MATH 101 (en)– Analysis I (English)
Notes for the course given in Fall 2021

Teacher:Roberto Svaldi
Head Assistant: Stefano Filipazzi

Notes written by Zsolt Patakfalvi & Roberto Svaldi

Thursday 30th November, 2023

This work is licensed under a Creative Commons "Attribution-NonCommercial-NoDerivatives 4.0 International" license.

Contents

1 Real functions of one variable

In this section, we are going to consider functions $f \colon E \to \mathbb{R}$ where E is a subset of $\mathbb{R}$ and study their properties. We first start by recalling general basic properties of functions.

1.1 Limits of functions and continuity

In this section we will define and discuss the notion of limit of a function at a given point. The notion of limit aims to give a mathematically precise measure of what the local behavior of a function is around a given point.

Example 1.1. The starting point of our investigation is the function $f(x) := \frac{\sin(x)}{x}$ near $x = 0$. At first sight, $f(0)$ would seem not to be defined, as x appears in the denominator of $\frac{\sin(x)}{x}$. On the other hand, looking at the graph of the function in Figure 1, it would appear that the

Images/sinx_sux_gr.png

Figure 1: $f(x) = \frac{\sin(x)}{x}$.

closer x is to 0, the closer $\frac{\sin(x)}{x}$ is to 1.
We can be even more precise if, for example, we consider the sequence (y_n), $y_n := \frac{1}{n}$, $n \in \mathbb{N}$, then $\lim\limits_{n\to\infty} y_n = 0$ and we can actually show that also the limit $\lim\limits_{n\to\infty} f(y_n)$ exists. Indeed, you proved in the exercise sheets that

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

So, even though $\frac{\sin(x)}{x}$ is not defined at $x = 0$, if we set

$$f(x) = \frac{\sin(x)}{x}, \text{ for } x \neq 0, \quad f(0) = 1,$$

then it would appear that $f(x)$ becomes a nice "continuous" function around at $x = 0$, meaning that we could draw the graph of f with just one continuous stroke of the pen.

The goal of this section is for us to turn the ideas contained in the previous example into some precise mathematical concepts and definitions and derive further consequences starting from those. In particular, we will define precisely why, in the previous example, $f(0) = 1$ makes $f(x)$ "continuous".

First, we need to make of how to unsure that a function f is defined around a point $x_0 \in \mathbb{R}$, so that it makes sense for us to talk about "the behavior of f around x_0".

Definition 1.2. A function $f\colon E \to \mathbb{R}$ is defined on *a punctured neighborhood* of $x_0 \in \mathbb{R}$ if for some positive real number $\delta \in \mathbb{R}^*_+$, E contains a set of the form $(x_0 - \delta, x_0 + \delta) \setminus \{x_0\}$.

Remark 1.3. Equivalently, we can restate the above definition in the following way:
A function $f\colon E \to \mathbb{R}$ is defined on *a punctured neighborhood* of $x_0 \in \mathbb{R}$ if for some positive real number $\delta \in \mathbb{R}^*_+$, an interval of the form $(x_0 - \delta, x_0 + \delta)$ is contained in $E \cup \{x_0\}$.

Example 1.4. The function $f(x) := \frac{\sin(x)}{x}$ is defined on any pointed neighborhood of 0. Indeed, f is defined on $E := \mathbb{R} \setminus \{0\}$ so that

$$(-\delta, +\delta) \setminus \{0\} \subseteq E, \quad \forall \delta \in \mathbb{R}^*_+.$$

We are then ready to give the formal definition of limit.

Definition 1.5. Let $f\colon E \to \mathbb{R}$ and $l \in \mathbb{R}$. Assume that E contains a punctured neighborhood of $x_0 \in \mathbb{R}$. Then, $\lim_{x \to x_0} f(x) = l$ if one of the following two equivalent conditions holds:

(1) For every $0 < \varepsilon \in \mathbb{R}$ there exists $\delta_\varepsilon \in \mathbb{R}^*_+$ such that

$$\forall x \in E, 0 < |x - x_0| < \delta_\varepsilon \Rightarrow |f(x) - l| < \varepsilon.$$

(2) For every sequence $(x_n) \subseteq E \setminus \{x_0\}$ for which $\lim_{n\to\infty} x_n = x_0$, we have $\lim_{n\to\infty} f(x_n) = l$.

Remark 1.6. Roughly speaking the two definitions mean the following:

(1) whenever f is defined at x and x is close to x_0, then $f(x)$ is close to l. More precisely: for every $\varepsilon > 0$ there is a $\delta > 0$ such that if x is closer to x_0 than δ (and f is defined at x), then $f(x)$ is closer to l than ε.

(2) whenever a sequence (y_n) is contained in $E \setminus \{x_0\}$ and it converges to x_0, then the sequence $(f(y_n))$, given by the values of the function f along the x_n, converges to l.

Remark 1.7. We explain why the two conditions in Definition 1.5 are equivalent.

(1) $\Rightarrow$ (2): Let us fix a sequence (y_n) which is contained in $E \setminus \{x_0\}$ and for which $\lim_{n\to\infty} x_n = x_0$. We have to show that $\lim_{n\to\infty} f(x_n) = l$. Let us fix $\varepsilon > 0$. Then, this yields a $\delta > 0$ as in definition (i). For this δ, there is an n_δ such that $|x_0 - x_n| < \delta$ for $n \geq n_\delta$, and hence for all such n, $|l - f(x_n)| < \varepsilon$.

NOT (1) $\Rightarrow$ NOT (2): The negation of (i) is that there is an $\varepsilon > 0$ such that for each $\delta > 0$ we can find $y_\delta \in (x_0 - \delta, x_0 + \delta[\setminus\{x_0\}$ such that $|f(y_\delta) - l| \geq \varepsilon$. Defining $y_n := y_{\frac{1}{n}}$, then the sequence (y_n) converges to x_0, but all values $f(y_n)$ have distance at least ε from l, so the sequence $(f(x_n))$ cannot converge to l.

We will work more often with the use definition (2) more as it is simpler. Luckily, it is almost always enough for proving that a limit does not exist. We usually use definition (1) only when (2) does not work.

Example 1.8. We show that $\lim_{x\to 2} x^2 = 4$ using point (1) of Definition 1.5. In order to do this, we proceed as follows: let us fix $\varepsilon > 0$; at this point, we need to find $\delta > 0$ such that

$$\text{if } 0 < |x - 2| < \delta \text{ then, } |x^2 - 4| < \varepsilon.$$

Let us note that

$$x^2 - 4 = (x - 2)(x + 2).$$

Furthermore, if $0 < |x-2| < 1$, then $3 < x+2 < 3$; thus, if $0 < |x-2| < 1$, then

$$|x^2-4| = |x-2||x+2| < 5|x-2|. \tag{1.8.a}$$

Therefore, taking $\delta = \min\left\{1, \frac{\varepsilon}{5}\right\}$, we conclude that, if $0 < |x-2| < \delta$ then

$$|x-2| < 1 \quad \text{and} \quad |x-2| < \frac{\varepsilon}{5}.$$

Hence, it follows that

$$\underbrace{|x^2-4| \quad < \quad 5|x-2|}_{\text{Using (1.8.a), since } |x-2|<1} \quad < \quad 5\frac{\varepsilon}{5} = \varepsilon.$$

Example 1.9. We can repeat the same computation as in the previous example, also using (2) of Definition 1.5. In general, proving the existence (and finiteness of the limit using sequence) can be rather tricky: you can try for example to use that definition to compute the limit of $\frac{\sin(x)}{x}$ at $x = 0$. Let (x_n) be a sequence such that $\lim_{n\to\infty} x_n = 2$. Then, by the algebraic properties of the limit, that is, by **??**, we know that $\lim_{n\to\infty} x_n^2 = \left(\lim_{n\to\infty} x_n\right)^2 = 2^2 = 4$.

Example 1.10. We can generalize the arguments from the previous examples to show that for any $x_0 \in \mathbb{R}$, $\lim_{x\to x_0} x^n = x_0^n$ for any $n \in \mathbb{N}$.

Using the notion of limit, we can also define the notion of continuity of a function f at a point in $D(f)$.

Definition 1.11. Let $f\colon E \to \mathbb{R}$ be a function, $E \subset \mathbb{R}$. The function f is *continuous* at a point $x_0 \in E$, if $\lim_{x\to x_0} f(x)$ exists, it is finite and $\lim_{x\to x_0} f(x) = f(x_0)$.

Remark 1.12. Implicit in Definition 1.11 is the fact that the limit of $f(x)$ at x_0 exists. In particular, the subset $E \subseteq \mathbb{R}$ on which f is assumed to be defined must contain not only x_0 but also a punctured neighborhood of x_0.

Remark 1.13. The main difference between Definition 1.5 and Definition 1.11 is the fact that, while in Definition 1.5 we do not require the function f to be defined at the point x_0 at which we are trying to compute the limit and when taking the limit we only look at the value of f on points of $D(f) \setminus \{x_0\}$, in the case of Definition 1.11, instead, we very much want to allow the value $f(x_0)$ to play a role. More precisely, we have the following characterization of continuity at a point, via conditions analogous to those in Definition 1.5.

Proposition 1.14. *Let $f\colon E \to \mathbb{R}$ be a function and let $x_0 \in E$. Assume that there exists a open interval of the form $(x_0 - \delta, x_0 + \delta[$, $\delta > 0$ contained in E. Then, f is continuous at x_0 if and only if one of the following two equivalent definitions hold:*

(1) For every $\varepsilon \in \mathbb{R}_+^*$ there is a $\delta_\varepsilon \in \mathbb{R}_+^*$ such that

$$\forall x \in E \text{ such that } |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

(2) For every sequence $(x_n) \subseteq E$ for which $\lim_{n\to\infty} x_n = x_0$, then $\lim_{n\to\infty} f(x_n) = f(x_0)$.

In view of the above proposition, Example 1.10 yields the following immediate corollary.

Corollary 1.15. *Fix $n \in \mathbb{N}$. Then $f(x) = x^n$ is continuous at every $x_0 \in \mathbb{R}$.*

Using the next definition, we can rephrase the previous corollary by saying that, for a fixed $n \in \mathbb{N}$, the function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous.

Definition 1.16. Let $f\colon E \to \mathbb{R}$. Assume that $\forall x_0 \in E$, E contains an open ball centered at x_0. Then we say that f *is continuous* if it is continuous at every $x_0 \in E$.

Example 1.17. (1) Let us define the function $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(x) := \begin{cases} 0 & x \neq 0, \\ 1 & x = 0. \end{cases}$$

Then, f is not continuous at $x = 0$.
In fact, $\lim_{x\to 0} f(x) = 0$, as in the definition we assumed $0 < |x - x_0| \leq \delta$, so the function value 1 for $x_0 = 0$ does not cause any problem.

(2) Let us define the function $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(x) := \begin{cases} 0 & x \in \mathbb{Q}, \\ 1 & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Then, the set of points of $\mathbb{R}$ at which f is continuous is empty.
For example, let us consider the point $0 \in \mathbb{R}$ and let us define the sequences $(y'_n), (y''_n) \subset \mathbb{R} \setminus \{0\}$, to be

$$y'_n = \frac{1}{n}, \quad y''_n = \frac{1}{\sqrt{n^2+1}}, \quad n \geq 1.$$

As $\forall n \in \mathbb{N}^*$, $\frac{1}{n} \in \mathbb{Q}$, $\frac{1}{\sqrt{n^2+1}} \in \mathbb{R} \setminus \mathbb{Q}$, then $\forall n \in \mathbb{N}^*$, $f(y'_n) = 0$, while $f(y''_n) = 1$.
Hence, $\lim_{n\to\infty} y'_n = 0 = \lim_{n\to\infty} y''_n$, while $\lim_{n\to\infty} f(y'_n) = 0$, $\lim_{n\to\infty} f(y''_n) = 1$, hence the limit $\lim_{x\to 0} f(x)$ does not exist, and moreover, f is not continuous at 0.
One can repeat the same reasoning at any point $x_0 \in \mathbb{R}$, by taking y'_n to be a sequence of rational numbers converging to x_0 (for example, $y'_n = x_0 + \frac{1}{n}$, if x_0 is rational, or y'_n to be the truncation of the decimal form of x_0 at the n-th decimal digit, if x_0 is irrational) and y''_n to be a sequence of irrational numbers converging to x_0 (for example, $y''_n = x_0 + \frac{1}{\sqrt{n^2+1}}$, if x_0 is rational, or $y''_n = x_0 + \frac{1}{n}$ if x_0 is irrational). Then, $\forall n \in \mathbb{N}^*$, $f(y'_n) = 0$, while $f(y''_n) = 1$, and $\lim_{n\to\infty} y'_n = x_0 = \lim_{n\to\infty} y''_n$, while $\lim_{n\to\infty} f(y'_n) = 0$, $\lim_{n\to\infty} f(y''_n) = 1$, hence the limit $\lim_{x\to x_0} f(x)$ does not exist, and moreover, f is not continuous at x_0.
This implies that $\lim_{x\to x_0} f(x)$ does not exist at any point $x_0 \in \mathbb{R}$, in particular, f is not continuous at any point of $\mathbb{R}$.

