition 3.1. *If f is differentiable at x, then f is continuous at x.*

Proof. Given that f is differentiable at x, the limit (3.1) exists, and therefore, for every $\epsilon > 0$, there exists $\delta > 0$ such that

$$|t - x| < \delta \quad \Rightarrow \quad \left| \frac{f(t) - f(x)}{t - x} - f'(x) \right| < \epsilon.$$

From this, it follows that
$$|f(t) - f(x) - (t-x)f'(x)| < |t-x|\epsilon.$$
Applying the reverse triangle inequality to the left hand side of the expression above we obtain
$$|f(t) - f(x)| < |t-x|(\epsilon + |f'(x)|).$$
Hence, letting $r := \min\{\delta, \epsilon/(\epsilon + |f'(x)|)\}$, we see that all $t \in B_r(x)$ verify
$$|f(t) - f(x)| < \epsilon,$$
and therefore f is continuous at x. □

The following well-known fact is simply the contrapositive of the previous proposition.

Corollary 3.2. *If f is discontinuous at x, then f is not differentiable at x.*

Definition 3.2 (*Local extrema*). A function f is said to have a ***local maximum*** (resp. ***local minimum***) at a point x_0, if there exists $r > 0$ such that
$$f(x_0) \geq f(x) \;\; \forall\, x \in B_r(x_0) \qquad (\text{resp. } f(x_0) \leq f(x) \;\; \forall\, x \in B_r(x_0))\,.$$
The point x_0 at which the function attains a local maximum (resp. minimum) is called a ***maximizer*** (resp. ***minimizer***).

One of the most common—and useful—applications of differentiation is as a tool for finding local extrema. The usual procedure of locating some of the candidate points by finding the zeros of the derivative is justified by the following result.

Proposition 3.2. *If the function f has a local maximum or minimum at a point x_0 and is differentiable at x_0, then*
$$f'(x_0) = 0.$$

Proof. We will assume that f has a local maximum at x_0; the proof for local minima is completely analogous. Since x_0 is a local maximizer, then there exists $r > 0$ such that for all t such that $|t - x_0| < r$ it follows that $f(t) \leq f(x_0)$. For any such t it then follows that $f(t) - f(x_0) \leq 0$ therefore
$$\text{if } t < x \;\;\Rightarrow\;\; 0 \leq \frac{f(t) - f(x_0)}{t - x} \qquad \text{while if } \;\; t > x \;\;\Rightarrow\;\; \frac{f(t) - f(x_0)}{t - x} \leq 0.$$
Since f is differentiable at x, the limit of the Newton quotient as $t \to x$ exists and must be equal to the derivative and to both of the one sided limits. Therefore
$$0 \leq \lim_{t \to x^-} \frac{f(t) - f(x_0)}{t - x} = f'(x) = \lim_{t \to x^+} \frac{f(t) - f(x_0)}{t - x} \leq 0,$$
and the result follows. □

Theorem 3.3 (*Rolle's theorem*). *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous over the compact interval $[a, b]$ and differentiable over (a, b). If $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. By hypothesis f is continuous over the compact interval $[a, b]$ therefore, from the extreme value theorem, there exist points $x_{\text{min}}, x_{\text{max}} \in [a, b]$ such that

$$f(x_{\text{min}}) \leq f(x) \qquad \text{and} \qquad f(x_{\text{max}}) \geq f(x) \quad \text{for all } x \in [a, b].$$

There are two possible cases:

1. If both x_{min} and x_{max} are located at the endpoints of the interval then, from the assumption that $f(a) = f(b)$, it would follow that $f(x_{\text{min}}) = f(x_{\text{max}})$ and thus the function would be constant, yielding $f'(x) = 0$ for all $x \in (a, b)$.
2. At least one of $x_{\text{min}}, x_{\text{max}}$ belongs to the interval (a, b). Call this point x^* Then, from the Proposition 3.2 it follows that $f'(x^*) = 0$.

□

Theorem 3.4 (***Mean value theorem***). *Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous over $[a, b]$ and differentiable over (a, b). Then, there exists some $c \in (a, b)$ such that*

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

Proof. We wish to use Rolle's theorem to aid us in our proof. However, in order to apply it we must first build an auxiliary function—involving f—that satisfies the hypotheses of Rolle's theorem. With that goal in mind, we pick a arbitrary parameter $r \in \mathbb{R}$, and define and auxiliary function of the form $g(x) := f(x) - rx$.

Given that f is continuous over $[a, b]$ and differentiable over (a, b), g will also satisfy both of these properties. We must then find the appropriate value for r that would yield $g(a) = g(b)$. In other words, we search for an r such that

$$f(a) - ra = f(b) - rb,$$

which clearly will be satisfied if and only if $r = \dfrac{f(b) - f(a)}{b - a}$.

We can now invoke Rolle's theorem for the function

$$g(x) := f(x) - \left(\frac{f(b) - f(a)}{b - a}\right) x$$

to ensure the existence of a point $c \in (a, b)$ such that

$$g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a} = 0,$$

which proves the result. □

Corollary 3.5. *If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant on (a, b).*

Proof. Let $x, y \in (a, b)$ and notice that f is continuous over $[x, y]$ and differentiable over (x, y). Then, by the mean value theorem, there exists $c \in (x, y)$ such that

$$f'(c) = \frac{f(x) - f(y)}{x - y},$$

but $f'(c) = 0$ for any $c \in (a, b)$ therefore, using this fact in the expression above, it follows that for any $x, y \in (a, b)$

$$f(x) = f(y).$$

□

We now prove a closely related result, sometimes called the ***generalized mean value theorem*** and some other times the ***Cauchy's mean value theorem***.

Theorem 3.6 (Generalized mean value theorem). *If the functions $f, g : \mathbb{R} \to \mathbb{R}$ are both continuous over the interval $[a, b]$ and differentiable over (a, b), then there exists a point $c \in (a, b)$ such that*

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Proof. Since both f and g are continuous over the interval $[a, b]$ and differentiable over (a, b), the function $h : \mathbb{R} \to \mathbb{R}$ defined as

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

is also continuous over $[a, b]$ and differentiable over (a, b). Moreover, simple algebraic manimpulations show that

$$\begin{aligned}
h(a) = \ & (g(b) - g(a))f(a) - (f(b) - f(a))g(a) \\
&= g(b)f(a) - g(a)f(b) \\
= \ & g(b)f(a) - g(b)f(b) + f(b)g(b) - g(a)f(b) \\
&= (g(b) - g(a))f(b) - (f(b) - f(a))g(b) \\
&= h(b).
\end{aligned}$$

We can therefore apply Rolle's theorem to obtain a point $c \in (a, b)$ such that

$$0 = h'(c) = (g(b) - g(a))f'(c) - (f(b) - f(a))g'(c),$$

from which the theorem follows. $\square$

The following result highlights what is perhaps the most important property of the derivative: it determines the best linear approximation to a (differentiable) function at one point. This property is often overlooked in one dimension, but is key to generalizing the concept of derivative to multiple dimensions and beyond.

Lemma 3.1 (*Fundamental lemma of differentiation*). *Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable at a point x_0. Then, there exists a real-valued function η, defined in an interval around zero such that η is continuous at zero and*

$$\eta(0) = 0, \tag{3.2a}$$
$$f(x) = f(x_0) + (x - x_0)(f'(x_0) + \eta(x - x_0)). \tag{3.2b}$$

Proof. We will simply define a function that satifies these two properties by construction, and will then show that it is indeed continuous at zero. Define

$$\eta(h) := \begin{cases} \frac{f(x_0+h)-f(x_0)}{h} - f'(x_0) & \text{if } h \neq 0 \\ 0 & \text{if } h = 0. \end{cases} \tag{3.3}$$

Letting $h = x - x_0$ in the definition above, it is easy to see that properties (3.2a) and (3.2b) are satisfied by construction, while from the definition of η and the fact that f is differentiable at x_0 we have

$$\lim_{h\to 0} \eta(h) = \lim_{h\to 0} \left(\frac{f(x_0 + h) - f(x_0)}{h} - f'(x_0) \right) = f'(x_0) - f'(x_0) = 0\,,$$

and thus η is continuous at 0 as desired. $\square$

Remark 3.1. *The value of the function η can be interpreted as a measure of how far from the derivative the Newton quotient is as h approaches zero. Clearly, if the function is differentiable, the "distance" must decrease smoothly as $h \to 0$ and should vanish on the limit. This is the intuitive interpretation of (3.2a) and the continuity of η at zero. Moreover, (3.2b) states that, if a function is differentiable at a point x_0, one can approximate it in the vicinity of x_0 by the straight line going through the point $(x_0, f(x_0))$ and with slope given by $f'(x_0)$ as*

$$f(x) \approx f(x_0) + (x - x_0) f'(x_0)$$

incurring an error that vanishes faster than the distance between the evaluation point x and the approximation point x_0, as measured by the magnitude of the term $(x - x_0)\eta(x - x_0)$. Due to this geometric interpretation, the lemma is also known as the **linear approximation lemma**.

We will now take advantage of the function provided by the previous lemma to prove the well-known chain rule.

Theorem 3.7 (*Chain rule*). *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous over the compact interval $[a, b]$ and that it is differentiable at some point $x_0 \in [a, b]$. Let $g : f([a, b]) \to \mathbb{R}$ and suppose that g is differentiable at the point $f(x_0)$. If we define*

$$\phi : [a, b] \to \mathbb{R} \qquad \text{as} \qquad \phi(x) := g(f(x)),$$

then ϕ is differentiable at x_0 and

$$\phi'(x_0) = g'(f(x_0)) \cdot f'(x_0). \tag{3.4}$$

Proof. We want to show that the limit as $h \to 0$ of the Newton quotient associated to the composition $\phi(x)$ exists and is indeed equal to the right hand side of (3.4). This will follow from a careful application of the linear approximation lemma. Since f is differentiable at x_0 and g is differentiable at $f(x_0)$, Lemma 3.1 guarantees the existence of functions η_f and η_g continuous at 0 and satisfying

$$\eta_f(0) = \eta_g(0) = 0, \tag{3.5a}$$

$$f(x_0 + h_f) - f(x_0) = h_f(f'(x_0) + \eta_f(h_f)), \tag{3.5b}$$

$$g(f(x_0) + h_g) - g(f(x_0)) = h_g\left(g'(f(x_0)) + \eta_g(h_g)\right), \tag{3.5c}$$

where we have defined $h_f := x - x_0$ and $h_g := f(x) - f(x_0)$. Let us then analyze the Newton quotient

$$\begin{aligned}
\frac{\phi(x_0 + h_f) - \phi(x_0)}{h_f} &= \frac{g(f(x_0 + h_f)) - g(f(x_0))}{h_f} \\
&= \frac{h_g(g'(f(x_0)) + \eta_g(h_g))}{h_f} && \text{(By (3.5c))} \\
&= \frac{(f(x) - f(x_0))(g'(f(x_0)) + \eta_g(h_g))}{h_f} && \text{(Since } h_g := f(x) - f(x_0)\text{)} \\
&= \frac{h_f(f'(x_0) + \eta_f(h_f))(g'(f(x_0)) + \eta_g(h_g))}{h_f} && \text{(From (3.5b))},
\end{aligned}$$

and therefore, expanding the product above we obtain

$$\frac{\phi(x_0 + h_f) - \phi(x_0)}{h_f} = g'(f(x_0)) \cdot f'(x_0) + \eta_g(h_g) \cdot f'(x_0) + \eta_f(h_f) \cdot g'(f(x_0)) + \eta_f(h_f) \cdot \eta_g(h_g).$$

Now, since f is continuous at x_0, we observe that

$$\lim_{h_f \to 0} h_g = \lim_{h_f \to 0} (f(x) - f(x_0)) = \lim_{h_f \to 0} (f(x_0 + h_f) - f(x_0)) = 0,$$

and thus, in view of the continuity of both η_g and η_f at zero, combined with (3.5a), it follows that

$$\lim_{h_f \to 0} \eta_f(h_f) = \lim_{h_f \to 0} \eta_g(h_g) = 0. \tag{3.6}$$

Therefore, letting $h_f \to 0$ in the Newton quotient we obtain

$$\begin{aligned}\lim_{h_f \to 0} \frac{\phi(x_0 + h_f) - \phi(x_0)}{h_f} &= \lim_{h_f \to 0} \left(g'\left(f(x_0)\right) \cdot f'(x_0) \right) \\ &+ \underbrace{\lim_{h_f \to 0} \left(\eta_g\left(h_g\right) \cdot f'(x_0) + \eta_f(h_f) \cdot g'\left(f(x_0)\right) + \eta_f(h_f) \cdot \eta_g\left(h_g\right) \right)}_{=0 \ \text{From (3.6)}},\end{aligned}$$

from which we obtain (3.4) as desired. ☐

We will now make immediate use of the chain rule to prove another useful result:

Theorem 3.8 (***Inverse function theorem***)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous and invertible on $[a, b]$, and differentiable at $x \in [a, b]$ with $f'(x_0) \neq 0$. Then the inverse function f^{-1} is differentiable at the point $f(x)$, and*

$$\left(f^{-1}\right)' \left(f(x)\right) = \frac{1}{f'(x)}. \tag{3.7}$$

Proof. The proof is a straightforward application of the chain rule. Since f is invertible at x we have that

$$f^{-1}\left(f(x)\right) = x\,.$$

Differentiating both sides of the identity above and applying the chain rule on the left hand side yields

$$\left(f^{-1}\right)' \left(f(x)\right) \cdot f'(x) = 1.$$

Since we assumed that $f'(x) \neq 0$, the result follows immediately from the equality above. ☐

3.1 Exercises

1. Let $f, g : \mathbb{R} \to \mathbb{R}$ be differentiable at a point $x \in \mathbb{R}$. Using only the definition of the derivative as a limit prove that:

   (a) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$.