Example 1.18. *We claim that* $\lim_{x\to 0} \cos(x) = 1$.
Indeed, let (x_n) be a sequence converging to 0. Then,

$$0 \leq |\cos(x_n) - 1| = \left|2\sin^2\left(\frac{x_n}{2}\right)\right| \leq 2\frac{x_n^2}{4} = \frac{x_n^2}{2},$$

using the inequality $|\sin(x)| \leq |x|$. So, squeeze theorem tells us that $\lim_{n\to\infty} |\cos(x_n) - 1| = 0$.

Example 1.19. *The limit* $\lim_{x\to 0} \sin\left(\frac{1}{x}\right)$ *does not exist.*
Indeed, consider the sequences $x_n := \frac{1}{\pi\left(2n+\frac{1}{2}\right)}$ and $x'_n := \frac{1}{\pi\left(2n+\frac{3}{2}\right)}$. Then, first $\lim_{n\to\infty} x_n = 0$ and $\lim_{n\to\infty} x'_n = 0$. However,

$$\lim_{n\to\infty} \sin\left(\frac{1}{x_n}\right) = \sin\left(\frac{1}{\frac{1}{\pi\left(2n+\frac{1}{2}\right)}}\right) = \sin\left(\pi\left(2n + \frac{1}{2}\right)\right) = 1,$$

but

$$\lim_{n\to\infty} \sin\left(\frac{1}{x'_n}\right) = \sin\left(\frac{1}{\frac{1}{\pi\left(2n+\frac{3}{2}\right)}}\right) = \sin\left(\pi\left(2n+\frac{3}{2}\right)\right) = -1.$$

So, point (2) of Definition 1.5 is not satisfied, and hence the limit does not exist

1.1.1 Limits and algebra

Definition 1.5 allows us to translate all the statements about limits of sequences to limits of functions. Indeed, let us say we are have functions f, g defined around a point $x_0 \in \mathbb{R}$ – but we are not necessarily assuming that f, g are defined at x_0 – and we want to prove that if l and k are the limits of $f(x)$ and $g(x)$ (at x_0), then $l + k$ is the limit of $(f + g)(x)$. Let us take a sequence (y_n) converging to x_0. We know that $\lim_{n\to\infty} f(y_n) = l$ and $\lim_{n\to\infty} g(y_n) = k$. But then, **??** implies that $\lim_{n\to\infty} f(x_n) + f'(x_n) = l + k$, that is, $\lim_{n\to\infty}(f + f')(x_n) = l + k$, which is exactly generalizing the statement about limit of sequences and addition, to the case of limit of functions.

Unsurprisingly, at this point, we can do the same with all other properties that we proved for limits of sequences. We collect all the statements one can show along the same arguments:

Proposition 1.20. *Let f and g be two functions such that a punctured neighborhood of x_0 is in the domain of both f and g. Assume that the limits of f and g at x_0 exist and they are l and k, respectively. Then,*

(1) the limit of $f + g$ exists at x_0 and $\lim_{x\to x_0}(f + g)(x) = l + k$

(2) the limit of $f \cdot g$ exists at x_0 and $\lim_{x\to x_0}(f \cdot g)(x) = l \cdot k$

(3) if $k \neq 0$, then the limit of $\frac{f}{g}$ exists at x_0 and $\lim_{x\to x_0}\left(\frac{f}{g}\right)(x) = \frac{l}{k}$

(4) if $f(x) \leq g(x)$ for any x in a punctured neighborhood of x_0, then $l \leq k$.

(5) **Squeeze Theorem***: if there is a third function $h(x)$ such that there is also a punctured neighborhood of x_0 in the domain of h, and:*

(i) on some punctured neighborhood of x_0 we have $f(x) \leq h(x) \leq g(x)$, and

(ii) $l = k$,

then $\lim_{x\to x_0} h(x) = l$.

Example 1.21. The main example for using point (5) of Proposition 1.20 is that $\lim_{x\to 0}\frac{\sin(x)}{x} = 1$. Indeed, we have already seen, cf. **??**, that

$$0 \leq \sin(x) \leq x \leq \tan(x), \quad \text{for } x \in [0, \frac{\pi}{2}],$$
$$\tan(x) \leq x \leq \sin(x) \leq 0, \quad \text{for } x \in [-\frac{\pi}{2}, 0];$$

which implies that

$$|\cos(x)| \leq \left|\frac{\sin(x)}{x}\right| \leq 1, \quad \text{for } x \in [-\frac{\pi}{2}, \frac{\pi}{2}]. \tag{1.21.b}$$

Since $\forall x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, $\cos(x) \geq 0$, whereas $\sin(x)$ and x are odd function, so that also $\forall x \in [-\frac{\pi}{2}, \frac{\pi}{2}] \setminus \{0\}$, $\frac{\sin(x)}{x} \geq 0$, the chain of inequalities in (1.21.b) holds also once we remove the absolute values. So, by the Squeeze Theorem for limits we can conclude that

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

since $\lim_{x \to 0} \cos(x) = 1$, see Example 1.18.

Example 1.22. *For any $k \in \mathbb{N}^*$, $\lim_{x \to 0} x^k \sin\left(\frac{1}{x}\right) = 0$.* Indeed, as $|\sin\left(\frac{1}{x}\right)| = 1$, $\forall x \in \mathbb{R}^*$, then

$$-x^k \leq x^k \sin\left(\frac{1}{x}\right) \leq x^k, \quad \forall x \in \mathbb{R}^*.$$

and the conclusion follows from the Squeeze Theorem.

The above proposition has all the nice consequences about continuity.

Proposition 1.23. *If $f, g \colon E \to \mathbb{R}$ are continuous functions at $x_0 \in E$. Then the following function are also continuous at x_0:*

(1) $\alpha f + \beta g$ for any $\alpha, \beta \in \mathbb{R}$,

(2) $f \cdot g$, and

(3) $\frac{f}{g}$, if $g|_E$ is nowhere zero (meaning that for all $x \in E : g(x) \neq 0$), then.

Example 1.24. We collect here some example of continuous functions, on their respective domains, that is, each of these functions is continuous at any point where they are defined:

- $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_r x^r$, that is, $p(x)$ is a polynomial in one variable x.
- $f(x) := \frac{1}{x}$ is continuous on $\mathbb{R} \setminus 0$.
- $\frac{x}{x^2 - 3x + 1}$ is continuous on $\mathbb{R} \setminus \left\{\frac{3 \pm \sqrt{5}}{2}\right\}$,
- In general, if $p(x)$ and $q(x)$ are two polynomials, then $\frac{p(x)}{q(x)}$ is continuous on $\{x \in \mathbb{R} | q(x) \neq 0\}$ (which is the whole real line minus finitely many points).

1.1.2 Limit and composition

Let us recall the following definition of composition of functions.

Definition 1.25. If $f \colon E \to \mathbb{R}$ and $g : G \to \mathbb{R}$ are functions such that $\mathrm{R}(f) \subseteq G$ then we may define the *composition* $g \circ f$ (order matters!!) of f with g by

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

Example 1.26. Let us take $f \colon \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2 + 1$, and $g : \mathbb{R} \to \mathbb{R}$ given by $g(y) = y^3 + y + 1$. Then we have

$$(g \circ f)(x) = (x^2 + 1)^3 + (x^2 + 1) + 1 = x^6 + 3x^4 + 4x^2 + 3. \tag{1.26.c}$$

Let us look at an example about whether composition of continuous functions is continuous or not.

Example 1.27. Consider the functions defined in Example 1.26. *We want to show that $g \circ f$ is continuous at $x = 0$.* As $(g \circ f)(0) = 3$ by (1.26.c), we then need to show that $\lim_{x \to 0} (g \circ f)(x) = 3$. and this can be immediately deduced from the last part of the equation in (1.26.c). So, indeed $g \circ f$ is continuous at $x = 0$.

In general the situation is just as nice as in Example 1.27.

Proposition 1.28. *Let $f : E \to \mathbb{R}$ and $g : G \to \mathbb{R}$ be two functions. Assume that:*

(1) $f(E) \subseteq G$,

(2) f is continuous at x_0,

(3) g is continuous at $y_0 := f(x_0)$.

Then $g \circ f$ is continuous at x_0.

Proof. We verify condition (2) of Proposition 1.14. Let $(z_n) \subseteq E$ be a sequence such that

$$\lim_{n \to \infty} z_n = x_0. \tag{1.28.d}$$

According to (1.28.d) and our assumption (2), then

$$\lim_{n \to \infty} f(z_n) = y_0. \tag{1.28.e}$$

Hence

$$\underbrace{\lim_{n \to \infty} (g \circ f)(z_n) = \overbrace{\lim_{n \to \infty} g(f(z_n))}_{\text{Definition 1.25}} = g(y_0)}^{\text{(1.28.e) and condition (3)}}.$$

□

Remark 1.29. Let us examine a bit further Example 1.27. In the proof of Proposition 1.28 we showed that if

$$\lim_{x \to x_0} f(x) = y_0 \qquad \text{and} \qquad \lim_{y \to y_0} g(y) = l, \tag{1.29.f}$$

then $\lim_{x \to x_0} g \circ f(x) = l$ holds under the assumption that f and g are continuous. We may be tempted to think that an analogous statement to Proposition 1.28 should also for the limit of a composition of functions, just assuming the condition in (1.29.f). However, as we will see in Example 1.31, this is not true. The reason is that in Definition 1.5, contrary to Proposition 1.14, there is nothing said about the behavior at x_0 and y_0. So, we have to assume that $f(x)$ avoids y_0 in a punctured neighborhood of x_0.

The precise statement about composition of functions, in regards to limits, is as follows.

Proposition 1.30. *Let $f : E \to \mathbb{R}$ and $g : G \to \mathbb{R}$ be functions and let $x_0 \in E$ be a point such that*

(1) $f(E) \subseteq G$,

(2) $\lim_{x \to x_0} f(x_0) = y_0$,

(3) $\lim_{y \to y_0} g(y_0) = l$

(4) there is a punctured neighborhood $(x_0 - \delta, x_0 + \delta) \setminus \{x_0\} \subseteq E$ such that for every x in this neighborhood, $f(x) \neq y_0$.

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

Proof. We use part (2) of Definition 1.5. Thus, let us fix a sequence $(z_n) \subseteq E\setminus\{x_0\}$ such that

$$\lim_{n\to\infty} z_n = x_0. \tag{1.30.g}$$

In particular, by throwing away finitely many elements of the sequence, we may assume that

$$(z_n) \subseteq (x_0-\delta, x_0+\delta)\setminus\{x_0\} \subseteq E. \tag{1.30.h}$$

By the assumption (2) in the statement of the proposition, and by (1.30.g), it follows that

$$\lim_{n\to\infty} f(z_n) = y_0. \tag{1.30.i}$$

Lastly, by our assumption (4) and (1.30.h) we have

$$(f(z_n)) \subseteq G\setminus\{y_0\}. \tag{1.30.j}$$

Hence, by our assumption (3) and by (1.30.j), we have

$$\lim_{n\to\infty} (g\circ f)(x_n) = \lim_{n\to\infty} g(f(x_n)) = l.$$

$\square$

The following example shows that condition (4) of Proposition 1.30 is necessary. That is, if we drop condition (4), the statement of Proposition 1.30 would not hold.

Example 1.31. Consider:

$$g(x) = \begin{cases} 0, & \text{for } x\neq 0, \\ 1, & \text{for } x = 0, \end{cases} \qquad \text{and} \qquad f(x) = \begin{cases} x\sin\left(\frac{1}{x}\right), & \text{for } x\neq 0 \\ 0, & \text{for } x = 0. \end{cases}$$

Then, $\lim_{x\to 0} f(x) = 0$ and $\lim_{x\to 0} g(x) = 0$. However, $\lim_{x\to 0}(g\circ f)(x) \neq 0$, because the following two sequences induce function value sequences with different limits:

$$x_n := \frac{1}{\pi n} \qquad \text{and} \qquad y_n := \frac{1}{\pi n + \frac{\pi}{2}},$$

as

$$\lim_{n\to\infty} (g\circ f)(x_n) = \lim_{n\to\infty} 1 = 1 \qquad \text{and} \qquad \lim_{n\to\infty} (g\circ f)(y_n) = \lim_{n\to\infty} 0 = 0.$$

Also, let us note that condition (4) of Proposition 1.30 below is not satisfied in this example, as $f(x) = 0$ for $x = \frac{1}{\pi n}$, so there is no punctured neighborhood of 0 on which the function of f avoids the value 0.

Example 1.32. A positive example for applying Proposition 1.30 is during the argument of showing that $\lim_{x\to 0} \dfrac{\sin(x^2)}{x^2} = 1$. Indeed, if we set $g(x) := \frac{\sin(x)}{x}$, and $f(x) = x^2$, then condition (4) of Proposition 1.30 is also satisfied, as $f(x)\neq 0$ for $x\neq 0$.

1.1.3 Infinite limits

Definition 1.33. A *neighborhood* of $+\infty$ (resp. $-\infty$) is an unbounded interval of the form $(a, +\infty)$ (resp. $(-\infty, a)$).

We extend the definition of limit to comprise the case where we allow ourselves to work with the extended real line $\overline{\mathbb{R}}$.

Definition 1.34. Let $x_0, l \in \overline{\mathbb{R}}$, and let $f\colon E \to \mathbb{R}$ be a function, $E \subset \mathbb{R}$. Assume that E contains a punctured neighborhood of x_0. We say that the limit of $f(x)$ at x_0 is l, if for any sequence $(y_n) \subseteq E \setminus \{x_0\}$[1], whenever $\lim_{n\to\infty} y_n = x_0$, then $\lim_{n\to\infty} f(y_n) = l$.

Example 1.35. *We show that* $\lim_{x\to 0} \frac{1}{x^2} = +\infty$. Indeed, if $(x_n) \subset \mathbb{R}^*$ is a sequence satisfying $\lim_{n\to\infty} x_n = 0$, then $\lim_{n\to\infty} \frac{1}{x_n^2} = +\infty$ by algebraic properties of limits of sequences.
On the other hand, the limit $\lim_{x\to 0} \frac{1}{x}$ does not exist. In fact, considering the sequence $x_n := \frac{1}{n}$, then $\lim_{n\to\infty} \frac{1}{x_n} = \lim_{n\to\infty} n = +\infty$, while for $y_n = \frac{-1}{n}$, then $\lim_{n\to\infty} \frac{1}{y_n} = \lim_{n\to\infty} -n = -\infty$.

Proposition 1.36. *Let $x_0 \in \overline{\mathbb{R}}$, and let $f, g\colon E \to \mathbb{R}$ be functions.*

(1) Addition rule. *Assume that the following conditions are satisfied:*

- $\lim_{x\to x_0} f(x) = +\infty$ *(resp. $-\infty$), and*
- $g(x)$ *is bounded from below (resp. from above)*

then $\lim_{x\to x_0} (f+g)(x) = +\infty$ *(resp. $-\infty$).*

(2) Product rule. *Assume that the following conditions are satisfied:*

- $\lim_{x\to x_0} |f(x)| = +\infty$,
- *there exists $\delta > 0$ such that $|g(x)| \geq \delta$ for all $x \in E$, and*
- $f(x)g(x) > 0$ *(resp. < 0) for all $x \in E$,*

then $\lim_{x\to x_0} f(x)g(x) = +\infty$ *(resp. $-\infty$).*

(3) First division rule. *If*

- $f(x)$ *is bounded,*
- $g(x)$ *is nowhere zero, and*
- $\lim_{x\to x_0} |g(x)| = +\infty$.

Then $\lim_{x\to x_0} \frac{f(x)}{g(x)} = 0$

(4) Second division rule. *If*

- $\lim_{x\to x_0} g(x) = 0$,
- *there is a $\delta > 0$ such that $|f(x)| \geq \delta$ for all $x \in E$, and*
- $f(x)/g(x) > 0$ *(resp. < 0) for all $x \in E$,*

then $\lim_{x\to x_0} \frac{f(x)}{g(x)} = +\infty$ *(resp. $-\infty$).*

[1]When $x_0 = \pm\infty$, then the condition that x_0 does not belong to E is automatically satisfied, since $E \subseteq \mathbb{R}$.

(5) Squeeze. *If* $f(x) \le g(x)$, *and*

- *if* $\lim_{x\to x_0} f(x) = +\infty$, *then* $\lim_{x\to x_0} g(x) = +\infty$
- *if* $\lim_{x\to x_0} g(x) = -\infty$, *then* $\lim_{x\to x_0} f(x) = -\infty$

Example 1.37. Here are a few examples.

- $\lim_{x\to 0} \underbrace{\frac{1}{x^2}}_{\to\infty} + \underbrace{\cos(x)}_{\text{bounded}} = +\infty$
- $\lim_{x\to +\infty} \underbrace{\cos(x)}_{\text{bounded}} \cdot \underbrace{(-x^2 + x^3)}_{\to+\infty} = -\infty$, since $-x^2 + x^3 = 0$ only for $x = 0, 1$.
- $\lim_{x\to +\infty} \frac{\overbrace{\arctan(x)}^{\text{bounded}}}{\underbrace{(-x)}_{\to-\infty}} = 0.$

Remark 1.38. (1) The assumptions stated in part (1) of Proposition 1.36 are important, as otherwise we can have all different kinds of limits. We give examples of this using the functions $f(x) = x^3$, $g(x) = x^2$ and $h(x) = x^3 + 1$. We have

- $\lim_{x\to+\infty} f(x) = \lim_{x\to+\infty} x^3 = +\infty$, $\lim_{x\to+\infty} - f(x) = \lim_{x\to+\infty} - x^3 = -\infty$
- $\lim_{x\to+\infty} h(x) = \lim_{x\to+\infty} (x^3 + 1) = +\infty$, $\lim_{x\to+\infty} - h(x) = \lim_{x\to+\infty} - (x^3 + 1) = -\infty$, and
- $\lim_{x\to+\infty} g(x) = \lim_{x\to+\infty} x^2 = +\infty$, $\lim_{x\to+\infty} - g(x) = \lim_{x\to+\infty} - x^2 = -\infty$.

On the other hand:

- $\lim_{x\to+\infty} f(x) - g(x) = \lim_{x\to+\infty} x^3 - x^2 = \lim_{x\to+\infty} x^2(x - 1) = +\infty$,
- $\lim_{x\to+\infty} g(x) - f(x) = -\infty$, and
- $\lim_{x\to+\infty} f(x) - h(x) = -1$.

In particular, never use addition law for limits of the type $(+\infty) + (-\infty)$.

(2) The above assumptions for point (2) of Proposition 1.36 are also important. We give examples of this using the functions $f(x) = x$, $g(x) = \frac{\cos(x)}{x}$ and $h(x) = (-1)^{[x]}$. We have

- $\lim_{x\to+\infty} |f(x)| = +\infty$,
- $|g(x)|$ is not bounded from below, and
- $|h(x)|$ is bounded from below, but $f(x)h(x) \not> 0$.

Then:

- $\lim_{x\to+\infty} f(x)g(x) = \lim_{x\to+\infty} \cos(x)$ does not exist, and
- $\lim_{x\to+\infty} f(x)h(x) = \lim_{x\to+\infty} x(-1)^{[x]}$ does not exist, on the other hand
- the product law applies to $(f(x)h(x))(h(x))$ and yields $\lim_{x\to+\infty} (f(x)h(x))(h(x)) = +\infty$

Never try to use product rule to limits of the type $0 \cdot \infty$.

(3) The assumptions of the first division rule are also important. One can can show that in the $\frac{\pm\infty}{\pm\infty}$ case anything can happen for example using $\frac{1}{x}$, $\frac{1}{x^2}$, $\frac{1}{x^3}$,$(-1)^{\left[\frac{1}{x}\right]}\frac{1}{x}$ with limit at 0:

- $\lim\limits_{x\to 0}\frac{\frac{1}{x}}{\frac{1}{x^2}} = \lim\limits_{x\to 0} x = 0$,
- $\lim\limits_{x\to 0}\frac{(-1)^{\left[\frac{1}{x}\right]}\frac{1}{x}}{\frac{1}{x}} = \lim\limits_{x\to 0}(-1)^{\left[\frac{1}{x}\right]}$ does not exist and bounded,
- $\lim\limits_{x\to 0}\frac{\frac{1}{x^2}}{\frac{1}{x}} = \lim\limits_{x\to 0}\frac{1}{x}$ does not exist and unbounded, and
- $\lim\limits_{x\to 0}\frac{\frac{1}{x^3}}{\frac{1}{x}} = \lim\limits_{x\to 0}\frac{1}{x^2} = +\infty$.

Similar examples show that the assumptions are important for the second division rule. Never try to use division rules to limits of the form $\frac{\pm\infty}{\pm\infty}$ and $\frac{0}{0}$.

1.1.4 One sided limits

The main question is how to make sense of limits such as at 0 of $\sqrt{x}$, as here the domain does not contain a punctured neighborhood of 0. The solution for this is the introduction of the notions of left and right limits.

Definition 1.39. A function $f : E \to \mathbb{R}$ is defined to the left (resp. to the right) of $x_0 \in \mathbb{R}$, if E contains an interval of the form $(x_0 - \delta, x_0[$ (resp. $(x_0, x_0 + \delta[)$.

Definition 1.40. Let $f\colon E \to \mathbb{R}$ be a function and $x_0 \in \mathbb{R}$. Assume that f is defined to the left (resp. right) of x_0. Let $l \in \overline{\mathbb{R}}$.

(1) We say that the limit of f for x that goes to x_0 from the *left* is l if for all sequences $(x_n) \subseteq \{x \in E | x < x_0\}$, whenever $\lim\limits_{n\to\infty} x_n = x_0$ then $\lim\limits_{n\to\infty} f(x_n) = l$. When this condition is satisfied, we write $\lim\limits_{x\to x_0^-} f(x) = l$.

(2) We say that the limit of f for x that goes to x_0 from the *right* is l if for all sequences $(x_n) \subseteq \{x \in E | x > x_0\}$, whenever $\lim\limits_{n\to\infty} x_n = x_0$ then $\lim\limits_{n\to\infty} f(x_n) = l$. When this condition is satisfied, we write $\lim\limits_{x\to x_0^+} f(x) = l$.

Example 1.41. Consider the function $f\colon \mathbb{R}_+ \to \mathbb{R}$ defined as $f(x) := \sqrt{x}$. *We claim that* $\lim\limits_{x\to 0^+}\sqrt{x} = 0$.
Indeed, fix a sequence $(x_n) \subseteq \mathbb{R}_+^*$ such that $\lim\limits_{n\to\infty} x_n = 0$. We have to show that then $\lim\limits_{n\to\infty}\sqrt{x_n} = 0$ too. So, we need to show that for each $\varepsilon > 0$, there is an n_0 such that for every integer $n \geq n_0$, $\sqrt{x_n} \leq \varepsilon$. However, we know that $\lim\limits_{n\to\infty} x_n = 0$. So, we know that there is an n_0 such that $|x_n| < \varepsilon^2$ for all $n \geq n_0$. But then, for any such n we also have $\sqrt{x_n} < \varepsilon$.