   (b) If $g(x) \neq 0$, then $(f/g)'(x) = \dfrac{g(x)f'(x) - f(x)g'(x)}{g^2(x)}$.

2. (a) Let $f(x) = x^2$ if $x \in \mathbb{Q}$, and $f(x) = 0$ if $x \in \mathbb{Q}^c$. Prove that f is differentiable at 0.

   (b) Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that $|f(x)| \leq x^2$ for all x. Prove that f is differentiable at 0.

   (c) We want to generalize the previous result by replacing x^2 with a more general function $g(x)$. Determine what property $g(x)$ must have, and use it to prove the result: "If g satisfies [property that you determined] and $|f(x)| \leq g(x)$ for all x, then $f(x)$ is differentiable at zero.

3. *This problem is a companion to problem 4 in Homework 4.*

   Suppose that the function $f : [0, 1] \to [0, 1]$ is differentiable on $[0, 1]$, and that $f'(x) \neq 1$ for all $x \in [0, 1]$. Prove that there is **exactly one** $x^* \in [0, 1]$ such that $f(x^*) = x^*$.

4. Prove that if $f'(x)$ is increasing, then every tangent line to the graph of f intersects it only once.

Chapter 4

Integration

In elementary calculus texts, the concept of integration is usually introduced through the discussion of the "intuitive" geometric idea of computing the area of a region enclosed by the horizontal axis, the graph of a function, and vertical line segments starting at the integration limits. For continuous and non negative functions, this geometric interpretation of the integral seems quite natural. The interpretation gets a little more complicated when dealing with continuous functions that may change sign. This new conceptual difficulty is quickly settled by *declaring* that the "area" of the regions where the graph lies below the horizontal axis will be assigned a negative value—although the notion of a "negative" surface should strike an inquisitive mind as not that natural anymore. Things get even murkier when the function in question has discontinuities... especially when the number of discontinuities becomes too large. Under these pathological circumstances, the supposedly "intuitive" concept of area below a curve becomes much less so, and the need for a riguourous definition of area becomes apparent. Such a definition is by no means trivial and we shall not attempt to define it here; we will leave that for a more advanced course on *measure theory.* Instead, we will now carefully construct *one* notion of integral that will match the intuition of area under a curve for functions that are non negative and continuous. The particular construction that we will study is due to Jean Gaston Darboux (1842–1917) and Bernhard Riemann (1826–1866).

Remark 4.1. *In this section we will consider* ***bounded*** *functions of the form $f : [a, b] \to \mathbb{R}$, where the interval $[a, b]$ is compact. We will* ***not*** *require the function to be continuous.*

Definition 4.1 (*Partition*). A ***partition*** P of an interval $[a, b]$ is a **finite** collection of points $x_0, \ldots, x_n$ satisfying the condition

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b.$$

Definition 4.2 (*Upper and lower sums*). For a bounded function defined over a compact interval $[a, b]$ and a corresponding partition P, we can define

$$m_i := \inf \{f(x) : x_{i-1} \leq x \leq x_n\} \qquad \text{and} \qquad M_i := \sup \{f(x) : x_{i-1} \leq x \leq x_n\}. \tag{4.1}$$

Note that m_i and M_i must be defined in terms of infima and suprema due to the fact that we are not requiring that f be continuous. We then define the ***lower sum*** of f with respect to the partition P, and the ***upper sum*** of f with respect to the partition P respectively as:

$$L(f, P) := \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \quad \text{(lower sum)} \quad \text{and} \quad U(f, P) := \sum_{i=1}^{n} M_i(x_i - x_{i-1}) \quad \text{(upper sum)}.$$

For non-negative functions, these quantities can be interpreted respectively as the sums of the areas of rectangles with bases of length $x_i - x_{i-1}$ and heights lying below and above the graph of f. One would hope that, as the number of rectangles grows and the width of their bases decreases, these numbers would approximate the intuitive value of the area under the curve.

Given that, for a fixed partition P, the inequality $m_i \leq M_i$ is trivially satisfied, it follows that $L(f,P) \leq U(f,P)$. However, if P and Q are *different* partitions of the same interval, it is not entirely clear that the analogous inequality $L(f,P) \leq U(f,Q)$ is also true. Spoiler alert: it is. However, since the points included in each partition may be different, it turns out that in order to establish such a relation we must first find some common ground between the two partitions in question. We will now set out to achieve that goal.

Definition 4.3 (*Refinement*). If P and Q are partitions of an interval $[a,b]$, we say that P is a ***refinement*** of Q if $Q \subset P$, with the inclusion being strict. In other words, P is a refinement of Q if P includes all the points of Q, and at least one more.

The proof of the following lemma is simple and is left as an excercise

Lemma 4.1. *Let A and B be bounded and non empty sets of real numbers such that $A \subset B$. Then*

$$\inf(B) \leq \inf(A) \leq \sup(A) \leq \sup(B).$$

We will make use of the result above to prove the following

Lemma 4.2. *Let P be a partition of the interval $[a,b]$, and Q be a refinement of P. Then*

$$L(f,P) \leq L(f,Q) \qquad \text{and} \qquad U(f,Q) \leq U(f,P). \tag{4.2}$$

Proof. The proof of the result uses an inductive argument. Let us start by considering that the refinement Q includes *exactly one* more point than P, and let us label that one point as x^*. Therefore, there exists one point $x_j \in P$ such that $x_{j-1} < x^* < x_j$ and the partitions then take the form

$$P := \{a = x_0, \dots x_{j-1}, x_j, \dots, x_n = b\} \qquad \text{and} \qquad Q := \{a = x_0, \dots x_{j-1}, x^*, x_j, \dots, x_n = b\}.$$

We then define

$$m^*_{j-1} := \inf\{f(x) : x_{j-1} \leq x \leq x^*\}\,, \qquad M^*_{j-1} := \sup\{f(x) : x_{j-1} \leq x \leq x^*\}\,,$$
$$m^*_{j} := \inf\{f(x) : x^* \leq x \leq x_j\}\,, \qquad M^*_{j} := \sup\{f(x) : x^* \leq x \leq x_j\}\,,$$

and note that—since the intervals $[x_{j-1}, x^*]$ and $[x^*, x_j]$ are both subsets of $[x_{j-1}, x_j]$—Lemma 4.1 yields

$$m_j \leq m^*_{j-1}\,, \qquad m_j \leq m^*_j\,, \tag{4.3a}$$
$$M_j \geq M^*_{j-1}\,, \qquad M_j \geq M^*_j\,. \tag{4.3b}$$

We now turn to the lower sum of f with respect to P. If we isolate the term corresponding to the interval containing x^* and split it into the contributions of the subintervals $[x_{j-1}, x^*]$ and $[x^*, x_j]$, it then follows that

$$\begin{aligned}
L(f,P) &= \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \\
&= \left(\sum_{i=1}^{j-1} m_i(x_i - x_{i-1})\right) + m_j(x^* - x_{j-1}) + m_j(x_j - x^*) + \left(\sum_{i=j+1}^{n} m_i(x_i - x_{i-1})\right) \\
&\leq \left(\sum_{i=1}^{j-1} m_i(x_i - x_{i-1})\right) + m^*_{j-1}(x^* - x_{j-1}) + m^*_j(x_j - x^*) + \left(\sum_{i=j+1}^{n} m_i(x_i - x_{i-1})\right) \quad \text{(From (4.3a))} \\
&= L(f,Q).
\end{aligned}$$

The last equality follows from the fact that $m_1, \dots, m^*_{j-1}, m^*_j, \dots, m_n$ are the infima of the function values at every subinterval of the partition Q, while the differences $(x_1 - x_0), \dots, (x^* - x_{j-1}), (x_j - x^*), \dots, (x_n - x_{n-1})$ are the lengths of all the subintervals in Q. Similarly, for the upper sum we obtain

$$\begin{aligned}
U(f,P) &= \sum_{i=1}^{n} M_i(x_i - x_{i-1}) \\
&= \left(\sum_{i=1}^{j-1} M_i(x_i - x_{i-1})\right) + M_j(x^* - x_{j-1}) + M_j(x_j - x^*) + \left(\sum_{i=j+1}^{n} M_i(x_i - x_{i-1})\right) \\
&\geq \left(\sum_{i=1}^{j-1} M_i(x_i - x_{i-1})\right) + M^*_{j-1}(x^* - x_{j-1}) + M^*_j(x_j - x^*) + \left(\sum_{i=j+1}^{n} M_i(x_i - x_{i-1})\right) \quad \text{(From (4.3b))} \\
&= U(f,Q).
\end{aligned}$$

Therefore, (4.2) holds if Q contains only one more point than P.

We then assume as induction hypothesis that (4.2) holds as well if Q contains k more points than P and study the case when Q contains $k+1$ more points than P. Let $x^*_1 < \ldots < x^*_{k+1}$ be the additional points in Q and $x_j \in P$ be the point such that $x_{j-1} < x^*_{k+1} < x_j$. Moreover, note that $P = P_* \cup P_{**}$, where

$$P_* := \{x_0, \dots, x_{j-1}\} \qquad \text{and} \qquad P_{**} := \{x_{j-1}, \dots, x_n\}$$

are respectively partitions of the intervals $[a, x_{j-1}]$ and $[x_{j-1}, b]$. Similarly, $Q = Q_* \cup Q_{**}$ where

$$Q_* := \{x_0, \dots, x^*_1, \dots, x^*_k, \dots, x_{j-1}\} \qquad \text{and} \qquad Q_{**} := \left\{x_{j-1}, x^*_{k+1} \dots, x_n\right\}$$

are respectively refinements of P_* (with k additional points) and P_{**} (with one additional point). Therefore, the previous analysis together with the induction hypothesis yield:

$$L(f,P_*) \leq L(f,Q_*)\,, \qquad L(f,P_{**}) \leq L(f,Q_{**})\,, \tag{4.4a}$$
$$U(f,P_*) \geq U(f,Q_*)\,, \qquad U(f,P_{**}) \geq U(f,Q_{**})\,. \tag{4.4b}$$

It then follows from the first two inequalities in (4.4a) that

$$L(f,P) = L(f,P_*) + L(f,P_{**}) \leq L(f,Q_*) + L(f,Q_{**}) = L(f,Q),$$

while from the last two in (4.4b) we obtain

$$U(f,P) = U(f,P_*) + U(f,P_{**}) \geq U(f,Q_*) + U(f,Q_{**}) = U(f,Q),$$

which concludes the proof. □

Recalling that, intuitively, lower sums underestimate the area under the graph of f and upper sums over estimate it, the previous result can be interpreted geometrically as follows. Whenever a partition is refined by the inclusion of new points, the estimation of the area below the graph of f that we obtain by using the new partition improves in the senste that both the under and over estimation become smaller.

We can now use the previous lemma to establish a ”common ground” to compare the upper and lower sums from any two different partitions. We will achieve this by considering a refinement of both partitions that includes all the points in both of them.