Proposition 1.42. *Let $f\colon E \to \mathbb{R}$ be a function such that there is a punctured neighborhood of x_0 contained in E, and both*

$$l_1 := \lim_{x\to x_0^-} f(x) \qquad \textit{and} \qquad l_2 := \lim_{x\to x_0^+} f(x)$$

exists. Then

$$\lim_{x\to x_0} f(x) = l \quad \Longleftrightarrow \quad l_1 = l_2 = l.$$

Example 1.43. Consider the function $f(x) = \{x\} =$. Both left and right limits exist at all points, and furthermore:

$$\lim_{x \to x_0^-} \{x\} = \begin{cases} \{x\} & \text{if } x \notin \mathbb{Z} \\ 1 & \text{if } x \in \mathbb{Z} \end{cases} \quad \text{and} \quad \lim_{x \to x_0^+} \{x\} = \begin{cases} \{x\} & \text{if } x \notin \mathbb{Z} \\ 0 & \text{if } x \in \mathbb{Z} \end{cases}$$

Hence, according to Proposition 1.42,

$$\lim_{x \to x_0} \{x\} \text{ exists } \Leftrightarrow x \notin \mathbb{Z}.$$

Example 1.44. Example 1.41 together with Proposition 1.42 show that the function $f(x) = \sqrt{|x|}$ is continuous in 0. It is not hard to show that f is actually continuous everywhere in $\mathbb{R}$.

1.1.5 Monotone functions

For a monotone function f, the left (resp. the right) limits always exists at any point x_0 at which the function is defined at the left (resp. the right) of x_0.

Proposition 1.45. *Let $f\colon E \to \mathbb{R}$ be a monotone function. Then, at each point $x_0 \in E$:*

(1) if f is defined on the left of x_0, $\lim_{x \to x_0^-} f(x)$ exists,

(2) if f is defined on the right of x_0, $\lim_{x \to x_0^+} f(x)$ exists, and

(3) if f is defined in a neighborhood of $\pm\infty$, then $\lim_{x \to \pm\infty} f(x)$ exists.

Proof. We treat only the increasing case, as the decreasing one follows from that by regarding $-f$ instead of f. Also, we treat only the first case as the others are similar. Set:

$$l := \sup\{f(x) | x \in E, x < x_0\}. \tag{1.45.k}$$

Let

$$(x_n) \subseteq \{x \in E | x < x_0\} \quad \text{such that} \quad \lim_{n \to \infty} x_n = x_0. \tag{1.45.l}$$

We have to show that $\lim_{n \to \infty} f(x_n) = l$. Fix a $\varepsilon > 0$. Then, by the definition of l, there is an $x' \in \{x \in E | x < x_0\}$, such that

$$f(x') > l - \varepsilon. \tag{1.45.m}$$

According to (1.45.l), there is an $n_0 \in \mathbb{N}$ such that for all integers $n \geq n_0$ we have

$$x' < x_n < x_0. \tag{1.45.n}$$

However, then for all integers $n \geq n_0$ we have:

$$l \geq \underbrace{f(x_n)}_{\text{(1.45.k) and (1.45.n)}} \geq \underbrace{f(x')}_{f \text{ is increasing and (1.45.n)}} \geq \underbrace{l - \varepsilon.}_{\text{(1.45.m)}}$$

This shows that $\lim_{n \to \infty} f(x_n) = l$ indeed.

□

Example 1.46. (1) Let

$$f(x) := \operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0. \end{cases}$$

Images/sgn_gr.png

Figure 2: $f(x) = \operatorname{sgn}(x)$.

Then

$$\lim_{x \to x_0^-} \operatorname{sgn}(x) = -1 \qquad \text{and} \qquad \lim_{x \to x_0^+} \operatorname{sgn}(x) = 1$$

Note that these limits exist and neither of them agree with $f(0) = 0$.

(2) Let $f(x) := \lfloor x \rfloor$. Then:

$$\lim_{x \to x_0^-} f(x) = \begin{cases} \lfloor x \rfloor & x \notin \mathbb{Z} \\ x - 1 & x \in \mathbb{Z} \end{cases} \qquad \text{and} \qquad \lim_{x \to x_0^+} f(x) = \begin{cases} \lfloor x \rfloor & x \notin \mathbb{Z} \\ x & x \in \mathbb{Z}. \end{cases}$$

So, the left and right limits exist, despite having different values whenever $x \in \mathbb{Z}$.

1.1.6 More on continuity

First, we note that there are more algebraic rules of continuity (we already discussed addition, multiplication and division in Proposition 1.23):

Proposition 1.47. *If $f, g \colon E \to \mathbb{R}$ are functions that are continuous at $x_0 \in E$, then so are:*

(1) $|f|$,

(2) $\max\{f, g\}$, *where*

$$\max\{f, g\}(x) := \max\{f(x), g(x)\}$$

(3) $\min\{f, g\}$ *(defined similarly),*

(4) $f^+ := \max\{f, 0\}$,

(5) $f^- := \min\{f, 0\}$.

Example 1.48. We can use for example the continuity of the absolute value for squeezing. For example, let

$$g(x) := \begin{cases} 1 & \text{for } x \in \mathbb{Q} \\ x & \text{for } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

We claim that $g(x)$ is continuous at $x_0 = 1$. The main idea is that we can try to apply the Squeeze Theorem for the limit of functions using the following chain of inequalities:

$$-|x-1| \leq g(x) - 1 \leq |x-1|.$$

According to point (1) of Proposition 1.47, the function $|x-1|$ is continuous everywhere over $\mathbb{R}$; thus,

$$\lim_{x \to 1} -|x-1| = \lim_{x \to 1} |x-1| = 0,$$

so that by Proposition 1.20(5), it follows that $\lim_{x \to 1} f(x) - 1 = 0$. Hence, $\lim_{x \to 1} f(x) = 1 = f(1)$ and f is continuous at $x_0 = 1$.

1.1.7 Uniform continuity and Lipschitzianity

We introduce a stronger version of continuity.

Definition 1.49. A function $f\colon E \to \mathbb{R}$ is said to be *uniformly continuous* if for every $\varepsilon > 0$ there is a $\delta_\varepsilon > 0$ such that for all $x, y \in E$ then

$$\text{if } |x-y| < \delta_\varepsilon \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Remark 1.50. The notion of uniform continuity defined above is much stronger than that of continuity, cf. Definition 1.11. More precisely, using the characterization of continuity for a function $f\colon E \to \mathbb{R}$ given in Proposition 1.14, for any $x_0 \in E$ and any $\varepsilon > 0$ there exists $\delta > 0$, which depends on ε and x_0, such that for any $x \in E$

$$\text{if } |x - x_0| < \delta \Longrightarrow |f(x) - f(x_0)| < \varepsilon.$$

In Definition 1.49, the for a fixed $\varepsilon > 0$, the existence of $\delta > 0$ is no longer dependent on the choice of a base point $x_0 \in E$; instead, at this point such choice can be made independently (or rather, uniformly) from the points of E: it just depends on the choice of ε.

In view of the observation of the previous remark, we immediately have the following proposing showing that uniform continuity is a stornger property than continuity.

Proposition 1.51. *If $f\colon E \to \mathbb{R}$ is uniformly continuous then it is continuous.*

Example 1.52. *The function $f(x) := x^2 : \mathbb{R} \to \mathbb{R}$ is not uniformly continuous.* On the other hand, we have already seen that it is continuous as it is a polynomial.
Indeed, for any $x, y \in \mathbb{R}$,

$$|x^2 - y^2| = |x+y| \cdot |x-y|.$$

So, for any $\varepsilon > 0$ and $\delta > 0$, we may chose $x, y \in \mathbb{R}$ such that $|x+y| > \frac{2\varepsilon}{\delta}$, and $|x-y| = \frac{\delta}{2}$ – to do that, it suffices to choose two real numbers x, y that are very large but very close to each other. Thus, it follows that

$$|x-y| < \delta \quad \text{and} \quad |x^2 - y^2| > \frac{2\varepsilon}{\delta}\frac{\delta}{2} = \varepsilon.$$

We will see in the next section that if we consider a continuous function f over a closed bounded interval $[a, b]$ – rather than on an unbounded domain as in this case, where x^2 is considered over $\mathbb{R}$ – then f is absolutely continuous.

Example 1.53. *We show that* $\cos(x) : \mathbb{R} \to \mathbb{R}$ *is uniformly continuous and hence continuous.* Indeed,

$$|\cos(x) - \cos(y)| = 2\left|\sin\left(\frac{x+y}{2}\right)\right|\left|\sin\left(\frac{x-y}{2}\right)\right| \leq 2\left|\sin\left(\frac{x-y}{2}\right)\right| \leq 2|x-y|.$$

So, if we set $\delta = \frac{\varepsilon}{2}$, then we have

$$|x-y| \leq \delta \Rightarrow |\cos(x) - \cos(y)| \leq 2|x-y| \leq 2\delta = 2\frac{\varepsilon}{2} = \varepsilon.$$

This result, together with Proposition 1.28, implies that also functions such as $\cos(x^2)$, $\cos^2(x)$, etc. are continuous.

We introduce now a property that makes it particularly easy to show that a function is uniformly continuous.

Definition 1.54. A function $f\colon E \to \mathbb{R}$ is said to be *Lipschitz* if there exists a positive real number C such that for every $x, y \in E$, $|f(x) - f(y)| \leq C|x-y|$.

When the conditions of Definition 1.54 are satisfied we say that C is a Lipschitz constant for the Lipschitz function f.

Proposition 1.55. *Let* $f\colon E \to \mathbb{R}$ *be a function which is Lipschitz with Lipschitz constant* C*, where* E *is an open interval. Then* f *is uniformly continuous on* E*; hence* f *is also continuous on* E*.*

Proof. For a fixed positive real number ε in Definition 1.49, it suffices to take $\delta := \frac{\varepsilon}{C}$. □

Example 1.56. Let $f\colon [0,1] \to \mathbb{R}$ be the function $f(x) = x^2$. Then for any $x, y \in [0,1]$,

$$|x^2 - y^2| = |x-y||x+y| \leq C|x-y|, \tag{1.56.o}$$

where $C := \sup\{|x+y| \mid x, y \in [0,1]$. By definition, $C \leq 2$ – it not hard to show that actually $C = 2$ – hence we can rewrite (1.56.o) as

$$|x^2 - y^2| = |x-y||x+y| \leq 2|x-y|.$$

Hence, f is Lipschitz and thus uniformly continuous. Let us notice that

We will see in Example 1.62 that there exist functions that are uniformly continuous but not Lipschitz.

1.1.8 Left and right continuity

Lastly, we introduce left and right continuity, and we use this to define continuity on a closed interval.

Definition 1.57. Let $f\colon E \to \mathbb{R}$ be a function, and $x_0 \in E$.

(1) f *is left continuous at* x_0, if $\lim_{x \to x_0^-} f(x) = f(x_0)$.

(2) f *is right continuous at* x_0, if $\lim_{x \to x_0^+} f(x) = f(x_0)$.

In Definition 1.16 we defined what it means to be continuous on an open interval. For functions the domains of which are closed intervals the definition has to use left and right limits as well at the two endpoints:

Definition 1.58. A function $f\colon [a,b] \to \mathbb{R}$ is *continuous* if:

(1) f is continuous at any point contained in (a,b);

(2) f is left continuous at b; and,

(3) f is right continuous at a.

Example 1.59. The function $f\colon [-1,1] \to \mathbb{R}$ defined as $f(x) := \sqrt{1-x^2}$ is continuous. Indeed this is true by the following (where we use that $g(y) = \sqrt{y}$ is continuous on $\mathbb{R}_+^*$, which will be a consequence of our general theorem about the continuity of the inverse. Indeed, by applying the statement of Theorem 1.74 to $f(x) = x^2$ one obtains that $f^{-1} = g$ is continuous on $\mathbb{R}_+^*$):

(1) if $-1 < c < 1$, then $\sqrt{1-x^2}$ at c is continuous because $\sqrt{1-x^2}$ is the composition of $\sqrt{y}$ and $1-x^2$, and the latter is continuous at c and the former is continuous at $1-c^2$ (as $1-c^2 > 0$).

(2) $\sqrt{1-x^2}$ is left continuous at 1, because for all (x_n) converging to 1 from the left we have $\lim\limits_{n\to\infty} \sqrt{1-x_n^2} = 0$, as $\lim\limits_{n\to\infty} 1-x_n^2 = 0$, and $\lim\limits_{x\to 0^+} \sqrt{y} = 0$ according to Example 1.41.

(3) $\sqrt{1-x^2}$ is right continuous at -1 by almost verbatim the same argument as the previous point, one only needs to take $\lim\limits_{n\to\infty} x_n = -1$ instead of 1.

1.1.9 Consequences of Bolzano-Weierstrass

In this subsection we shall show how continuous functions defined over bounded closed intervals behave nicely. The proofs of all the results illustrated in this subsection heavily relies on Bolzano-Weierstrass **??**. As we will be assuming, throughout this section, that the domain $D(f)$ of a function f is closed bounded interval, given a sequence $(x_n) \subseteq D(f)$, by **??** we will always be able to assume that we can pass to a converging subsequence $(x_{n_k}) \subseteq (x_n)$ whose limit belongs to $D(f)$, since we are assuming $D(f)$ is closed.

We start by showing that a continuous function defined over a closed bounded interval is always uniformly continuous.

Theorem 1.60. *Let $a, b \in \mathbb{R}$. If $f\colon [a,b] \to \mathbb{R}$ is continuous, then f is uniformly continuous.*

Proof. Assume that f is not uniformly continuous. Then there is a $\varepsilon > 0$ such that for every $\frac{1}{n}$ there are x_n and $y_n \in [a,b]$ such that $|x_n - y_n| \le \frac{1}{n}$ and $|f(x_n) - f(y_n)| > \varepsilon$. By Bolzano-Weierstrass (**??**) we may assume that $\lim\limits_{n\to\infty} x_n = x_0 \in [a,b]$. However, then the condition $|x_n - y_n| \le \frac{1}{n}$ yields that we have also $\lim\limits_{n\to\infty} y_n = x_0$. Using again $|x_n - y_n| \le \frac{1}{n}$ together with the continuity of f we obtain that $|f(x_0) - f(x_0)| \ge \varepsilon$. This is a contradiction. □

Remark 1.61. For Theorem 1.60 to hold true, it is very important that the domain of f is a closed bounded interval $[a,b]$ for $a, b \in \mathbb{R}$. We have already seen that the function $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ is not uniformly continuous on any interval of the form $(a, +\infty[$, $a \in \mathbb{R} \cup \{-\infty\}$.

Example 1.62. The following example shows that there exists uniformly continuous functions that are not Lipschitz.
Let us consider $f\colon \mathbb{R}_+ \to \mathbb{R}$, $f(x) = \sqrt{x}$. Let us fix a real number $a > 0$, and define $g\colon [0,a] \to \mathbb{R}$ by $g = f|_{[0,a]}$, $g(x) = \sqrt{x}$. Theorem 1.60 shows that g is uniformly continuous, as we are taking the domain of definition of g to be a closed bounded interval. To show that g is not Lipschitz, it suffices to show that for any $C \in \mathbb{R}_+^* 0$ there exists $s, t \in [0,a]$, $s \neq t$ such that $|g(s) - g(t)| > C|s-t|$, or, equivalently, $\frac{|g(s)-g(t)|}{|s-t|} > C$. Let us fix $C > 0$. Then, there exists $t \in]0,1[$ such that $\frac{1}{\sqrt{t}} > C$, since $\lim_{x\to 0^+} \frac{1}{\sqrt{x}} = +\infty$. Taking $s = 0$, then $\frac{|g(s)-g(t)|}{|s-t|} = \frac{\sqrt{t}}{t} = \frac{1}{\sqrt{t}} > C$, which is what we wanted to prove. It is not hard to show, that actually even $f(x) = \sqrt{x}$ is uniformly continuous but not Lipschitz. The proof is left as an exercise.

Theorem 1.63. *If $f\colon [a,b] \to \mathbb{R}$ is continuous for some $a, b \in \mathbb{R}$, then there are $c, d \in [a,b]$ such that*

$$\begin{aligned} M &:= \sup_{x\in[a,b]} f(x) = \max_{x\in[a,b]} f(x) = f(c), \\ m &:= \inf_{x\in[a,b]} f(x) = \min_{x\in[a,b]} f(x) = f(d). \end{aligned}$$

Remark 1.64. The above theorem can be restated by saying that for a given function $f\colon [a,b] \to \mathbb{R}$, if f is continuous then the range $\mathrm{R}(f)$ of f is a closed and bounded interval, $\mathrm{R}(f) = [c,d]$.

Proof. We only prove the existence of $\max_{x\in[a,b]} f(x)$, the case of $\min_{x\in[a,b]} f(x)$ follows similarly.
First we prove that f is bounded from above, so that $\sup_{x\in[a,b]} f(x)$ must exist. Assume, by contradiction, that f is not bounded from above. That means that for each integer $n > 0$ there is $x_n \in [a,b]$ such that $f(x_n) \geq n$. As $(x_n) \subseteq [a,b]$ is a bounded sequence, by Bolzano-Weierstrass **??**, there exists a convergent subsequence $(x_{n_k}) \subseteq (x_n)$. Set $c := \lim\limits_{k\to\infty} x_{n_k}$. Then $c \in [a,b]$, and the following chain of equalities yields a contradiction:

$$\mathbb{R} \ni f(c) = \underbrace{\lim_{k\to\infty} f(x_{n_k})}_{f\,:\,[a,b]\to\mathbb{R} \text{ is continuous}} = \underbrace{+\infty}_{f(x_{n_k})\geq n_k \geq k}.$$

This concludes the statement that f is bounded from above.
Having proved that f is bounded from above, $\sup\limits_{x\in[a,b]} f(x)$ makes sense. Thus, we must prove that $\sup\limits_{x\in[a,b]} f(x) = \max\limits_{x\in[a,b]} f(x)$. By definition of supremum, there exists a sequence $(y_n) \subseteq [a,b]$ such that $f(y_n) \geq M - \frac{1}{n}$. In particular, $\lim\limits_{n\to\infty} f(y_n) = M$. By Bolzano-Weierstrass **??**, there exists a convergent subsequence $(y_{n_k}) \subseteq (y_n)$. Set $c := \lim\limits_{k\to\infty} y_{n_k}$. Then $c \in [a,b]$, and

$$f(c) = \underbrace{\lim_{k\to\infty} f(y_{n_k})}_{f\,:\,[a,b]\to\mathbb{R} \text{ is continuous}} = \underbrace{\lim_{n\to\infty} f(y_n)}_{??} = M.$$

□

Remark 1.65. The conclusion of the above theorem does not hold, if we do not assume that the domain of f is a closed bounded interval $[a,b]$, $a, b \in \mathbb{R}$. For example, take $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) := \frac{1}{x^2+1}$. Then f does not attain its minimum as $\mathrm{R}(f) =]0,1]$: in fact, $f(x) > 0$, $\forall x \in \mathbb{R}$ and f converges to 0 as x goes to $\pm\infty$.

Theorem 1.66 (Intermediate value theorem)**.** *Let $a, b \in \mathbb{R}$. If $f\colon [a,b] \to \mathbb{R}$ is continuous, then it takes each value between $M := \max\limits_{x\in[a,b]} f(x)$ and $m := \min\limits_{x\in[a,b]} f(x)$ at least once. More precisely, for each $c \in [m, M]$, there exists $d \in [a,b]$ such that $f(d) = c$.*

Idea. We give only the idea and we refer to the precise proof to page 81-82 of the book.

We know by the above theorem that there are $a', b' \in [a,b]$ such that $m = f(a')$ and $M = f(b')$. Hence, by replacing a with a' and b with b' (and some algebraic manipulation in the case when $b' < a'$), we may assume that $f(a) = m$, $f(b) = M$ and $m < c < M$. Then, the idea is to consider

$$S := \{x \in [a,b] | f(x) < c\}$$

Set $d := \sup S$. By the definition of sup, there is a sequence $(x_n) \subseteq S$ converging to d from the left and let y_n be any sequence converging to d from the right. Applying continuity to the first sequence shows that $f(d) \le c$, and by applying it to the second one shows that $f(d) \ge c$. So, $f(d) = c$.

$\square$

Example 1.67. In other words, Theorem 1.66 says that $\mathrm{R}(f) = [m, M]$. Hence, for example, the image of an interval $[a, b]$ via a continuous function f (whose domain contains $[a, b]$) cannot be $[c, d] \cup [e, d]$, $c < d < e < f$ – that is, it cannot be the union of two disjoint intervals.

Example 1.68. If $f : \mathbb{R} \to \mathbb{R}$ is a continuous function, such that $f(0) = 1$, $f(1) = 3$, $f(2) = -1$, then f attains the value 2 at least two times. Indeed, our assumptions say that the maximum of $f|_{[0,1]}$ is at least 3 and the minimum of $f|_{[0,1]}$ is at most 1. Hence, Theorem 1.66 applied to $f|_{[0,1]}$ yields that there is at least one $c \in [0, 1]$ such that $f(c) = 2$. Similarly, Theorem 1.66 applied to $f|_{[0,1]}$ yields that there is at least one $d \in [1, 2]$ such that $f(d) = 2$. Furthermore, $c \neq d$, because $c = d$ can only happen if $c = d = 1$. However, $f(c) = 3 \neq 2$. Hence, c and d are two distinct real numbers at which f takes the value 2.

We will apply the above theoretical result to find solutions of equations of the form $f(x) = x$. For example one can ask, if there is a solution of $\cos(x) = x$ for some $x \in \left[0, \frac{\pi}{2}\right]$. Corollary 1.69 lets us answer this question.

Corollary 1.69 (Banach fixed point theorem for closed intervals)**.** *Let $a, b \in \mathbb{R}$. If $f : [a, b] \to [a, b]$ is a continuous function, then there exists $x \in [a, b]$ such that $f(x) = x$.*

Given a set S and function $f : S \to S$, an element $s \in S$ such that $f(s) = s$ is called a *fixed point*.

Proof. Set $g(x) := f(x) - x$. Then $g(a) = f(a) - a \ge 0$ and $g(b) = f(b) - b \le 0$. So, by the intermediate value theorem, there is a real number $c \in [a, b]$ such that $0 = g(c)$. This is equivalent to $f(c) = c$. $\square$

Example 1.70. The function $\cos(x)|_{[0,\frac{\pi}{2}]} : \left[0, \frac{\pi}{2}\right] \to \mathbb{R}$ can be regarded as $\cos(x)|_{[0,\frac{\pi}{2}]} : \left[0, \frac{\pi}{2}\right] \to \left[0, \frac{\pi}{2}\right]$, since $\mathrm{R}(\cos(x)|_{[0,\frac{\pi}{2}]}) = [0, 1] \subset \left[0, \frac{\pi}{2}\right]$. Then the above theorem says that there is a fixed point x for which $\cos(x) = x$.

1.2 Monotonicity and invertibility of continuous functions

Let us recall the following definition.

Definition 1.71. Let $f : E \to \mathbb{R}$ be a function, $E \subset \mathbb{R}$.

(1) f is *strictly increasing* if $f(x) < f(y)$ for all $x < y$ in E.

(2) f is *strictly decreasing* if $f(y) > f(x)$ for all $x < y$ in E.

(3) $f : E \to \mathbb{R}$ is *strictly monotone*, if it is strictly increasing or strictly decreasing.