Theorem 4.1. *Let P and Q be different partitions of the interval $[a, b]$. Then*

$$L(f, P) \leq U(f, Q). \tag{4.5}$$

Proof. We define the new partition $R := P \cup Q$ of the interval $[a, b]$ and note that it is a refinement of both, since $P \neq Q$. Therefore, the previous lemma ensures that

$$L(f, P) \leq L(f, R) \leq U(f, R) \leq U(f, Q).$$

□

A very important consequence of the previous result is that the set of lower sums of a bounded function over a compact interval is bounded above, while the set of its upper sums is bounded below. This follows from the pprevious theorem by picking the trivial partition T consisting only of the endpoints of the interval so that, for any other partition P we have

$$(b-a) \inf_{a \leq x \leq b} (f(x)) = L(f, T) \leq U(f, P) \quad \text{and} \quad L(f, P) \leq U(f, T) = (b-a) \sup_{a \leq x \leq b} (f(x)) .$$

Knowing that the set of lower sums is non-empty and bounded from above, and that the set of upper sums is non empty and bounded from below we can now make use of their respective supremum and infimum to define

Definition 4.4 (***Upper and lower integrals***)**.** Let $[a, b] \subset \mathbb{R}$ be compact and $f : [a, b] \to \mathbb{R}$ be bounded. We define the ***upper integral*** and ***lower integral*** of f over $[a, b]$ respectively by

$$\overline{\int_a^b} f := \inf\{U(f, P) : P \text{ is a partition of } [a, b]\} \qquad \text{(Upper integral)}\,,$$

$$\underline{\int_a^b} f := \sup\{L(f, P) : P \text{ is a partition of } [a, b]\} \qquad \text{(Lower integral)}\,.$$

Definition 4.5 (***Integrability***)**.** Let $[a, b] \subset \mathbb{R}$ be compact and $f : [a, b] \to \mathbb{R}$ be bounded. We say that f is ***Riemann integrable*** or simply ***integrable*** if

$$\overline{\int_a^b} f = \underline{\int_a^b} f\,.$$

When this occurs we then drop the distinction between the upper and lower integrals and refer to either simply as the ***integral*** of f denoting them by the familiar notation

$$\int_a^a f.$$

The quantity above is also known as the ***definite integral***.

The definition above matches the intuitive idea that, as long as the function is not pathologically strange, the smallest possible over estimate and the largest possible under estimate of the area under its curve should coincide and be equal to the area.

Since partitions with more points result in more rectangles with narrower bases, one can also interpret the definition of integrability as the fact that when we approximate the area using finer and fine partitions, the decreasing upper and the increasing lower estimates should squeeze the true value of the area and eventually meet at that common value. As the following result shows, these two notions are in fact equivalent. Its proof is left as an excercise.

Theorem 4.2. *Let $[a, b] \subset \mathbb{R}$ be compact and $f : [a, b] \to \mathbb{R}$ be bounded. Then, f is* **integrable** *if and only if for every $\epsilon > 0$ there exists a partition P of $[a, b]$ such that*

$$U(f, P) - L(f, P) < \epsilon\,. \tag{4.6}$$

Unfortunately, not every function lends itself to integration under this definition, as the following example will demonstrate.

Example 4.1. Consider the characteristic function of the rational numbers, also known as the ***Dirichlet function***,

$$\chi_{\mathbb{Q}}(x) := \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \in \mathbb{Q}^c \end{cases}.$$

Due to the density of the rationals in $\mathbb{R}$, any interval of positive length includes infintely many rationals and irrationals. Therefore, for any subinterval I of any partition P, of any interval $[a, b]$ we have that

$$\inf\{\chi_{\mathbb{Q}}(x) : x \in I\} = 0 \qquad \text{and} \qquad \sup\{\chi_{\mathbb{Q}}(x) : x \in I\} = 1.$$

Therefore for any partition P we have that $U(f, P) - L(f, P) = 1$, and $\chi_{\mathbb{Q}}$ is therefore a ***non integrable function***. This highligts, even at this early stage, the need for a stronger definition of integral. Nevertheless, the Riemann-Darboux integral is already flexible and powerful enough to justify its use and study.

Having constructed our integral with the geometric notion of the area under the curve of a continuous function (and having shown that there indeed exist functions that are not integrable) the following reasuring result guarantees that our construction will indeed be effective for continuous functions.

Theorem 4.3. *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then f is integrable.*

Before proving the statement, we discuss the reasoning behind the proof. We will use the notation in (4.1) for the infimum and supremum of a function over a subinterval. Given $\epsilon > 0$, we would like to show that there exists a partition P such that (4.6) is satisfied. We start by analyzing the difference between the upper and lower sums of f for some partition $P := \{a = x_0, \ldots, x_n = b\}$. If somehow we were able to uniformly bound the difference between the maximum and minimum values of f over every subinterval such that $M_i - m_i \leq C$ for some $C > 0$ it would then follow that

$$U(f, P) - L(f, P) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) \leq C \sum_{i=1}^{n} (x_i - x_{i-1}) = C(b - a).$$

If, additionally, the constant C were "small enough" we could then make the difference between the upper and lower sum arbitrarily small. Therefore, we must be able to choose a partition such that, over every subinterval, the difference between the maximum and minimum value of the function is "uniformly small". The fact that the function is continuous over a compact interval will allow us to do that. As we shall see now.

Proof. Since f is continuous over the compact interval $[a, b]$, it is uniformly continuous. Therfore, given $\epsilon > 0$ there exists $\delta > 0$ such that for all $x, y \in [a, b]$ if holds that $|x - y| < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \epsilon/(b - a)$. We then pick a partition P such that for all i we have $x_i - x_{i-1} < \delta$ and therefore $M_i - m_i < \epsilon/(b - a)$, this yields

$$U(f, P) - L(f, P) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) \leq \frac{\epsilon}{b - a} \sum_{i=1}^{n} (x_i - x_{i-1}) = \frac{\epsilon}{b - a}(b - a) = \epsilon,$$

and the function is integrable. $\square$

Continuous functions are not the only functions that are integrable. The class of, not necessarily continuous, monotonic functions is as well.

Theorem 4.4. *Let $f : [a, b] \to \mathbb{R}$ be monotonic. Then f is integrable.*

The idea of the proof is similar as that of the previous theorem. Lets consider the case of a monotonically increasing function since the argument for a decreasing function is completely analogous. If we analyze the difference between the lower and an upper sum of f with respect to a partition P, and use the fact , we obtain

$$U(f, P) - L(f, P) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) = \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))(x_i - x_{i-1}).$$

Note that, since the function is increasing

$$\sum_{i=1}^{n} (M_i - m_i) \leq \lim_{x\to b} f(x) - \lim_{x\to a} f(x).$$

Thus, if we could bound the length of each subinterval by some constant $C > 0$ we would obtain

$$U(f, P) - L(f, P) \leq C \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \leq C(\lim_{x\to b} f(x) - \lim_{x\to a} f(x)).$$

Therefore, if C is small enough, the difference between upper and lower sums can be narrowed. This observation leads to the proof.

Proof. Let $\epsilon > 0$ and $D := \lim_{x\to b} f(x) - \lim_{x\to a} f(x)$. We then choose a partition P such that, for all $x_i \in P$ we have $x_i - x_{i-1} < \epsilon/D$. It follows that, for this partition, we have

$$U(f, P) - L(f, P) \leq \frac{\epsilon}{D} \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) < \frac{\epsilon}{D} D = \epsilon$$

and the function is integrable. □

Proving the ***linearity of the integral*** is surprisingly non-trivial. While it does not need any advanced result, it does require a lot of careful work with upper and lowers sums over partitons of the interval. We will start by recalling that the upper and lower sums of a function with respect to a partition are defined in terms of infima and suprema of function values over each subinterval in the partition. In view of this, it is easy to see how Lemma 2.1 implies the following:

Proposition 4.1. *Let $[a, b]$ be bounded, $P := \{a = x_0, \ldots, x_n = b\}$ be a partition of $[a, b]$, and $f : [a, b] \to \mathbb{R}$ be bounded. Then*

- *If $c > 0$, then $L(cf, P) = cL(f, P)$ and $U(cf, P) = cU(f, P)$.*
- *If $c < 0$, then $L(cf, P) = cU(f, P)$ and $U(cf, P) = cL(f, P)$.*

Proof. If $c > 0$ we have that

$$L(cf, P) = \sum_{i=1}^{n} \inf\{cf(x) : x_{i-1} \leq x \leq x_i\}(x_i - x_{i-1}) = c\sum_{i=1}^{n} \inf\{f(x) : x_{i-1} \leq x \leq x_i\}(x_i - x_{i-1}) = cL(f, P),$$

while for the upper sum it follows that

$$U(cf,P)=\sum_{i=1}^{n}\sup\{cf(x):x_{i-1}\le x\le x_i\}(x_i-x_{i-1})=c\sum_{i=1}^{n}\sup\{f(x):x_{i-1}\le x\le x_i\}(x_i-x_{i-1})=cU(f,P).$$

Similarly, if $c<0$ we observe that

$$L(cf,P)=\sum_{i=1}^{n}\inf\{cf(x):x_{i-1}\le x\le x_i\}(x_i-x_{i-1})=c\sum_{i=1}^{n}\sup\{f(x):x_{i-1}\le x\le x_i\}(x_i-x_{i-1})=cU(f,P),$$

while for the upper sum it follows that

$$U(cf,P)=\sum_{i=1}^{n}\sup\{cf(x):x_{i-1}\le x\le x_i\}(x_i-x_{i-1})=c\sum_{i=1}^{n}\inf\{f(x):x_{i-1}\le x\le x_i\}(x_i-x_{i-1})=cL(f,P).$$

□

We are now in the position to use the previous result to prove the firs step towards linearity of the integral

Theorem 4.5. *If $c\in\mathbb{R}$ and $f:[a,b]\to\mathbb{R}$ is integrable, then $cf:[a,b]\to\mathbb{R}$ is integrable and*

$$\int_a^b cf=c\int_a^b f.$$

Proof. First let us note that if $c=0$ the result is trivial, so we first assume that $c>0$. Then

$$\begin{aligned}
\underline{\int_a^b} cf &= \sup\{L(cf,P):P\text{ is a partition of }[a,b]\}\\
&= c\sup\{L(f,P):P\text{ is a partition of }[a,b]\} && \text{(from Proposition 4.1)}\\
&= c\int_a^b f && \text{(since } f \text{ is integrable)}\\
&= c\inf\{U(f,P):P\text{ is a partition of }[a,b]\} && \text{(since } f \text{ is integrable)}\\
&= \inf\{U(cf,P):P\text{ is a partition of }[a,b]\} && \text{(from Proposition 4.1)}\\
&= \overline{\int_a^b} cf\,.
\end{aligned}$$

On the other hand, if $c<0$ we have

$$\begin{aligned}
\underline{\int_a^b} cf &= \sup\{L(cf,P):P\text{ is a partition of }[a,b]\}\\
&= c\inf\{U(f,P):P\text{ is a partition of }[a,b]\} && \text{(from Proposition 4.1)}\\
&= c\int_a^b f && \text{(since } f \text{ is integrable)}\\
&= c\sup\{L(f,P):P\text{ is a partition of }[a,b]\} && \text{(since } f \text{ is integrable)}\\
&= \inf\{U(cf,P):P\text{ is a partition of }[a,b]\} && \text{(from Proposition 4.1)}\\
&= \overline{\int_a^b} cf\,.
\end{aligned}$$

□

Before proving the second part of the linearity of the integral, we need to establish one further auxiliary result

Lemma 4.3. *Let $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ be bounded. Then*

$$\inf\{f(x) : x \in [a,b]\} + \inf\{g(x) : x \in [a,b]\} \leq \inf\{f(x) + g(x) : x \in [a,b]\} \tag{4.7a}$$

$$\sup\{f(x) + g(x) : x \in [a,b]\} \leq \sup\{f(x) : x \in [a,b]\} + \sup\{g(x) : x \in [a,b]\} \tag{4.7b}$$

Proof. Since

$$\inf\{f(x) : x \in [a,b]\} \leq f(x)\,\forall\, x \in [a,b] \qquad \text{and} \qquad \inf\{g(x) : x \in [a,b]\} \leq g(x)\,\forall\, x \in [a,b]$$

we have that

$$\inf\{f(x) : x \in [a,b]\} + \inf\{g(x) : x \in [a,b]\} \leq \{f(x) + g(x) : x \in [a,b]\}.$$

so that the sum of the infima is a lower bound for the set of sum values. Since the infimum is the largest lower bound of a set, (4.7a) follows from the ineqiality above. Analoglously, since

$$\sup\{f(x) : x \in [a,b]\} \geq f(x)\,\forall\, x \in [a,b] \qquad \text{and} \qquad \sup\{g(x) : x \in [a,b]\} \geq g(x)\,\forall\, x \in [a,b]\,,$$

it follows that

$$\sup\{f(x) : x \in [a,b]\} + \sup\{g(x) : x \in [a,b]\} \geq \{f(x) + g(x) : x \in [a,b]\}.$$

So that the sum of the suprema is an upper bound for the set of sum values. The inequality 4.7b follows from this and from the observation that the supremum is the smallest possible upper bound of a set. ☐

Remark 4.2. *In the result above, the values of the set $\{f(x) + g(x) : x \in [a,b]\}$ are restricted to share the same common value of the argument x so that the supremum is then taken over values of the sum with the same argument. On the other hand, the supremum of the sets $\{f(x) : x \in [a,b]\}$ and $\{g(x) : x \in [a,b]\}$ may happen for different values of the argument x. This explains the use of inequalities (rather than equalities) in the result above. For instance, consider the function $f(x) = x$ and $g(x) = 1 - x$ over the interval $[0,1]$. The maximum value of f is $f(1) = 1$ and the maximum value of g is $g(0) = 1$. Therefore*

$$\sup\{f(x) : x \in [0,1]\} + \sup\{g(x) : x \in [0,1]\} = 2.$$

On the other hand, $f(x) + g(x) = x + (1 - x) = 1$, and therefore

$$\sup\{f(x) + g(x) : x \in [0,1]\} = 1.$$

We now move on to the second part of the linearity. We will make use of Lemmas 4.2 and 4.3 and the characterization of integrability given by Theorem 4.2.