Corollary 1.72. *Let $a, b \in \overline{\mathbb{R}}$. If $f : (a, b) \to \mathbb{R}$ is strictly monotone and continuous, then the range $\mathrm{R}(f)$ is an open interval.*

Proof. Set

$$\begin{aligned} S &:= \sup\{\, f(x) \mid x \in (a, b) \,\}, \\ I &:= \inf\{\, f(x) \mid x \in (a, b) \,\}. \end{aligned}$$

First, we show that $S, I \notin \mathrm{R}(f)$. We only prove statement about S since the statement about I can be proven analogously. So, let us assume by contradiction that $S = f(c)$ for some $c \in (a, b)$. Choose $c < d \in (a, b)$ – here we are using that the interval is open!. Then, as f is

strictly increasing $f(d) > f(c) = S$, which is a contradiction with the definition of S.
We now show that $\mathrm{R}(f) = (I, S)$. Let us fix $p \in (I, S)$. By the definition of S and I, there exist $c, d \in (a, b)$ such that $f(c) < p < f(d)$. Then the Intermediate Value Theorem 1.66 implies that $p \in \mathrm{R}(f)$, since $p \in [f(c), f(d)] \subset \mathrm{R}(f)$. □

Theorem 1.73. *Let $f\colon E \to F$ be a continuous function on an interval E. Then, f is strictly monotone if and only if it is injective.*

Proof. We do not prove this in class, read the proof from page 84-85 of the book.

□

Theorem 1.74. *If $f\colon E \to F$ is continuous, strictly monotone and surjective function between intervals E, F. Then f^{-1} is also continuous.*

Let us recall that in the hypotheses of Theorem 1.74, the inverse function f^{-1} exists by Theorem 1.73.

Proof. We only show the case when E is an open interval (a, b), for some $a, b \in \overline{\mathbb{R}}$. In this case, F is also an open interval according to Corollary 1.72. Fix $0 < \varepsilon \in \mathbb{R}$ and $y_0 \in F$. Set $x_0 := f^{-1}(y_0)$. According to Corollary 1.72, there exist $c, d \in \mathbb{R}$ such that

$$\mathrm{R}(f|_{(x_0-\varepsilon, x_0+\varepsilon)}) = (c, d) \tag{1.74.a}$$

In particular, there exists $\delta > 0$ such that for every $y \in F$

$$\text{if } |y - y_0| \leq \delta \Rightarrow y \in (c, d). \tag{1.74.b}$$

For example, it suffices to take $\delta := \frac{\min\{|c-y_0|,|d-y_0|\}}{2}$: that is a choice of δ for which the above condition is satisfied.

We show that with the above choice of δ the definition of the continuity of f^{-1} at y_0 is satisfied. That is, for every $y \in F$,

$$|y - y_0| \leq \delta \Rightarrow \underbrace{y \in (c, d)}_{(1.74.b)} \Rightarrow \underbrace{|f^{-1}(y) - x_0| \leq \varepsilon.}_{(1.74.a)}$$

□

Example 1.75. Neither of the functions $\sin(x)$, $\cos(x)$, $\tan(x)$ and $\operatorname{cotan}(x)$ are invertible if considered as functions $\mathbb{R} \to \mathbb{R}$, as they are not injective in view of their periodicity. However, if we restrict their domains adequately they become strictly montone, and then, according to Theorem 1.74, their inverses are continuous too:

(1) $\arcsin(x)$ is the inverse of $\sin(x)|_{[-\frac{\pi}{2},\frac{\pi}{2}]}$. For example, $\arcsin\left(-\frac{1}{2}\right) = -\frac{\pi}{6}$, and $\arcsin\left(-\frac{1}{2}\right) \neq \frac{7\pi}{6}$, despite having $\sin\left(\frac{7\pi}{6}\right) = -\frac{1}{2}$ too.

(2) $\arccos(x)$ is the inverse of $\cos(x)|_{[0,\pi]}$.

(3) $\arctan(x)$ is the inverse of $\tan(x)|_{[-\frac{\pi}{2},\frac{\pi}{2}]}$.

2 Differentiation

Let $f\colon E \to \mathbb{R}$ be a real valued one variable function. We would like to approximate it with a linear one. That is, we would like to write

$$f(x) = f(x_0) + a(x - x_0) + r(x), \tag{2.0.a}$$

where a is a real number, and the error function $r(x)$ is small in a neighborhood of x_0. The question is: how small would we like $r(x)$ to be so that we obtain a "good" approximation? What kind of function do then realize formula (2.0.a) with our chosen conditions on $r(x)$?

Well, if we want our approximation to at least compute the right value of f at x_0, since

$$\lim_{x \to x_0} x - x_0 = 0,$$

we need to impose that $\lim\limits_{x \to x_0} r(x) = 0$. Even better, we would like $r(x)$ to be smaller than a linear function, otherwise the linear approximation in (2.0.a) will not be very precise. But what does it precisely mean that $r(x)$ should be smaller than a linear function? The precise mathematical wording is the following:

$$\lim_{x \to x_0} \frac{r(x)}{x - x_0} = 0. \tag{2.0.b}$$

Even better, taking $r_1(x) := \frac{r(x)}{x - x_0}$, we can rewrite the above condition as

$$r(x) = (x - x_0) r_1(x), \quad \text{and} \quad \frac{r_1(x)}{x - x_0} = 0. \tag{2.0.c}$$

The graph of the function $g(x) := f(x_0) + a(x - x_0)$ is a line in the cartesian plane. Considering

Figure 3: A differentiable function and the tangent line to the graph.

the graph of $f(x)$, then if we can show that that for f the error function $r(x)$ is smaller than linear, that is, if $r(x)$ satisfies the condition of (2.0.b), then the line representing the graph of g will be tangent to the graph of f at the point $(x_0, f(x_0))$.

At this point the central question is: for what functions f do a and $r(x)$ exists satisfying (2.0.a), (2.0.b), respectively?

If both (2.0.a) and (2.0.b) hold, then

$$\frac{r(x)}{x - x_0} + a = \frac{f(x) - f(x_0)}{x - x_0}, \quad x \neq x_0, \tag{2.0.d}$$

and, moreover, by taking the limit for $x \to x_0$ on both sides of this equation, using (2.0.b), it follows that

$$a = \underbrace{\lim_{x\to x_0} \frac{r(x)}{x - x_0} + a}_{\text{by (2.0.b)}} = \underbrace{\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}}_{\text{by (2.0.d)}}. \tag{2.0.e}$$

So, the existence of the real number a together with the sub-linear[2] behavior of the error term described in (2.0.b) imply that the limit on the right of (2.0.e) exists and it is finite. This discussion motivates the following definition.

Definition 2.1. Let $f\colon E \to \mathbb{R}$ be a function and let $x_0 \in E$.

(1) The function f is differentiable at x_0, if the limit

$$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{2.1.f}$$

exists and it is finite. We call the value of the limit in (2.1.f) the derivative of f at x_0 and we denote it by $f'(x_0)$.

(2) We say that $f\colon E \to \mathbb{R}$ is differentiable if it is differentiable at all points $x_0 \in E$.

(3) The function

$$f'\colon \{x \in E \mid f \text{ is differentiable at } x\} \to \mathbb{R}, \quad x \mapsto f(x)'$$

is called the derivative function of f. the domain of f' is composed of all points of E where the above limit exists.

Remark 2.2. (1) The derivative $f'(x_0)$ of f at x_0 can be also defined to be the unique real number c satisfying

$$f(x) = f(x_0) + c \cdot (x - x_0) + r(x), \tag{2.2.g}$$

where the function $r(x)$ satisfies $\lim_{x\to x_0} \frac{r(x)}{x-x_0} = 0$. As above, we can write $r(x) = (x - x_0)r_1(x)$ and $\lim_{x\to 0} r_1(x) = 0$. In the reminder of this section, we will also use the notation $\varepsilon_1(x)$ to denote the function $r_1(x)$.

(2) The definition of the derivative $f'(x_0)$ in Definition 2.1 can be summarized from a geometrical viewpoint by saying that the derivative is the limit (when it exists) for $x \to x_0$ of the slope of the unique line passing through $(x_0, f(x_0))$ and the point $(x, f(x))$ corresponding to x on the graph.

Example 2.3. Constant functions are differentiable everywhere. In fact, if $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = C$, $\forall x \in \mathbb{R}$, $C \in \mathbb{R}$, then for $x_0 \in \mathbb{R}$

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

Example 2.4. *We show that* $(x^2)' = 2x$.
For any $x_0 \in \mathbb{R}$, we need to compute the limit $\lim_{x\to x_0} \frac{x^2 - x_0^2}{x - x_0}$. Thus,

$$\lim_{x\to x_0} \frac{x^2 - x_0^2}{x - x_0} = \lim_{x\to x_0} \frac{(x - x_0)(x + x_0)}{x - x_0} = \lim_{x\to x_0} x + x_0 = 2x_0.$$

[2] Sublinear stands for "less than linear", that is, the condition defined in (2.0.b)

`Images/secants.jpg`

Example 2.5. *Similarly, if $a \in \mathbb{Z}_+$, then $(x^a)' = ax^{a-1}$.*
Indeed,

$$\begin{aligned}
\lim_{x\to x_0} \frac{x^a - x_0^a}{x - x_0} &= \lim_{x\to x_0} \frac{(x-x_0)(x^{a-1} + x^{a-2}x_0 + x^{a-3}x_0^2 + \cdots + x^1 x_0^{a-2} + x_0^{a-1})}{x - x_0} \\
&= \lim_{x\to x_0} x^{a-1} + x^{a-2}x_0 + x^{a-3}x_0^2 + \cdots + x x_0^{a-2} + x_0^{a-1} \\
&= \underbrace{\lim_{x\to x_0} x^{a-1} + \lim_{x\to x_0} x^{a-2}x_0 + \lim_{x\to x_0} x^{a-3}x_0^2 + \cdots + \lim_{x\to x_0} x x_0^{a-2} + \lim_{x\to x_0} x_0^{a-1}}_{\text{by the addition rule for finite limits and the fact that } \forall c \in \mathbb{N},\ \lim_{x\to x_0} x^c = x_0^c} \\
&= a x_0^{a-1}.
\end{aligned}$$

Example 2.6. *We show that* $\sin(x)' = \cos(x)$.

$$\begin{aligned}
\lim_{x\to x_0} \frac{\sin(x) - \sin(x_0)}{x - x_0} &= \lim_{x\to x_0} \frac{2\cos\left(\frac{x+x_0}{2}\right)\sin\left(\frac{x-x_0}{2}\right)}{x - x_0} \\
&= \lim_{x\to x_0} \cos\left(\frac{x+x_0}{2}\right) \cdot \underbrace{\lim_{x\to x_0} \frac{\sin\left(\frac{x-x_0}{2}\right)}{\frac{x-x_0}{2}}}_{\lim_{t\to 0}\frac{\sin(t)}{t}=1} = \cos(x_0),
\end{aligned}$$

where we could break up the limit in the multiplication thanks to Proposition 1.20.

Example 2.7. *Similarly,* $\cos(x)' = -\sin(x)$.

$$\begin{aligned}
\lim_{x\to x_0} \frac{\cos(x) - \cos(x_0)}{x - x_0} &= \lim_{x\to x_0} \frac{-2\sin\left(\frac{x+x_0}{2}\right)\sin\left(\frac{x-x_0}{2}\right)}{x - x_0} \\
&= \lim_{x\to x_0} -\sin\left(\frac{x+x_0}{2}\right) \cdot \underbrace{\lim_{x\to x_0} \frac{\sin\left(\frac{x-x_0}{2}\right)}{\frac{x-x_0}{2}}}_{\lim_{t\to 0}\frac{\sin(t)}{t}=1} = -\sin(x_0),
\end{aligned}$$

where we could break up the limit in the multiplication thanks to Proposition 1.20.

Differentiability is a stronger condition than continuity, as the following proposition readily shows.