Theorem 4.6. *If f and g are integrable over $[a,b]$, then $f + g : [a,b] \to \mathbb{R}$ is integrable and*

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g\,.$$

Proof. We start by recalling that if P is a partition of $[a,b]$ and Q is a refinement of P, Lemma 4.2 ensures that $U(f,P) \geq U(f,Q)$ and $L(f,P) \leq L(f,P)$. These two inequalities together imply

$$U(f,P) - L(f,P) \geq U(f,Q) - L(f,Q). \tag{4.8}$$

We then let $\epsilon > 0$ be arbitrary. Since f and g are both integrable over $[a,b]$, Theorem 4.2 guarantees the existence of partions P_f and P_g of $[a,b]$ such that

$$\epsilon/2 > U(f,P_f) - L(f,P_f) \qquad \text{and} \qquad \epsilon/2 > U(g,P_g) - L(g,P_g).$$

We take now the refinement of both P_f and P_g given by $Q := P_f \cup P_g$ and use (4.8) and the inequalities above to arrive at

$$\epsilon/2 > U(f,P_f) - L(f,P_f) \geq U(f,Q) - L(f,Q) \qquad \text{and} \qquad \epsilon/2 > U(g,P_g) - L(g,P_g) \geq U(g,Q) - L(g,Q).$$

Adding these two inequalities we see that

$$\begin{aligned} \epsilon &> (U(f,P_f) + U(g,P_g)) - (L(f,P_f) + L(g,P_g)) \\ &\geq (U(f,Q) + U(g,Q)) - (L(f,Q) + L(g,Q)) \\ &\geq U(f+g,Q) - L(f+g,Q) && \text{(From Lemma 4.3).} \end{aligned}$$

Therefore $f+g$ is integrable. It remains to show that the integral of the sum is equal to the sum of the integrals. With that goal in mind, we note that for any partition P the following two inequalities must hold

$$\begin{aligned} L(f,P) + L(g,P) &\leq \int_a^b f + \int_a^b g \leq U(f,P) + U(g,P) \\ L(f,P) + L(g,P) \underbrace{\leq}_{\text{By } 4.7a} L(f+g,P) &\leq \int_a^b (f+g) \leq U(f+g,P) \underbrace{\leq}_{\text{By } 4.7b} U(f,P) + U(g,P) \end{aligned}$$

Subtracting the second row from the first one yields

$$L(f,P) + L(g,P) - (U(f,P) + U(g,P)) \leq \int_a^b f + \int_a^b g - \int_a^b (f+g) \leq U(f,P) + U(g,P) - (L(f,P) + L(g,P)) \tag{4.9}$$

If we let $P := P_f + P_g$ be the refinement used in the previous part then we have that

$$0 \leq U(f,P) + U(g,P) - (L(f,P) + L(g,P)) < \epsilon$$

and therefore (4.9) implies that

$$-\epsilon \leq \int_a^b f + \int_a^b g - \int_a^b (f+g) \leq \epsilon.$$

Since for any ϵ we can pick a partition for which the inequality above is satisfied, we conclude that

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g$$

as desired. □

We have now built a relatively large family of integrable functiond. We have shown that continuous functions over a compact interval are integrable, that monotonic functions are as well, as long as they are bounded, that multiplying an integrable function by a constant number or adding it to another integrable function results in a new integrable function, the fact that the product of integrable functions is also integrable will be dealt with in Problem 5.

Consider a finite set of real numbers $A := \{a_1, \dots, a_n\}$. When computing the mean value of the set A, we add all the numbers in the set, and then divide the result by the "size" of the set—namely the cardinality of A. This leads to the formula for the average value $\langle A \rangle$ of the set A

$$\langle A \rangle = \frac{1}{n} \sum_{i=1}^{n} a_i.$$

Above, the notation $\langle A \rangle$ is one of many ways to denote the average value of a set. By analogy, if f is an integrable function over $[a, b]$ and we consider the set $\{f(x) : x \in [a, b]\}$ of function values over the interval, we can define the mean value by integrating the function over the inteval $[a, b]$ and dividing the result by the lenght of the interval, leading to the definition of the ***mean value*** of the function f over $[a, b]$ as:

$$\langle f \rangle := \frac{1}{b-a} \int_a^b f.$$

This value is well defined for all integrable functions regardless of wether they are continuous or not. However, as we shall now see, if in addition to being integrable, the function is continuous over $[a, b]$ there is in fact one point in the interval $[a, b]$ where the function value is equal to its mean value.

Theorem 4.7 (***Mean value theorem for integrals***)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous. Then, there exists one point $x \in [a, b]$ such that*

$$f(x) = \frac{1}{b-a} \int_a^b f.$$

Proof. We first note that, since f is continuous over $[a, b]$, Theorem 4.3 guarantees that it will be integrable over the interval as well, so that $\int_a^b f$ exists. Moreover, since f is continuous over the compact interval $[a, b]$, the extreme value theorem gives the existence of points $x_m, x_M \in [a, b]$ where the function attains its minimum and maximum values. Therefore, for every $x \in [a, b]$ we have that

$$m := f(x_m) \leq f(x) \qquad \text{and} \qquad f(x) \leq f(x_M) =: M. \tag{4.10}$$

Moreover, recall that for an integrable function over an interval $[a, b]$ and any partition P of such interval it holds that

$$L(f, P) \leq \int_a^b f \leq U(f, P).$$

Therefore, by taking P to be the trivial partition that includes only the endpoints a and b the inequality above implies that

$$f(x_n)(b-a) = m(b-a) \leq \int_a^b f \leq M(b-a) \leq f(x_M)(b-a),$$

where we used (4.10) to express the minimum and the maximum of f. The last inequality can be written in the form

$$f(x_m) \leq \frac{1}{b-a} \int_a^b f \leq f(x_M),$$

which, recalling that f was assumed to be continuous over $[a, b]$, can be combined with the intermediate value theorem to then guarantee that there exists $x \in (x_m, x_M) \subset [a, b]$ such that

$$f(x) = \frac{1}{b-a} \int_a^b f,$$

as we were trying to prove. $\square$

4.1 Exercises

Problems 1, 2, 3, 4, and 5 all showcase important properties of the integral. Therefore, solving them and becoming confortable with their conclusions should be consider essential.

1. The following five statements are closely related to each other. The proof of one should be a relatively easy consequence of the previous one. Thus, if you make sure that your proof for the first tatement is correct, the rest of the dominoes should fall.

   (a) Let $a < b < c$ and f be integrable on $[a, c]$. Prove that f is integrable on $[a, b]$ and on $[b, c]$ and that
   $$\int_a^c f = \int_a^b f + \int_b^c f. \tag{4.11}$$

   (b) Let $a < b < c < d$ and f be integrable on $[a, d]$. Prove that f is integrable on $[b, c]$ and that
   $$\int_a^d f = \int_a^b f + \int_b^c f + \int_c^d f.$$
   [Hint: This should follow easily from the previous part.]

   (c) Let f be integrable on $[a, b]$ and $0 \leq f(x)$ for all $x \in [a, b]$. Prove that $0 \leq \int_a^b f(x)\, dx$.

   (d) Prove that if f and g are integrable on $[a, b]$ and $f(x) \leq g(x)$ for all $x \in [a, b]$, then
   $$\int_a^b f(x)\, dx \leq \int_a^b g(x)\, dx.$$
   This result is referred to as the ***monotonicity of the integral***.

   (e) Prove that if f is integrable on $[a, b]$, then
   $$\left| \int_a^b f(x)\, dx \right| \leq \int_a^b |f(x)|\, dx.$$
   This result is sometimes called the ***triangle inequality for integrals***.

2. Let $[a, b] \subset \mathbb{R}$ be compact and $f : [a, b] \to \mathbb{R}$ be bounded. Prove that f is integrable if and only if for every $\epsilon > 0$ there exists a partition P of $[a, b]$ such that
   $$U(f, P) - L(f, P) < \epsilon\, .$$

3. Let $f : [a, b] \to \mathbb{R}$ be integrable. Prove that
   $$\int_a^b f = -\int_b^a f.$$

4. (a) Let $f : [a, b] \to \mathbb{R}$ be bounded and equal to zero exept at a finite number of points in $[a, b]$. Prove that f is integrable on $[a, b]$ and that
$$\int_a^b f = 0.$$

   (b) Let $f : [a, b] \to \mathbb{R}$ be integrable and $g : [a, b] \to \mathbb{R}$ be a function such that $g(x) = f(x)$ except for a finite number of points in $[a, b]$. Prove that g is integrable and that
$$\int_a^b g = \int_a^b f.$$
   [Hint: This should follow easily from part a)].

5. (a) Let $f : [a, b] \to \mathbb{R}$ be bounded, such that there exists $B > 0$ with $|f(x)| \leq B$ for all $x \in [a, b]$. Show that for any partition P of $[a, b]$
$$U(f^2, P) - L(f^2, P) \leq 2B\left(U(f, P) - L(f, P)\right).$$
   [Hint: $f^2(x) - f^2(y) = (f(x) + f(y))\,(f(x) - f(y))$].

   (b) Show that if f is integrable on $[a, b]$, so is f^2.

   (c) Show that if f and g are integrable on $[a, b]$, so is fg.

6. In what follows, let $f[a, b] \to \mathbb{R}$ be integrable and non-negative.

   (a) Prove that if $[c, d] \subset [a, b]$, then $\int_c^d f \leq \int_a^b f$.

   (b) Prove that if f is continuous on $[a, b]$ and $\displaystyle\int_a^b f = 0$ then $f = 0$.

   (c) If, in addition to being continuous over $[a, b]$, f is such that $\displaystyle\int_a^b fg = 0$ for all continuous functions $g : [a, b] \to \mathbb{R}$. Prove that $f = 0$. [Hint in view of part b), there is an obvious choice for g]. This result is known as the ***fundamental lemma of the calculus of variations***.

7. Suppose f and g are continuous on $[a, b]$ such that $\int_a^b f = \int_a^b g$. Prove there exists $x \in (a, b)$ such that $f(x) = g(x)$.

8. Let f and g be continuous functions on $[a, b]$.

   (a) Prove that if $0 \leq g(x)$ for all $x \in [a, b]$, then there exists $x \in (a, b)$ such that
$$\int_a^b g(t) f(t)\, dt = f(x) \int_a^b g(t)\, dt$$

   (b) Use the previous part to derive a proof of the mean value theorem for integrals.

   (c) Show by exhibiting an example that the conclusion of the first part may not hold if g changes sign over $[a, b]$.

Chapter 5

The fundamental theorems of calculus

In this chapter we will focus on the interaction between differentiation and integration. We start by noting that the process of integration can be used to define new real-valued functions as follows. If $f : [a, b] \to \mathbb{R}$ is integrable we define

$$F(x) := \int_a^x f \qquad \text{for } \; x \in [a, b]. \tag{5.1}$$

We will devote the reminder of this chapter to study the properties of this function. In particular its continuity and differentiability.

Theorem 5.1. *If $f : [a, b] \to \mathbb{R}$ is integrable, then the function $F(x) : [a, b] \to \mathbb{R}$ given by (5.1) is continuous.*

Proof. We start by considering a point $c \in (a, b)$ and $h > 0$, and observing that

$$F(c + h) - F(c) = \int_a^{c+h} f - \int_a^c f = \int_c^{c+h} f \qquad \text{(From (4.11))}.$$

We then notice that f is bounded (as our definition of integrability applied only to bounded functions) and therefore there exists $M > 0$ such that $|f(x)| \leq M$ for all $x \in [a, b]$. Therefore, using the monotonicity property, we have that

$$-Mh \leq \int_c^{c+h} f \leq Mh,$$

Which implies that

$$|F(c + h) - F(c)| \leq Mh.$$

In view of the last inequality, given $\epsilon > 0$ it is enough to let $h < \epsilon/M$ to guarantee that for all $x \in [c, c + h]$ it will follow that $|F(c + h) - F(c)| < \epsilon$. If $h < 0$, then

$$F(c + h) - F(c) = \int_a^{c+h} f - \int_a^c f = \int_{c+h}^c f\,,$$

but f is bounded, and the length of the interval $[c+h, c]$ is $-h$ (the minus sign stems from the fact that $h < 0$) from which we obtain

$$(-M)(-h) \leq \int_c^{c+h} f \leq -Mh.$$

From here, we see again that

$$|F(c + h) - F(c)| \leq -Mh,$$

and thus, given $\epsilon > 0$ we only need to take $-h < \epsilon/M$ to ensure that for all $x \in [c+h, c]$ it follows that $|F(c+h) - F(c)| < \epsilon$. This proves the continuity of F over the open interval (a, b). Finally, to prove continuity at the endpoints, it is enough to observe that at the left endpoint a, the one sided limit $h > 0$ must be taken, while at the right endpoint b, the one sided limit $h < 0$ proves the result. Hence, F is continuous over the closed interval $[a, b]$. □

The previous result states that, upon integration, an integrable function f that may not be continuous itself, gives rise to a continuous function. It turns out that if the function f is continuous, integrating it will give rise to a differentiable function. However, before precisely stating and proving this, we will first prove an auxiliary result.

Lemma 5.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous at a point c. Define*

$$m_h := \inf\{f(x) : c \leq x \leq x+h\} \qquad \textit{and} \qquad M_h := \sup\{f(x) : c \leq x \leq x+h\}. \tag{5.2}$$

Then

$$\lim_{h\to 0} m_h = f(c) = \lim_{h\to 0} M_h.$$

Before attempting to prove the statement, we first discuss the conclusion intuitively. The result states that, if we have a function defined over an interval containing a point of continuity c, and we shrink the interval into said point of continuity, the value of the supremum and infimum over the interval converge to the value $f(c)$. Clearly continuity must play a key role, since it would be possible to have an otherwise constant function, which is undefined at a point x_0, and a sequence of intervals that shrink into x_0. For such a function, the infimum and supremum over any interval containing x_0 would remain constant, but the limit as the interval shrinks would not equal the value of the function at x_0 (since the function is undefined at the point).

Proof. We start by considering the case $h > 0$. Given an arbitrary $\epsilon > 0$ we observe that, since f is continuous at c, it is always possible to find δ such that

$$|c - x| < \delta \;\Rightarrow\; |f(c) - f(x)| < \epsilon/2. \tag{5.3}$$

For such a δ, we define

$$m_\delta := \inf\{f(x) : c \leq x \leq x+\delta\} \qquad \text{and} \qquad M_\delta := \sup\{f(x) : c \leq x \leq x+\delta\},$$

and remark that from the definitions in terms of infimum and supremum, it follows that there exist points x_* and x^* in the interval $[c, c+\delta]$ such that

$$f(x_*) - m_\delta < \epsilon/2 \qquad \text{and} \qquad M_\delta - f(x^*) < \epsilon/2.$$

Now, recalling that $h < \delta$ implies that $m_\delta < m_h$ and $M_h < M_\delta$, we have for all $h < \delta$

$$f(x_*) - m_h < f(x_*) - m_\delta < \epsilon/2\,, \tag{5.4a}$$
$$M_h - f(x^*) < M_\delta - f(x^*) < \epsilon/2. \tag{5.4b}$$

We can now combine (5.3) and (5.4a) to conclude that for all $h < \delta$

$$|f(c) - m_h| \leq |f(c) - f(x_*)| + |f(x_*) - m_h| < \epsilon,$$

while putting (5.3) and (5.4b) together yields

$$|f(c) - M_h| \leq |f(c) - f(x^*)| + |f(x^*) - M_h| < \epsilon.$$

The case $h < 0$ is analyzed in a completely analogous manner. Therefore, we have shown that for any $\epsilon > 0$ it is possible to find $\delta > 0$ such that if $|h| \leq \delta$, then

$$|f(c) - m_h| < \epsilon \qquad \text{and } |f(c) - M_h| < \epsilon$$

which proves the result. □

We will make use of this lemma in proving our next major result.

Theorem 5.2 (***First fundamental theorem of calculus***). *Let $f : [a, b] \to \mathbb{R}$ be integrable over $[a, b]$ and continuous for some $c \in (a, b)$. Then, the function $F(x) : [a, b] \to \mathbb{R}$ given by (5.1) is differentiable at $x = c$, and*

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

Proof. Let us consider $h > 0$, the interval $[c, c + h]$, and the associated infimum m_h and supremum M_h as defined in (5.2). Recalling that

$$F(c + h) - F(c) = \int_c^{c+h} f,$$

and that $m_h \leq f \leq M_h$ over the interval $[c, c + h]$, we see that

$$m_h h \leq F(c + h) - F(c) \leq M_h h.$$

Since $h > 0$ this implies that

$$m_h \leq \frac{F(c + h) - F(c)}{h} \leq M_h.$$

Appliying the conclusion of Lemma 5.1 to the inequalities above, yields

$$f(c) \leq \lim_{h \to 0} \frac{F(c + h) - F(c)}{h} \leq f(c).$$

The case $h < 0$ is treated analogously. □

Definition 5.1 (***Antiderivative***). If $f : [a, b] \to \mathbb{R}$ is integrable and continuous over (a, b), then, by Theorem 5.2, the function $F : (a, b) \to \mathbb{R}$ given by (5.1) is differentiable and satisfies

$$F'(x) = f(x) \qquad \text{for all } x \in (a, b).$$

In this case, F is called an **antiderivative** of f.

Remark 5.1. *Note that if f is not continuous, the function F may not be differentiable. Therefore the denomination of F as the antiderivative of f* ***only applies when f is continuous****.*

Antiderivatives, when they exist, are not unique since for any constant c, if F is an antiderivative of f it follows that $(F(x) + c)' = f(x)$ and therefore $F(x) + c$ is also an antiderivative of $f(x)$.

Theorem 5.3 (***Second fundamental theorem of calculus***). *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$, differentiable on (a, b) and such that $f'(x)$ is integrable over $[a, b]$. Then*

$$\int_a^b f' = f(b) - f(a). \tag{5.5}$$

Proof. Let $P := \{t_0 = a, t_1, \ldots, t_{n-1}, t_n = b\}$ be a partition of the interval $[a, b]$ and observe that

$$f(b) - f(a) = \sum_{i=1}^{n} \left(f(t_i) - f(t_{i-1}) \right). \tag{5.6}$$

On the other hand, since f is continuous over $[a, b]$ and differentiable over (a, b), these same properties are true within any subinterval $[t_{i-1}, t_i]$ in the partition P. Therefore, applying the mean value theorem for each subinterval, we have that for all $i \in \{1, \ldots, n\}$ there exists $x_i \in (t_{i-1}, t_i)$ such that

$$f(t_i) - f(t_{i-1}) = f'(x_i)(t_i - t_{i-1}). \tag{5.7}$$

Moreover, using the notation introduced in (4.1) for the infimum m_i and supremum M_i of the values of f ver the subinterval $[t_{i-1}, t_i]$, we have that

$$m_i \leq f(x_i) \leq M_i. \tag{5.8}$$

Combining (5.6), (5.7), and (5.8) it follows that

$$L(f', P) = \sum_{i=1}^{n} m_i(t_i - t_{i-1}) \leq \sum_{i=1}^{n} f'(x_i)(t_i - t_{i-1}) = f(b) - f(a) \leq \sum_{i=1}^{n} m_i(t_i - t_{i-1}) = U(f', P).$$

Moreover, from the definition of the integral of f' it also follows that

$$L(f', P) \leq \int_a^b f' \leq U(f', P).$$

Therefore, subtracting the two inequalities above, we obtain that, for any partition P it must hold

$$-\big(U(f', P) - L(f', P)\big) \leq \int_a^b f' - (f(b) - f(a)) \leq U(f', P) - L(f', P).$$

However, since f' is integrable, for any $\epsilon > 0$ we can always find a partition such that $U(f', P) - L(f', P) < \epsilon$ and therefore, combining this with the inequality above we obtain that, for any $\epsilon > 0$ we must have

$$-\epsilon \leq \int_a^b f' - (f(b) - f(a)) \leq \epsilon,$$

which proves (5.5). □

The first fundamental theorem of calculus tells how the properties of a function defined by an integral relate to the properties of the integrand. The second fundamental theorem of calculus tells us how to compute a definite integral if we happen to know the antiderivative of the integrand. We can use these two results to prove two additional useful and familiar tools from calculus.

Theorem 5.4 (***Integration by parts***). *Let f, g be real-valued functions continuous over the interval $[a, b]$, differentiable over the interval (a, b) and such that their derivatives f' and g' are integrable over the interval $[a, b]$. Then, the derivative $(fg)'$ is integrable and*

$$\int_a^b f'g + \int_a^b g'f = f(b)g(b) - f(a)g(b).$$

Proof. Since we have assumed f, g, f', and g' to be all integrable, and by Problem 5 on Chapter 4 the product of integrable functions is integrable, we have that $(fg)' = f'g + g'f$ is integrable. Therfore, by the second fundamental theorem of calculus, it follows that

$$\int_a^b f'g + \int_a^b g'f = \int_a^b (fg)' = f(b)g(b) - f(a)g(b).$$

□

Theorem 5.5 (***Change of variable***)**.** *Let $u : [a, b] \to \mathbb{R}$ be continuous over $[a, b]$ and differentiable over (a, b), $I \subset \mathbb{R}$ be an open interval containing the image of $[a, b]$ under u (i.e. such that $u([a, b]) \subset I$ and let $f : I \to \mathbb{R}$ be continuous over I. Then the following holds*

$$\int_a^b f(u(x)) \cdot u'(x)\, dx = \int_{u(a)}^{u(b)} f(u)\, du\,. \tag{5.9}$$

Proof. Fix a point $c \in I$. Since f is continuous over f, it follows by the first fundamental theorem of calculus that the function

$$F(u) := \int_c^u f(u)\, du$$

is differentiable for all $u \in I$ and $F'(u) = f(u)$. Moreover, since u is differentiable for $x \in (a, b)$, defining $g(x) := (F \circ u)(x)$ from the chain rule see that

$$g'(x) = (F \circ u)'(x) = F'(u(x)) \cdot u'(x) = f(u(x)) \cdot u'(x).$$

Therefore it follows that

$$\begin{aligned}
\int_a^b f(u(x)) \cdot u'(x)\, dx &= \int_a^b g'(x)\, dx \\
&= g(b) - g(a) && \text{(By the second fundamental theorem of calculus)} \\
&= F(u(b)) - F(u(a)) && \text{(From the definition of } g\text{)} \\
&= \int_c^{u(b)} f(x)\, dx - \int_c^{u(a)} f(x)\, dx && \text{(From the definition of } F\text{)} \\
&= \int_{u(a)}^{u(b)} f(u)\, du,
\end{aligned}$$

as desired. □

5.1 Exercises

1. (a) Complete the proof of Lemma 5.1 by considering the case $h < 0$. Follow carefully the argument provided in the text, as it only needs some minor modifications to deal with $h < 0$.

   (b) Complete the proof of The first fundamental theorem of calculus (Theorem 5.2) by considering the case $h < 0$. Follow carefully the argument provided in the text, as it only needs some minor modifications to deal with $h < 0$.

2. Let $f : \mathbb{R} \to \mathbb{R}$ be integrable on $[a, b]$.

(a) Show that there exists $x \in [a,b]$ such that $\displaystyle\int_a^x f(t)\,dt = \int_x^b f(t)\,dt$.

(b) Show, by constructing an example, that it may not always be possible to find such an x in (a,b).

3. Let f be a continuous function on $\mathbb{R}$ and define

$$F(x) := \int_{x-1}^{x+1} f(t)\,dt \quad \text{for} \quad x \in \mathbb{R}.$$

Show that F is differentiable and compute F'.

4. Let f be integrable on $[a,b]$, $c \in (a,b)$ and

$$F(x) := \int_a^x f \qquad \text{for all } x \in [a,b].$$

Prove the following. You can use the fundamental theorem of calculus.

(a) If f is differentiable at c, then F is also differentiable at c.

(b) If f is differentiable at c, then F' is continuous at c.

5. Prove that if h is continuous, f and g are differentiable, and

$$F(x) := \int_{f(x)}^{g(x)} h(t)\,dt,$$

then $\quad F'(x) = h(g(x)) \cdot g'(x) - h(f(x)) \cdot f'(x)$.

[Hint: try to split F in two parts, each of them having a constant limit of integration]

Chapter 6

Sequences and series of functions

Definition 6.1 (***Pointwise convergence***)**.** Let $A \subset \mathbb{R}$, $f : A \to \mathbb{R}$, and $\{f_n\}_{n=1}^{\infty}$ be a sequence of functions such that $f_n : A \to \mathbb{R}$ for all $n \in \mathbb{N}$. We say that the sequence **converges pointwise** on A to f if, for every $\epsilon > 0$ and $x \in A$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$

$$|f_n(x) - f(x)| < \epsilon.$$

In this case we can write "$f_n \to f$ *pointwise on* A".