Proposition 2.8. *If $f\colon E \to \mathbb{R}$ is differentiable at x_0, then it is continuous at x_0.*

Proof. This is a consequence of the following computation:

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} \underbrace{f(x_0) + (x-x_0)f(x_0)' + r(x)}_{\text{by (2.2.g)}} = \underbrace{f(x_0) + \lim_{x\to x_0} r(x)}_{\text{Proposition 1.20 and } \lim_{x\to x_0} x-x_0=0}$$

$$= f(x_0) + \underbrace{\lim_{x\to x_0} \frac{r(x)}{x-x_0}}_{=0 \text{by Remark 2.2}} \cdot \lim_{x\to x_0} (x-x_0) = f(x_0).$$

□

Example 2.9. The viceversa of Proposition 2.8 is not true. That is, if f is continuous at x_0, it does not necessarily have to be differentiable.
For example, let us consider the function $f(x) := |x|$. The function f is continuous on $\mathbb{R}$, in particular, it is continuous at $x_0 = 0$. On the other hand, f is not differentiable at 0, because that would imply that $\lim\limits_{x\to 0} \frac{|x|}{x}$ exists. However, since

Images/abs_val_gr.png

Figure 4: $f(x) = |x|$.

$$\lim_{x\to 0^-} \frac{|x|}{x} = \lim_{x\to 0^-} \frac{-x}{x} = -1 \neq 1 = \lim_{x\to 0^+} \frac{x}{x} = \lim_{x\to 0^+} \frac{|x|}{x}.$$

Proposition 1.42 implies that f is not differentiable at 0.

Example 2.10. (1) The function $f(x) := \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases}$ is not differentiable at 0, since it is not continuous at 0. On the other hand, outside 0, f is differentiable, since over $\mathbb{R}^*$, f is constant.

(2) The function $f(x) := \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R}\setminus\mathbb{Q} \end{cases}$ is not differentiable at any point of $\mathbb{R}$ since it is not continuous at any point of $\mathbb{R}$.

2.1 Computing derivatives

In this section we show how to compute derivatives. We first start by studying how derivatives behave with respect to the usual algebraic operations on $\mathbb{R}$, and then continue by studying how to compute derivatives with respect to composition and taking the inverse.

2.1.1 Addition

Proposition 2.11. *If $f, g\colon E \to \mathbb{R}$ are differentiable at x_0, then so is $\alpha f + \beta g$ for any $\alpha, \beta \in \mathbb{R}$, and furthermore*

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0).$$

Proof.

$$\begin{aligned}(\alpha f + \beta g)'(x_0) &= \lim_{x\to x_0} \frac{(\alpha f + \beta g)(x) - (\alpha f + \beta g)(x_0)}{x - x_0} \\ &= \alpha \underbrace{\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}}_{=f'(x_0)} + \beta \underbrace{\lim_{x\to x_0} \frac{g(x) - g(x_0)}{x - x_0}}_{=g'(x_0)} = \alpha f'(x_0) + \beta g'(x_0)\end{aligned}$$

where we could split the limit of the sum into the sum of the limits by the assumption on the differentiability of f, g at x_0, using Proposition 1.20. □

Example 2.12. $\left(5x^3 + 6x^2\right)' = \left(5x^3\right)' + \left(6x^2\right)' = 15x^2 + 12x$

2.1.2 Multiplication

Proposition 2.13. *If $f, g\colon E \to \mathbb{R}$ are differentiable at x_0, then so is $f \cdot g$, and furthermore*

$$(f \cdot g)'(x_0) = (fg' + f'g)$$

Proof.

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

where the fact that $\lim_{x\to x_0} f(x) = f(x_0)$ follows from Proposition 2.8 and we could split the limits of sum and multiplications using the differentiability of f, g at x_0 and Proposition 1.20. □

Example 2.14.

$$\left(x^2 \cos(x)\right)' = \left(x^2\right)' \cos(x) + x^2 \left(\cos(x)\right)' = 2x\cos(x) + x^2(-\sin(x)) = x(2\cos(x) - x\sin(x))$$

2.1.3 Division

Proposition 2.15. *If $f, g\colon E \to \mathbb{R}$ are differentiable at x_0, and $g(x_0) \neq 0$, then $\frac{f}{g}$ is also differentiable at x_0, and furthermore*

$$\left(\frac{f}{g}\right)'(x_0) = \left(\frac{gf' - fg'}{g^2}\right)(x_0)$$

In particular,

$$\left(\frac{1}{g}\right)'(x_0) = \left(\frac{-g'}{g^2}\right)(x_0)$$

Proof. We compute the now familiar limit

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

Grouping together, in the denominator of the last member of the previous equation, those terms that depend on $g(x_0)$ and $f(x_0)$, respectively, we obtain,

$$\lim_{x\to x_0}\frac{f(x)g(x_0)-f(x_0)g(x_0)+f(x_0)g(x_0)-f(x_0)g(x)}{g(x)g(x_0)(x-x_0)}$$
$$=\frac{g(x_0)}{\underbrace{g(x_0)\cdot\lim_{x\to x_0}g(x)}_{=g(x_0)}}\left(\underbrace{\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}}_{=f'(x_0)}\right)$$
$$-\frac{f(x_0)}{\underbrace{g(x_0)\cdot\lim_{x\to x_0}g(x)}_{=g(x_0)}}\left(\underbrace{\lim_{x\to x_0}\frac{g(x)-g(x_0)}{x-x_0}}_{=g'(x_0)}\right)$$
$$=\left(\frac{gf'-fg'}{g^2}\right)(x_0),$$

where the fact that $\lim_{x\to x_0}g(x)=g(x_0)$ follows from Proposition 2.8 and we could split the limits of sum and multiplications using the differentiability of f,g at x_0 and Proposition 1.20. □

Example 2.16. If $b>0$ is an integer and $x\neq 0$, then:

$$\left(\frac{1}{x^b}\right)'=-\frac{\left(x^b\right)'}{x^{2b}}=-\frac{bx^{b-1}}{x^{2b}}=\frac{-b}{x^{b+1}}.$$

That is, by setting $a=-b$ we obtain $(x^a)'=ax^{a-1}$. In particular, this shows that

$$(x^a)'=ax^{a-1}$$

holds for all integer a, not just the non-negative ones.

Example 2.17. If $x\neq k\pi+\frac{\pi}{2}$ for any $k\in\mathbb{Z}$, or, equivalently, if $\cos(x)\neq 0$, then

$$\tan(x)'=\left(\frac{\sin(x)}{\cos(x)}\right)'=\frac{\cos(x)(\sin(x))'-\sin(x)(\cos(x))'}{\cos(x)^2}$$
$$=\frac{\cos(x)\cos(x)-\sin(x)(-\sin(x))}{\cos(x)^2}=\frac{1}{\cos^2(x)}.$$

2.1.4 Composition of functions and derivatives

Proposition 2.18. *Let $f\colon E\to\mathbb{R}$, $g\colon G\to\mathbb{R}$ be functions such that that $f(E)\subseteq G$. Assume that f is differentiable at $x_0\in E$, and g is differentiable at $f(x_0)$ then $g\circ f\colon E\to\mathbb{R}$ is differentiable at x_0, and*

$$(g\circ f)'(x_0)=g'(f(x))\cdot f'(x_0).$$

Idea of the proof.

$$\begin{aligned}\lim_{x\to x_0}\frac{g(f(x))-g(f(x_0))}{x-x_0} &= \lim_{x\to x_0}\frac{g(f(x))-g(f(x_0))}{f(x)-f(x_0)}\frac{f(x)-f(x_0)}{x-x_0}\\ &= \lim_{x\to x_0}\frac{g(f(x))-g(f(x_0))}{f(x)-f(x_0)}\cdot\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} = g'(f(x_0))\cdot f'(x_0)\end{aligned}$$

□

Example 2.19. Let $f(x)=x^2$ and $g(y)=\cos(y)$. Then $f'(x)=2x$ and $g'(y)=-\sin(y)$. In particular,

$$\cos(x^2)' = (g\circ f)'(x) = (g'\circ f)(x)\cdot f'(x) = -\sin(x^2)2x.$$

Example 2.20. Let $a\in\mathbb{Z}$, $b\in\mathbb{Z}^+$, $f(x)=x^a$ and $g(y)=y^{\frac{1}{b}}$. Then according to Example 2.16 and Example 2.22, $f(x)'=ax^{a-1}$ and $g(y)'=\frac{1}{b}y^{\frac{1}{b}-1}$. Hence,

$$\left(x^{\frac{a}{b}}\right)' = \left((x^a)^{\frac{1}{b}}\right)' = (g\circ f)'(x) = (g'\circ f)(x)\cdot f'(x) = \frac{1}{b}\left(x^a\right)^{\frac{1}{b}-1}ax^{a-1} = \frac{a}{b}x^{\frac{a}{b}-a+a-1} = \frac{a}{b}x^{\frac{a}{b}-1}$$

So, *the formula $(x^r)'=rx^{r-1}$ holds also when r is any rational number* (as it did for $r\in\mathbb{Z}$ in Example 2.16).

2.1.5 Inversion of functions and derivatives

Proposition 2.21. *Let $f\colon E=(a,b)\to F$ be a bijective continuous function (so f is strictly monotone, and f^{-1} exists and is continuous by Theorem 1.74), and let $x_0\in E$ be such that $f'(x_0)\neq 0$. Then f^{-1} is differentiable at $y_0:=f(x_0)$, and we have*

$$\left(f^{-1}\right)'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}$$

Proof. The idea behind the proof of the proposition is that if we set $y=f(x)$ and $y_0=f(x_0$, we have

$$\frac{f^{-1}(y)-f^{-1}(y_0)}{y-y_0} = \frac{1}{\frac{y-y_0}{f^{-1}(y)-f^{-1}(y_0)}} = \frac{1}{\frac{f(f^{-1}(y))-f(f^{-1}(y_0))}{f^{-1}(y)-f^{-1}(y_0)}}.$$

Check page 109 for the precise proof. □

Example 2.22. If $f(x)=x^b$ for some integer $b\geq 1$, then $f^{-1}(y)=\sqrt[b]{y}=y^{\frac{1}{b}}$. So, $f'(x)=bx^{b-1}$, and

$$\left(y^{\frac{1}{b}}\right)' = \left(f^{-1}\right)'(y) = \frac{1}{f'(f^{-1}(y))} = \frac{1}{b\left(y^{\frac{1}{b}}\right)^{b-1}} = \frac{1}{b}y^{-\frac{b-1}{b}} = \frac{1}{b}y^{\frac{1}{b}-1}.$$

So, for $c=\frac{1}{b}$ (where $b\in\mathbb{Z}^*_+$), the formula for $(y^c)'=cy^{c-1}$. That is the formula is the same as in the case of c being an integer.

Example 2.23. Let $f(x)=\sin(x)|_{[-\frac{\pi}{2},\frac{\pi}{2}]}\colon\left[-\frac{\pi}{2},\frac{\pi}{2}\right]\to[-1,1]$. Then f is invertible and $f^{-1}(y)=\arcsin(y)$. Also, $f'(x)=\cos(x)|_{[-\frac{\pi}{2},\frac{\pi}{2}]}$, thus, for any $y\in]-1,1[$,

$$\arcsin'(y) = \frac{1}{\cos(\arcsin(y))} = \frac{1}{\sqrt{1-\sin^2(\arcsin(y))}} = \frac{1}{\sqrt{1-y^2}}.$$

2.1.6 The exponential function

For our last example in this section, we will discuss in details the exponential and logarithmic functions. Let us remind the reader that we defined

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

Definition 2.24. For $x \in \mathbb{R}$, we define

$$e^x := \sum_{k=0}^{\infty}\frac{x^k}{k!}.$$

Remark 2.25. Applying Definition 2.24, to $x = 0$ yields $e^0 = 1$. Furthermore, according to **??**, $e^1 = e$.