Remark 6.1. *The definition above says that the sequence of functions f_n converges pointwise to f if, for every value of $x \in A$, the sequence of real numbers that is obtained by evaluating every f_n at x converges to the number that is obtained by evaluating f at the same value of x. Note that, since the value of both the limit $f(x)$ and each of the terms in the sequence, $f_n(x)$, depend on the point x where the functions are evaluated, the value of N in the definition above will, in general, also depend on x as well as on ϵ.*

Example 6.1. Consider the family of functions $f_n : \mathbb{R} \to \mathbb{R}$ given by $f_n(x) := (\cos(x))^n$. The graph of some of these functions for different values of n is shown in Figure 6.1. Each of these functions is continuous and can be differentiated infinitely many times for any value of x. We now study the pointwise convergence of the sequence.

First, let $m \in \mathbb{Z}$ and observe that $|\cos x| < 1$ for any $x \neq m\pi$. Therefore, it is easy to show that for such values of x, $f_n(x) \to 0$ as $n \to \infty$. For even multiples of π, the value of cosine is equal to 1 and thus $f_n(x) \to 1$ if $x = 2m\pi$. However, if $x = (2m+1)\pi$ then as $n \to \infty$ the value of cosine alternates between 1 and -1; hence, the sequence is not convergent for these values of x. It can therefore be proven that the poitwise limit is given by

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 1 & \text{if } x = 2m\pi \\ \text{DNE} & \text{if } x = (2m+1)\pi \\ 0 & \text{otherwise} \end{cases} \qquad (\text{for } m \in \mathbb{Z}).$$

The pointwize limit above has countably many discontinuities and is not even defined for countably many values of x. This example highlights one drawback of pointwise convergence: many of the properties of the functions in the sequence (like continuity and differentiability) are not inherited by the pointwise limit.

Definition 6.2 (***Uniform convergence***)**.** Let $A \subset \mathbb{R}$, $f : A \to \mathbb{R}$, and $\{f_n\}_{n=1}^{\infty}$ be a sequence of functions such that $f_n : A \to \mathbb{R}$ for all $n \in \mathbb{N}$. We say that the sequence **converges uniformly** on A to f if, for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ and $x \in A$

$$|f_n(x) - f(x)| < \epsilon.$$

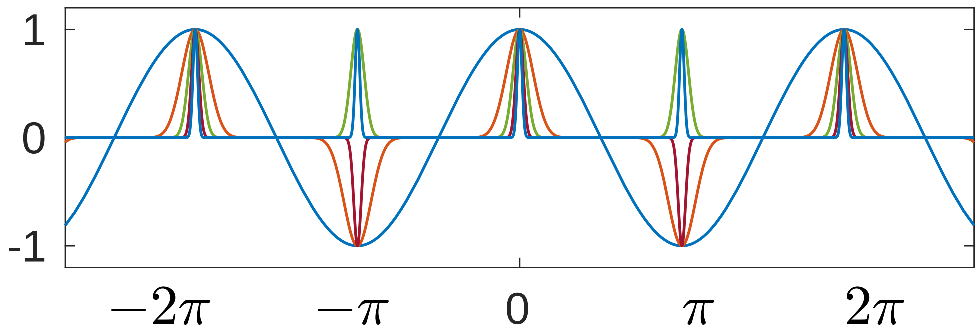

Figure 6.1: Plot of $(\cos(x))^n$ for increasing values of n. The sequence converges pointwise to 0 for values of x different from integer multiples of π, it converges pointwise to 1 for even multiples of π and fails to converge for odd multiples of π.

In this case we can write "$f_n \to f$ *uniformly on* A".

Remark 6.2. *Note that the "only" difference between the definition of pointwise convergence and uniform convergence is that the index N depends on the point x in the pointwise case, but is independent of x in the uniform case. This subtle difference has major implications. The definition of uniform convergence has a natural geometric interpretation: given $\epsilon > 0$, there exists some N such that for all $n \geq N$ and all $x \in A$ we have*

$$f(x) - \epsilon < f_n(x) < f(x) + \epsilon.$$

As shown in Figure 6.2, this inequality implies that for each ϵ the graphs of infinitely many terms of the sequence lie inside of a tubular region (with width 2ϵ) around the graph of the limiting function.

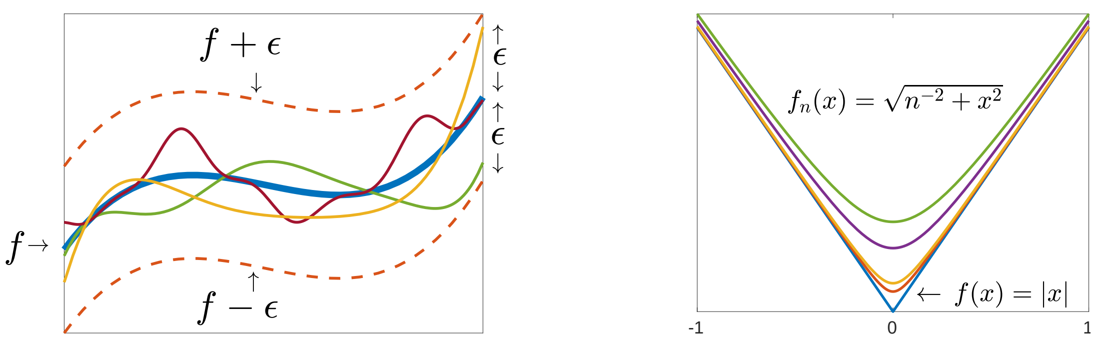

Figure 6.2: Left: If a sequence of functions converges uniformly to a limit function f, for any $\epsilon > 0$ the graphs of infinitely many functions f_n are contained within a strip of length 2ϵ around the graph of the limit. Right: The sequence $f_n(x) := \sqrt{n^{-2} + x^2}$ converges uniformly to $f(x) = |x|$ (depicted in blue) and every f_n is differentiable; nevertheless, the limit is not differentiable at $x = 0$.

Definition 6.3 (***Uniform Cauchy sequence***)**.** We say that a sequence of functions $f_n : A \subset \mathbb{R} \to \mathbb{R}$ is **uniformly Cauchy** if for every $\epsilon > 0$ there exist $N \in \mathbb{R}$ such that for all $n, m \geq N$ and $x \in A$ it holds that

$$|f_n(x) - f_m(x)| < \epsilon.$$

As the following theorem shows, the definition above provides us with an alternative method of determining wether a sequence of functions converges uniformly *without the need to the pointwise limit.*

Theorem 6.1 (***Cauchy criterion* for uniform convergence).** *Let $A \subset \mathbb{R}$, $f : A \to \mathbb{R}$, and $\{f_n\}_{n=1}^{\infty}$ be a sequence of functions such that $f_n : A \to \mathbb{R}$ for all $n \in \mathbb{N}$. The sequence f_n converges uniformly to a function $f : A \to \mathbb{R}$ if and only if the sequence f_n is uniformly Cauchy.*

Proof. $\boxed{\Rightarrow}$ Since the sequence converges uniformly to f, given $\epsilon > 0$, there exists N independent from x such that for all $n, m \geq N$ we have

$$|f_n(x) - f(x)| < \epsilon/2 \qquad \text{and} \qquad |f_m(x) - f(x)| < \epsilon/2.$$

Therefore, for all $n, m \geq N$ and $x \in A$

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f_m(x) - f(x)| < \epsilon/2 + \epsilon/2 = \epsilon$$

and the sequence is uniformly Cauchy.

$\boxed{\Leftarrow}$ Since the sequence is uniformly Cauchy, we know that for all $x \in A$ the sequence of real numbers $f_n(x)$ is Cauchy. Since every Cauchy sequence of real numbers is convergent (as proven in Problem 4 from Chapter 2) we know that for every $x \in A$ there exists a $y \in \mathbb{R}$ such that

$$\lim_{n\to\infty} f_n(x) = y.$$

Moreover, this limit is uniform. Therefore, we can define $f(x) : A \to \mathbb{R}$ by

$$f(x) := \lim_{n\to\infty} f_n(x).$$

It follows from the fact that the limit is uniform that for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$ and $x \in A$

$$|f_n(x) - f(x)| < \epsilon$$

and the sequence converges uniformly to f on A. □

The next theorem (whose proof will be left as an exercise) provides yet another way of verifying if a sequence of functions converges uniformly to a limit.

Theorem 6.2. *A sequence of functions $f_n(x) : A \subset \mathbb{R} \to \mathbb{R}$ converges uniformly to a function $f : A \to \mathbb{R}$ if and only if*

$$\lim_{n\to\infty} \sup\{|f(x) - f_n(x)| : x \in A\} = 0.$$

Example 6.1 showcases that pointwise convergence of a family of continuous functions is not enough to ensure the continuity of the limit. As we shall soon see, uniform convergence of a sequence is strong enough to overcome the issue. However, before stating and proving that result, we will first attempt a naïve proof that uses *only pointwise convergence* and will identify where the problem arises. This will then shed light on why uniform convergence suceeds where pointwise convergence does not.

Example 6.2. Assume that every function of a sequence $f_n : A \subset \mathbb{R} \to \mathbb{R}$ is continuous at some point $x_0 \in A$, and that the sequence converges *pointwise* to a limit function f. To show that f is continuous at x_0, we would have to prove that if x and x_0 are *close enough*, then $|f(x) - f(x_0)|$ can be made arbitrarily small. To aid us in this task we have only two tools:

1. Pointwise convergence which, *for a fixed point* x allows us to squeeze the distance $|f_n(x) - f(x)|$ by increasing the value of n.
2. Continuity of each f_n at x_0 wich, *for a fixed value of* n, allows us to squeeze the distance $|f_n(x) - f_n(x_0)|$ by reducing the distance between x and x_0.

Note that the term $|f(x) - f(x_0)|$, which is what we want to control, involves none of our tools: the functions f_n do not appear in the expression and thus we can not use n to reduce the distance; moreover, the function f is not known to be continuous (yet) and thus reducing the distance between x and x_0 is not useful. We must then introduce the useful terms f_n into the mix by "adding zeros" and using the triangle inequality in the following way:

$$|f(x_0) - f(x)| \leq \underbrace{|f(x_0) - f_n(x_0)|}_{①} + \underbrace{|f_n(x_0) - f_n(x)|}_{②} + \underbrace{|f_n(x) - f(x)|}_{③}.$$

Since the term ① involves a single, *fixed* value of x_0, then pointwise convergence guarantees that we can find some value of $N(x_0)$ such that, *for this particular point* x_0, the value of ① can be made smaller than $\epsilon/3$ as long as $n \geq N(x_0)$. The notation $N(x_0)$ deliberately stresses the fact that pointwise convergence implies that the value of N will, in general, depend on the point x_0. Therefore, given a particular value of ϵ, choosing a point x_0 will immediately fix the value of $N(x_0)$ required to bound the term ①.

We then move to the second term. The expression in ② involves *the same* function f_n evaluated at two different points x and x_0. Due to the fact that each f_n is continuous at x_0, we then know that *for each* function f_n, there exists $\delta(n, x_0) > 0$ such that if $|x - x_0| < \delta(n, x_0)$, then $|f_n(x_0) - f_n(x)|$ will be smaller than $\epsilon/3$. As before, the notation $\delta(n, x_0)$ deliberately stresses the fact that the value of δ will, in general, depend on the particular function f_n and on the point of continuity x_0. At this moment, it is very important to remark two points: 1) our admissible values of n were already fixed by the choice of N made to control ①, and 2) the value of δ has also now been fixed in order to control ②.

We then move to the final term ③, which involves different functions evaluated at the same point x. We would like to use convergence to make this term small, however, this would require changing the value of N that has already been fixed (to control ①). At this point it is tempting to simply choose the largest value of N of the ones required for ① and ③. However, taking the largest of the two potentially changes the value of N, which will then affect the value of δ required to control ② (since δ depends on the index n as well). This new change in δ limits the interval where we can choose x from, therefore potentially affecting our control of the term ③. If the point x in ③ changes the value of N will change in turn and the cycle starts all over again.

The source of the problem is that we need to control three different terms but have only two parameters that can be used: $N(x)$ and $\delta(n, x)$. If N were independent of the point x the problem would be solved, since then fixing the value of N to control ① would at once also control ③, *for any* choice of δ. Therefore, if pointwise convergence is strenghtened into *uniform* convergence, continuity can be proven by a reasoning like the one above, as we shall now see. Proves that involve a reasoning like the one that we just described are often referred to as an $\epsilon/3$ *proof* (read as *epsilon thirds*).

Theorem 6.3 (***Uniform limit theorem***)**.** *Let $f_n : A \subset \mathbb{R} \to \mathbb{R}$ be a sequence of functions all of which are continuous at some $x_0 \in A$ and such that f_n converges uniformly on A to some function $f : A \subset \mathbb{R} \to \mathbb{R}$. Then f is continuous at x_0.*

Proof. Let $\epsilon > 0$ and observe that, since $f_n \to f$ uniformly on A, it is possible to find $N \in \mathbb{N}$ such that for all $y \in A$ and $n \geq N$ we have

$$|f(y) - f_n(y)| < \epsilon/3.$$

Moreover, since very f_n is continuous at x_0, for the value of N found above there is δ_N such that

$$|x - x_0| < \delta_N \quad \Rightarrow |f_N(x) - f_N(x_0))| < \epsilon/3.$$

Therefore, we have that if $|x - x_0| < \delta_N$

$$|f(x_0) - f(x)| \leq |f_N(x_0) - f_N(x_0)| + |f_N(x_0) - f_N(x)| + |f_N(x) - f(x)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon,$$

and f is continuous at x_0. □

We now explore the connection between uniform convergence and differentiability. As we saw in Example 6.1, pointwise convergence of a sequence of differentiable functions is not strong enough to preserve the differentiability of the limit. In view of the uniform limit theorem 6.3, one would be tempted to believe that uniform convergence of a sequence of differentiable functions might be enough to guarantee the differentiablity of the limit. As the following example will show, this is in fact not the case.

Example 6.3. Consider the sequence of functions defined by $f_n(x) = \sqrt{1/n^2 + x^2}$ defined over $\mathbb{R}$. The right panel of Figure 6.2 shows the plots of some functions in the sequence for different values of n. Each of these functions is differentiable with $f_n'(x) = \dfrac{x}{\sqrt{1/n^2 + x^2}}$.

However, notice that for all $n \in \mathbb{N}$

$$x^2 < 1/n^2 + x^2 = f_n^2(x) \leq 1/n^2 + 2x/n + x^2 = (1/n + x)^2.$$

Therefore, recalling that $|y| = \sqrt{y^2}$ for all y, the inequalities above imply

$$|x| < f_n(x) \leq |1/n + x| \leq 1/n + |x| \quad \Rightarrow \quad 0 < f_n(x) - |x| \leq 1/n.$$

Since $1/n$ can be made arbitrarily small *independently of* x, this implies that $f_n(x) \to |x|$ uniformly. But $|x|$ is not differentiable at $x = 0$, while all of the functions $f_n(x)$ is differentiable for all x.

The previous example shows that uniform convergence is not enough to preserve differentiability of the limit. It turns out that much more is required to ensure the differentiability of the limit. The proof is also an $\epsilon/3$ argument similar to the one appearing in the proof of the uniform limit theorem 6.3, where uniform convergence saves the day.

Theorem 6.4 (*Differentiable limit theorem*). *Let $f_n : A \subset \mathbb{R} \to \mathbb{R}$ be a sequence of differentiable functions converging pointwise to some continuous function $f : A \subset \mathbb{R} \to \mathbb{R}$. Moreover, assume that the sequence of derivatives f_n' converges uniformly to some function $g : A \subset \mathbb{R} \to \mathbb{R}$. Then the pointwise limit f is differentiable and $f'(x) = g(x)$.*

Remark 6.3 (*Interchanging limits and derivatives*). *One consequence of the previous result is that, under the assumptions of the theorem,"limits and derivatives commute" since*

$$\lim_{n\to\infty} f_n'(x) = g(x) = f'(x) = \left(\lim_{n\to\infty} f_n(x)\right)'.$$

Remark 6.4. *Before starting the proof we will dissect the argument to make it more clear. We want to show that the derivative of the pointwise limit f is equal to the uniform limit of the derivatives g. In other words, we want to show that*

$$f'(x) = \lim_{y\to x} \frac{f(y) - f(x)}{y - x} = g(x).$$

This in turn implies showing that the difference $\left|g(x) - \frac{f(y)-f(x)}{y-x}\right|$ *can be made arbitrarily small by choosing* y *"close enugh" to* x*. To aid us in the task, we can use: pointwise convergence of* f_n*, uniform convergence of* f_n' *and differentiability of* f_n*. With this in mind, we estimate*

$$\left|g(x) - \frac{f(y)-f(x)}{y-x}\right| \leq \underbrace{\left|g(x) - f_n'(x)\right|}_{①} + \underbrace{\left|f_n'(x) - \frac{f_n(y)-f_n(x)}{y-x}\right|}_{②} + \underbrace{\left|\frac{f_n(y)-f_n(x)}{y-x} - \frac{f(y)-f(x)}{y-x}\right|}_{③}.$$

The term ① can be made small using uniform convergence of $f_n' \to g$*, which fixes the value of* N*. The term ② can then be controlled using the differentiability of each* f_n *by choosing* $\delta(N, x)$*. Finally ③ can be controlled using the pointwise convergence* $f_n \to f$*. However, note that the value of* N *that was fixed in the first step was chosen for the convergence of* f' *and therefore it must be adjusted for the last term (as it pertains to the convergence of* f_n *and not of its derivative* f_n'*). Once again, choosing the maximum value of* N *required to control both ① and ③, would then change the value of* $\delta(N, x)$ *used for ②, this would then may alter the point* y *appearing in ③ which would require changing* N *once more and restarting the cycle. The problem in this case stems from the fact that, now the value of* N *seems to change not depending on whether we are trying to control termswith functions* f_n *or their derivatives* f_n'*. If only we could connect the uniform convergence of the derivatives with uniform convergence of the functions, that would solve the issue. As we shall see, uniform convergence of the derivatives is so strong that in fact we will be able to establish such a connection. This will require a little of additional work.*

Proof. Define the function $g = f_n - f_m$ and observe that, since every f_n and f_m are differentiable, g is too. Therefore, we can pick any two points $x < y \in A$ and apply the mean value theorem to obtain the existence of $c \in (x, y)$ such that $g'(c) = (g(x) - g(y))/(x-y)$. Rewriting this expression in terms of f_n and f_m yields

$$f_n'(c) - f_m'(c) = \frac{f_n(x)-f_m(x)}{x-y} - \frac{f_n(y)-f_m(y)}{x-y} = \frac{f_n(x)-f_n(y)}{x-y} - \frac{f_m(x)-f_m(y)}{x-y},$$

where the last equality comes from a simple rearrangement of the terms. This implies that

$$\left|f_n'(c) - f_m'(c)\right| = \left|\frac{f_n(x)-f_n(y)}{x-y} - \frac{f_m(x)-f_m(y)}{x-y}\right|.$$

But the sequence f_n' is uniformly Cauchy (since it is uniformly convergent), therefore the equality above implies that the sequence $(f_n(x) - f_n(y))/(x-y)$ is also uniformly Cauchy. On the other hand, it is clear from the pointwise convergence of f_n to f that $(f_n(x) - f_n(y))/(x-y)$ converges pointwise to $(f(x) - f(y))/(x-y)$. Therefore, since the sequence is uniformly Cauchy, the pointwise limit is indeed a uniform limit (see problem 1) and thus

$$\frac{f_n(x)-f_n(x)}{x-y} \longrightarrow \frac{f(x)-f(x)}{x-y} \qquad \text{uniformly on } A.$$

Now let $\epsilon > 0$ and use the uniform convergence described above together with the uniform convergence of $f_n' \to g$ to pick an $N \in \mathbb{N}$ such that, for all $x, y \in A$ and $n \geq N$

$$\left|g(x) - f_n'(x)\right| \leq \epsilon/3 \qquad \text{and} \qquad \left|\frac{f_n(y)-f_n(x)}{y-x} - \frac{f(y)-f(x)}{y-x}\right| \leq \epsilon/3.$$

While, from the differentiability of f_n, it is possible to find δ such that if $|x-y| < \delta$ we have

$$\left|f_n'(x) - \frac{f_n(y)-f_n(x)}{y-x}\right| \leq \epsilon/3.$$

Putting these three ingredients together we conclude that, for any $\epsilon > 0$ there exists δ such that if $|x-y| < \delta$ then

$$\left|g(x) - \frac{f(y)-f(x)}{y-x}\right| \leq |g(x) - f_n'(x)| + \left|f_n'(x) - \frac{f_n(y)-f_n(x)}{y-x}\right| + \left|\frac{f_n(y)-f_n(x)}{y-x} - \frac{f(y)-f(x)}{y-x}\right| < \epsilon$$

establishing that f is differentiable at x and $f'(x) = g(x)$. $\square$

We now turn our attention towards the relation between convergence and integrability. Since integrability is a less finnicky property than differentiability or continuity (for instance, even discontinuous functions can be integrable), one would be tempted to believe that the pointwise limit of a sequence of integrable functions would be integrable. However, as the following example shows, this is not the case.

Example 6.4. Since the rational numbers between 0 and 1 are countable, it is possible to order them in a list tagged by natural indices. Let $\{q_1, q_2, q_3, \ldots\}$ be one such ordering of the rationals between 0 and 1 and $\mathbb{Q}_n := \{q_1, \ldots, q_n\}$. We then define the sequence of functions

$$f_n(x) := \begin{cases} 1 & \text{if } x \in \mathbb{Q}_n \\ 0 & \text{if } x \notin \mathbb{Q}_n \end{cases} \qquad \text{for } x \in [0,1].$$

Since every f_n is defined over a compact interval, is bounded, and is equal to 0 except for a finite number of points, Problem 4, they are all integrable. However, as $n \to \infty$ we have that f_n converges pointwise to $f\chi_{\mathbb{Q}\cap[0,1]}$ the Dirichlet function restricted to $[0,1]$. As it was shown in Example 4.1, this function is not integrable.

A related question is whether, if the pointwise limit happens to be integrable, the limit commutes with the integral. Unfortunately, if the convergence is only pointwise this is not always possible, as we will see now.

Example 6.5. Consider the sequence of functions $f_n(x) = nx^n$ for $x \in [0,1)$. Each function is integrable and

$$\int_0^1 f_n(x)\,dx = \frac{n}{n+1} \qquad \text{which implies} \qquad \lim_{n\to\infty} \int_0^1 f_n(x)\,dx = 1.$$

On the other hand we have that

$$\lim_{n\to\infty} \left|\frac{f_{n+1}(x)}{f_n(x)}\right| = \lim_{n\to\infty} \frac{n+1}{n}|x| < \lim_{n\to\infty} \frac{n+1}{n} = 1.$$

By the ratio test, this implies that for each $f_{n+1} \to 0$ pointwise, which is clearly integrable. Hence

$$\int_0^1 \lim_{n\to\infty} f_n(x)\,dx = \int_0^1 0\,dx = 0 \neq 1,$$

and thus pointwise limits and integrals do not commute.

Fortunately, uniform convergence is enough to ensure that the uniform limit of a sequence of integrable functions will be integrable and that the limit and the integral can be exchanged.

Theorem 6.5. *Let $f_n : [a,b] \to \mathbb{R}$ be a sequence of integrable functions that converge uniformly to a function f. Then the limit f is integrable and*

$$\int_a^b f = \lim_{n\to\infty} \int_a^b f_n(x)\,dx.$$

Proof. We will start by proving that the limit function f is integrable. The first order of business is to verify that f is bounded, since we have not defined the integral for unbounded functions. Note that every f_n is bounded (since it is integrable) and let M_n be the bound for the n-th function, so that $|f_n(x)|$. Now, since the sequence f_n is uniformly convergent to f, there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ and all $x \in [a, b]$ we have $|f(x) - f_N(x)| < 1$. Using the reverse triangle inequality, the previous statement ensures that for all $n \geq N$ and all $x \in [a, b]$ it follows that $|f(x)| < |f_N(x)| + 1 \leq M_N + 1$. Which proves that f is bounded with bound $M := M_N + 1$.

We then choose a partition $P := \{t_0, t_1, \ldots, t_J\}$ of $[a, b]$ and denote $\Delta_i := t_i - t_{i-1}$ and

$$m_i := \inf\{f(x) : x \in [t_{i-1}, t_i]\}, \qquad M_i := \sup\{f(x) : x \in [t_{i-1}, t_i]\},$$
$$\widetilde{m}_i := \inf\{f_n(x) : x \in [t_{i-1}, t_i]\}, \qquad \widetilde{M}_i := \sup\{f_n(x) : x \in [t_{i-1}, t_i]\},$$

Since $f_n \to f$ uniformly, we know that for any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ and all $x \in [a, b]$ we have $|f(x) - f_n(x)| \leq \frac{\epsilon}{4(b-a)}$. The reason behind the unusual choice $\frac{\epsilon}{4(b-a)}$ will become clear soon. The important point is that the value of $|f(x) - f_n(x)|$ can bemade arbitrarily small. Rearranging the last inequality leads to

$$-\frac{\epsilon}{4(b-a)} + f_n(x) \leq f(x) \leq \frac{\epsilon}{4(b-a)} + f_n(x) \quad \forall x \in [a, b],$$

from which we can deduce that, for every subinterval (t_{i-1}, t_i) in a partition $P := \{t_0, t_1, \ldots, t_J\}$ of $[a, b]$, it holds

$$\widetilde{m}_i \leq m_i + \frac{\epsilon}{4(b-a)} \qquad \text{and} \qquad M_i \leq \widetilde{M}_i + \frac{\epsilon}{4(b-a)}.$$

Therefore, subtracting the first inequality from the second one we obtain

$$0 \leq M_i - m_i \leq \widetilde{M}_i - \widetilde{m}_i + \frac{\epsilon}{2(b-a)}.$$

Multiplying both sides of the inequality above by $t_i - t_{i-1}$ and adding over all the subintervals in the partition P leads to

$$0 \leq U(f, P) - L(f, P) \leq U(f_n, P) - L(f_n, P) + \frac{\epsilon}{2}.$$

However, since f_n is integrable, we can choose a particion such that $U(f_n, P) - L(f_n, P) \leq \frac{\epsilon}{2}$ which finally leads to

$$0 \leq U(f, P) - L(f, P) < \epsilon,$$

and we conclude that f is integrable.

Finally use that $f_n \to f$ unifformly to pick $N \in \mathbb{N}$ such that, for all $n \geq N$ we have that $|f - f_n| < \frac{\epsilon}{b-a}$ to see that

$$\left| \int_a^b f - \int_a^b f_n \right| \leq \int_a^b |f - f_n| < \int_a^b \frac{\epsilon}{b-a} = \epsilon.$$

This proves that

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b f,$$

as we were trying to show. □

Since series are defined as limits of finite sums, the results that we have proven for sequences can now be easily translated into series, as we shall now see. Consider a sequence of functions $f_n : A \subset \mathbb{R} \to \mathbb{R}$. For each n we define the function $S_n : A \subset \mathbb{R} \to \mathbb{R}$ by

$$S_n(x) := \sum_{i=1}^{n} f_i(x)$$

and will call it the n-th ***partial sum***. The collection of partial sums defines a new sequence of functions $\{S_n\}$ whose convergence properties we can study. Whenever the sequence of partial sums $\{S_n\}$ converges pointwise we define the function

$$\sum_{n=1}^{\infty} f_n(x) := \lim_{n\to\infty} S_n(x) \qquad \text{for } x \in A,$$

(which is referred to as a **series**) and say that **the series converges**. If the sequence $\{S_n\}$ converges uniformly we say that the limit is a ***uniformly convergent series***.

Series of functions are very useful in both pure an applied mathematics as a tool for approximation. A complicated function can often be interpreted as the limit of a sum of simpler functions. Taylor series and Fourier series are two well-known examples of such an approximation technique which is also often called a ***series expansion***. We would then want to use the properties (continuity, differentiabily, integrability) of each of the terms of the series to deduce properties of the limiting function—since each of the terms in the series is typically easier to differentiate or integrate.

The theorems that we proved previously for sequences can now be used to determine under what conditions the limit inherits the properties of the summands and when can the infinite sum commutes with derivatives, limits or integrals. The proofs of all of the following results are simple consequences of the analogous results for sequences of functions.

Theorem 6.6. *Let $f_n : A \subset \mathbb{R} \to \mathbb{R}$ be a continuous function for every $n \in \mathbb{N}$. If the series $\sum_{n=1}^{\infty} f_n$ converges uniformly then*

$$f(x) := \sum_{n=1}^{\infty} f_n(x)$$

is continuous for every $x \in [a, b]$.

Proof. Since every f_n is continuous and the sum of continuous functions is continuous, every partial sum $S_n(x) = \sum_{i=1}^{n} f_i$ is a continuous function. Therefore, since $f(x) := \sum_{n=0}^{\infty} f_n(x)$ is the uniform limit of continuous functions, its continuity is a consequence of the uniform limit theorem 6.3. ☐

Theorem 6.7 (***Differentiation term by term***). *Let $f_n : A \subset \mathbb{R} \to \mathbb{R}$ be a differentiable function for every $n \in \mathbb{N}$ and let the series $\sum_{n=0}^{\infty} f_n$ converge to a continuous function f. If the series of derivatives $\sum_{n=1}^{\infty} f_n'$ converges uniformly to a function g, then f is differentiable on A and*

$$f'(x) := \sum_{n=1}^{\infty} f_n'(x).$$

Proof. Since every f_n is differentiable and a finite sum of differrentiable functions is differentiable, every partial sum $S_n(x) = \sum_{i=1}^{n} f_i(x)$ is differentiable and continuous. Moreover, denoting the n-th partial sum of the derivatives by $G_n(x) := \sum_{i=1}^{n} f_i'(x)$, the linearity of the derivative yields

$$S_n'(x) = \left(\sum_{i=1}^{n} f_i(x)\right)' = \sum_{i=1}^{n} f_i'(x) = G_n(x).$$

Therefore, we have that every function in the sequence $\{S_n\}$ is differentiable and coverges to a continuous function f and the sequence of derivatives $S_n' = G_n$ converges uniformly to a function g. These are all the hypotheses of the differentiable limit theorem 6.4 and therefore we have that $f(x) := \sum_{n=1}^{\infty} f_n(x)$ is differentiable and

$$f'(x) = g(x) = \sum_{n=1}^{\infty} f_n'(x).$$

□

Theorem 6.8 (***Integration term by term***). *Let $f_n : [a, b] \to \mathbb{R}$ be integrable for every $n \in \mathbb{N}$. If the series $\sum_{n=1}^{\infty} f_n$ converges uniformly then $f(x) := \sum_{n=1}^{\infty} f_n(x)$ is integrable on $x[a, b]$ and*

$$\int_a^b f = \sum_{n=1}^{\infty} \int_a^b f_n.$$

Proof. Given that every f_n is integrable and the sum of integrable functions is integrable, every partial sum $S_n := \sum_{i=1}^{n} f_i$ is integrable and, by the linearity of the integrals

$$\int_a^b \sum_{i=1}^{n} f_i = \sum_{i=1}^{n} \int_a^b f_i.$$

Since the sequence S_n converges uniformly, by Theorem 6.5 the limit function f is integrable and

$$\int_a^b f = \lim_{n\to\infty} \int_a^b S_n = \lim_{n\to\infty} \int_a^b \sum_{i=1}^{n} f_i = \lim_{n\to\infty} \sum_{i=1}^{n} \int_a^b f_i = \sum_{i=1}^{\infty} \int_a^b f_i,$$

as we were trying to prove. □

All the preceding theorems underline the importance of the uniform convergence of a series expansion. We now prove two results that prove very useful when trying to verify if a series is ideed uniformly convergent.

Theorem 6.9 (***Uniform Cauchy criterion for series***). *Let $f_n : A \subset \mathbb{R} \to \mathbb{R}$ be a sequence of functions. The series $\sum^{\infty} f_n$ is uniformly convergent if and oly if for any $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n, m > M$ and all $x \in A$*

$$|f_n(x) + \cdots + f_m(x)| < \epsilon.$$

Proof. The series $\sum^{\infty} f_n$ is uniformly convergent if and only if the sequence of partial sums $S_n(x) = \sum_{i=1}^{n} f_i(x)$ is uniformly Cauchy (by Theorem 6.1). This happens if and only if for any given $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n, m > M$ and all $x \in A$

$$\epsilon > |S_m(x) - S_n(x)| = \left| \sum_{i=1}^{m} f_i(x) - \sum_{i=1}^{n} f_i(x) \right| = |f_n(x) + \cdots + f_m(x)|$$

as desired. □

Theorem 6.10 (***Weierstrass M-test***). *Let $f_n : A \subset \mathbb{R} \to \mathbb{R}$ be a sequence of functions such that, for each n, there exists $M_n > 0$ such that $|f_n(x) \leq M_n|$ for all $x \in A$. If the series $\sum_{n=1}^{\infty} M_n$ converges, then the series $\sum_{n=1}^{\infty} f_n$ is uniformly convergent.*

Proof. Since the series of bounds M_n is convergent, the sequence of its partial sums is Cauchy, and therefore, given $\epsilon > 0$ there exists $M \in \mathbb{N}$ such that, for all $n, m \geq M$

$$\epsilon > \left| \sum_{i=1}^{m} M_i - \sum_{i=1}^{n} M_i \right| = \left| \sum_{i=n}^{m} M_i \right| = \sum_{i=n}^{m} M_i \geq \sum_{i=n}^{m} |f_i(x)| \geq \left| \sum_{i=n}^{m} f_i(x) \right| = |f_n(x) + \cdots + f_m(x)| .$$

Therefore, by The uniform Cauchy criterion for series (Theorem 6.9), the series $\sum_{n=1}^{\infty} f_n$ is uniformly convergent. □

6.1 Exercises

1. Let $f_n : \mathbb{R} \to \mathbb{R}$ be a family of functions that converge pointwise to some function $f : \mathbb{R} \to \mathbb{R}$. Prove that if $\{f_n\}$ is uniformly Cauchy, then $f_n \to f$ uniformly.
2. Let $f_n(x) := \frac{x^n}{n+x}$ for $x \in \mathbb{R}$ and $n \in \mathbb{N}$.
   (a) Find $\lim_{n\to\infty} f_n(x)$ (in the pointwise sense).
   (b) Does $f_n \to f$ uniformly on $[0, 1]$? (Prove or provide a counter example).
   (c) Does $f_n \to f$ uniformly on $[0, \infty)$? (Prove or provide a counter example).
3. Let $\{f_n\}$ and $\{g_n\}$ be sequences of functions that respectively converge uniformly to functions f and g.
   (a) Prove that the sequence $\{f_n + g_n\}$ is uniformly convergent.
   (b) Prove that if there exists $M > 0$ such that $|f_n(x)| \leq M$ and $|g_n(x)| \leq M$ for all x and n (we say that f and g are uniformly bounded), then the sequence $\{f_n g_n\}$ converges uniformly.
   (c) Provide an example that shows that the product $\{f_n g_n\}$ may not converge uniformly if the assumption of uniform boundedness is dropped.
4. Let $f : \mathbb{R} \to \mathbb{R}$ be uniformly continuous. Define a sequence $f_n(x) := f(x + 1/n)$.
   (a) Prove that $f_n \to f$ uniformly.
   (b) Provide an example that shows that the hypothesis on uniform continuity can not be relaxed into simple continuity.
5. The Differentiable limit theorem remains true with a slightly weaker set of hypotheses than the ones that we used to prove it.

   Let $f_n : A \subset \mathbb{R} \to \mathbb{R}$ be a sequence of differentiable functions converging pointwise to some function $f : A \subset \mathbb{R} \to \mathbb{R}$ **that is continuous at one point** $x_0 \in A$. Moreover, assume that the sequence of derivatives f_n' converges uniformly to some function $g : A \subset \mathbb{R} \to \mathbb{R}$. Prove that the pointwise limit f is differentiable and $f'(x) = g(x)$. [Hint follow the proof of the Differentiable limit theorem in the text and try to identify where and how is continuity used. Then modify it to accomodate for continuity at one point.]
6. Prove that a sequence of functions $f_n(x) : A \subset \mathbb{R} \to \mathbb{R}$ converges uniformly to a function $f : A \to \mathbb{R}$ if and only if
$$\lim_{n\to\infty} \sup\{|f(x) - f_n(x)| : x \in A\} = 0.$$
7. Let $f_n : [a, b] \to \mathbb{R}$ be a sequence of continuous functions.
   (a) Suppose that, for each $x \in [a, b]$, the sequence of real numbers $\{f_n(x)\}$ is decreasing. Prove that if $f_n \to 0$ pointwise on $[a, b]$, then $f_n \to 0$ uniformly on $[a, b]$. [Hint: If not, there exists $\epsilon > 0$ such that for every n there is a point x_n in $[a, b]$ and a function f_n in the sequence such that $f_n(x_n) \geq \epsilon$. (This is simply the negation of uniform convergence) Use this to obtain a contradiction.]
   (b) Suppose that, for each $x \in [a, b]$, the sequence of real numbers $\{f_n(x)\}$ is increasing. Prove that if $f_n \to f$ pointwise on $[a, b]$ and if f is continuous on $[a, b]$, then $f_n \to f$ uniformly on $[a, b]$. (This result is known as *Dini's theorem*).

Bibliography

[1] S. Abbott. *Understanding Analysis.* Springer New York, 2015.

[2] R. Bartle and D. Sherbert. *Introduction to Real Analysis.* Wiley, 1982.

[3] E. Landau. *Foundations of Analysis.* AMS Chelsea Publishing Series. American Mathematical Society, 2001.

[4] F. Morgan. *Real Analysis.* American Mathematical Society, 2005.

[5] C. C. Pugh. *Real Mathematical Analysis.* Springer New York, 2002.

[6] K. Ross. *Elementary Analysis.* Springer, 2014.

[7] T. Sánchez-Vizuet. Minimal analysis II: A concise introduction to real analysis of multiple variables, 2025.

[8] M. Spivak. *Calculus.* Publish or Perish, fourth edition, 2008.

Alphabetical Index

R

S

T

U

W