Proposition 2.26. *For any $x, y \in \mathbb{R}$, $e^{x+y} = e^x \cdot e^y$.*

Proof. This is an exercise in Week 10 exercise sheet. □

Corollary 2.27. *For any $x \in \mathbb{R}$, $e^{-x} = \frac{1}{e^x}$.*

Proof.

$$e^x \cdot e^{-x} = \underbrace{e^{x+(-x)}}_{\text{Proposition 2.26}} = e^0 = \underbrace{1}_{\text{Remark 2.25}}.$$

Dividing by e^x yields the statement (e^x cannot be 0, since then $e^x \cdot e^{-x} = 1$ could not hold). □

Corollary 2.28. *For every $x \in \mathbb{R}$, $e^x > 0$.*

Proof. For $x \geq 0$, then all the terms in the infinite sum in Definition 2.24 is at least zero, and the first term is 1. This implies the statement for $x \geq 0$.

So, we may assume from now that $x < 0$. We have $e^x = \frac{1}{e^{-x}}$ by Corollary 2.27. However, as now $-x > 0$ holds, the previous paragraph tells us that $e^{-x} > 0$, and hence also $\frac{1}{e^{-x}} > 0$. □

Proposition 2.29. $(e^x)' = e^x$

Proof. We need to show that

$$\lim_{x\to x_0}\frac{e^x - e^{x_0}}{x - x_0} = e^{x_0}.$$

This is equivalent to showing that

$$0 = \lim_{x\to x_0}\frac{e^x - e^{x_0}}{x - x_0} - e^{x_0} = \left(\lim_{x\to x_0}\frac{e^{x-x_0} - 1}{x - x_0} - 1\right)e^{x_0}.$$

By setting $y = x - x_0$, what we need to show is that

$$\lim_{y\to 0}\frac{e^y - 1}{y} - 1 = 0. \tag{2.29.a}$$

However, for $0 < |y| \leq 1$:

$$0 \leq \left|\frac{e^y - 1}{y} - 1\right| = \left|\frac{\sum_{k=0}^{\infty}\frac{y^k}{k!} - 1}{y} - 1\right| = \left|\sum_{k=2}^{\infty}\frac{y^{k-1}}{k!}\right| \leq \sum_{k=2}^{\infty}\frac{|y|^{k-1}}{k!} \leq |y|\sum_{k=2}^{\infty}\frac{|y|^{k-2}}{(k-2)!}$$

$$= |y|\sum_{k=0}^{\infty}\frac{|y|^k}{k!} \leq |y|e$$

Hence, by Squeeze Theorem Proposition 1.20(5), it follows that (2.29.a) holds indeed. □

Proposition 2.30. *We have* $\lim_{x\to+\infty} e^x = +\infty$, *and* $\lim_{x\to-\infty} e^x = 0$.

Proof. According to Definition 2.24, for all $x > 0$, $e^x \geq 1 + x$. As $\lim_{x\to+\infty} 1 + x = +\infty$, squeeze (point (5) of Proposition 1.36) shows that $\lim_{x\to+\infty} e^x = +\infty$. Then Corollary 2.27, Corollary 2.28 and point (3) of Proposition 1.36 show that $\lim_{x\to-\infty} e^x = 0$. □

Proposition 2.31. *The function* $e^x : \mathbb{R} \to \mathbb{R}$ *is strictly increasing.*

Proof. Choose $y > x \in \mathbb{R}$. We have to show that $e^y > e^x$. This is shown by the following computation:

$$e^y - e^x = \underbrace{(e^{y-x} - 1)}_{\substack{>1 \text{ by Definition 2.24, using} \\ y > x}} \cdot \underbrace{e^x}_{\substack{>0 \text{ by} \\ \text{Corollary 2.28}}} > 0$$

□

Corollary 2.32. *The range* $R(f)$ *of* $f := e^x : \mathbb{R} \to \mathbb{R}$ *is* $(0, +\infty) = \mathbb{R}^*_+$.

Proof. Follows immediately from Proposition 2.30 and Proposition 2.31. □

Definition 2.33.

(1) We define the (natural) logarithm function $\log(x) : \mathbb{R}^*_+ \to \mathbb{R}$ to be the inverse of the exponential function $f(x) = e^x$.

(2) For any $a \in \mathbb{R}^*_+$, the a-based exponential functions a^x is defined as

$$a^x := e^{x\cdot\log(a)}.$$

The logarithm in base a of x is the inverse function of the a-based exponential function a^x,

$$\log_a(x) := \frac{\log(x)}{\log(a)}.$$

(3) For any $a \in \mathbb{R}$, the a-th power functions is defined as

$$x^a := e^{a\cdot\log(x)}.$$

Remark 2.34. In the special cases where the functions of Definition 2.36 have been already defined (so x^a when $a \in \mathbb{Q}$, and a^x when $a = e$), they agree with the previously defined functions. This will be an exercise on the exercise sheet.

Example 2.35. (1) If $f(x) = e^x$, then $f^{-1}(x) = \log(x)$ and $f'(x) = e^x$ (Proposition 2.29). Hence:

$$(\log(x))' = \frac{1}{e^{\log(x)}} = \frac{1}{x}.$$

(2) Let $h : \mathbb{R}^*_+ \to \mathbb{R}$, $h(x) := x^x$. Then, defining $f(x) = x\log(x)$, $g(y) := e^y$,

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

Thus,

$$h'(x) = (f \circ g)'(x) = f'(g(x))g'(x) = x^x(\log(x) + 1).$$

Definition 2.36. (1) The *hyperbolic trigonometric* functions are defined below, and they are called hyperbolic sine/cosine/tangent/cotangent:

$$\sinh(x) := \frac{e^x - e^{-x}}{2}$$
$$\cosh(x) := \frac{e^x + e^{-x}}{2}$$
$$\tanh(x) := \frac{\sinh(x)}{\cosh(x)}$$
$$\coth(x) := \frac{\cosh(x)}{\sinh(x)}$$

The domains of all the above functions is $\mathbb{R}$, except for coth which is defined over $\mathbb{R}^*$, since $\sinh(0) = 0$.

Proposition 2.37. *We have:*

(1) $\sinh(x)' = \cosh(x)$

(2) $\cosh(x)' = \sinh(x)$

(3) $\tanh(x)' = \frac{1}{\cosh(x)^2}$

(4) $\coth(x)' = \frac{-1}{\sinh(x)^2}$

(5) $(x^a)' = ax^{a-1}$

(6) $(a^x)' = \log(a) \cdot a^x$

(7) $\log_a(x)' = \frac{1}{\log(a) \cdot x}$

The proof is left as an exercise.

2.2 One sided derivatives

Definition 2.38. If $f : E \to \mathbb{R}$ is a function and $x_0 \in E$ for which the then we say that the left (resp. right) derivative of f exists at x_0 if the function

$$\frac{f(x) - f(x_0)}{x - x_0} : E \setminus \{x_0\} \to \mathbb{R}$$

admits a left (resp. right limit). The value of this limit is then the left (resp. right) derivative.

Example 2.39. For $f(x) = |x|$ at $x = 0$ the left derivative is -1 and the right derivative is 1. In fact,

$$\lim_{x \to 0^+} \frac{f(x) - f(0)}{x - 0} = \frac{x}{x} = 1, \quad \lim_{x \to 0^-} \frac{f(x) - f(0)}{x - 0} = \frac{-x}{x} = -1.$$

As in the case of left and right limits, we can use left and right derivatives to decide when a function is differentiable at a given point.

Proposition 2.40. *Let $f : E \to \mathbb{R}$ be a function and $x_0 \in E$ a real number. Then f is differentiable at a point x_0 if and only if both its left and right derivatives exist and they agree. Furthermore, then the value of the derivative is the same as the common value of the left and the right derivatives.*

Proof. This is an immediate consequence of Proposition 1.42, Definition 2.1 and Definition 2.38. ☐

Example 2.41. (1) Let us consider the function $f: \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) := \begin{cases} x^2, & x \geq 0 \\ x^3, & x < 0 \end{cases}.$$

The function f is differentiable at 0 with derivative $f'(0) = 0$. Indeed,

$$\lim_{h \to 0^+} \frac{f(0+h) - f(0)}{h} = \lim_{h \to 0^+} \frac{h^2 - 0}{h} = \lim_{h \to 0^+} h = 0.$$

Similarly,

$$\lim_{h \to 0^-} \frac{f(0+h) - f(0)}{h} = \lim_{h \to 0^-} \frac{h^3 - 0}{h} = \lim_{h \to 0^-} h^2 = 0.$$

Since the left derivative and the right derivative exist and agree at $x_0 = 0$, we can conclude that

$$\lim_{h \to 0} \frac{f(0+h) - f(0)}{h} = 0.$$

Thus, f is differentiable at 0 with $f'(0) = 0$.

(2) Let us consider the function $f: \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) := \begin{cases} x+1 & \text{for } x \geq 0 \\ x & \text{for } x < 0 \end{cases}.$$

The function f is not continuous at $x_0 = 0$, as $\lim_{x \to 0^-} f(x) = 0$ whereas $\lim_{x \to 0^+} f(x) = 1$. As a differentiable function is continuous, then f is not differentiable at 0. On the other hand, f is differentiable outside of 0, since on a sufficiently small neighborhood of any point $x_0 \neq 0$, f is given by a linear function and we know that linear functions are differentiable.

2.3 Higher derivatives

Given a function f, we may try to iterate inductively the process of taking the derivative of f, thus obtaining what we will call the second derivative of f, the third, derivative of f, etc.

Definition 2.42. Let $f: E \to \mathbb{R}$ be a function.

(1) The second derivative f'' of f is the function

$$\begin{aligned} f'': \{x \in E \mid f' \text{ is differentiable at } x\} &\to \mathbb{R} \\ x &\mapsto f''(x) := (f')'(x). \end{aligned}$$

(2) Assume that the n-th derivative $f^{(n)}$ of f has been defined. Then the $(n+1)$-st derivative $f^{(n+1)}$ of f is the function

$$\begin{aligned} f^{(n+1)}: \{x \in E \mid f^{(n)} \text{ is differentiable at } x\} &\to \mathbb{R} \\ x &\mapsto f^{(n+1)}(x) := (f^{(n)})'(x). \end{aligned}$$

The n-th derivative of f at $x \in E$ is denoted by $f^{(n)}(x)$. For the first, second and third derivative of f, we will adopt the notation f', f'', f''' rather than $f^{(1)}, f^{(2)}, f^{(3)}$.

Example 2.43. (1) The second derivative of $f(x) = \arctan(x)$ is $f''(x): \mathbb{R} \to \mathbb{R}$,

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

(2) Let us consider the function $f\colon \mathbb{R}^* \to \mathbb{R}$ defined by $f(x) = e^{\frac{1}{x}}$. Then,

$$f''(x) = (f')'(x) = (e^{\frac{1}{x}} \cdot (-\frac{1}{x^2}))'$$
$$= (e^{\frac{1}{x}})' \cdot (-\frac{1}{x^2}) + e^{\frac{1}{x}} \cdot (-\frac{1}{x^2})' = e^{\frac{1}{x}}(\frac{1}{x^4} + \cdot\frac{2}{x^3}).$$

Example 2.44. Here we show an example of a function $f(x)$ such that $f(x)$ is differentiable two times, but not three times. That is, $f'(x)$ and $f''(x)$ exist for every $x \in \mathbb{R}$, but $f'''(0)$ does not exist.
Let us consider $f(x) := |x^3|$. Then, $f'(x)$ exists for all $x \in \mathbb{R}$, and:

$$f'(x) = \begin{cases} 3x^2 & x \geq 0 \\ -3x^2 & x < 0. \end{cases}$$

This is immediate at $x \neq 0$ from the formula

$$f(x) = \begin{cases} x^3 & \text{for } x \geq 0 \\ -x^3 & \text{for } x \leq 0 \end{cases}$$

To conclude the above first claim we just have to compute the left and the right derivatives of $f(x)$ at $x = 0$, and show that both are 0. Indeed:

$$\lim_{x \to 0^+} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^+} \frac{x^3 - 0}{x} = \lim_{x \to 0^+} x^2 = 0,$$

and

$$\lim_{x \to 0^-} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^-} \frac{-x^3 - 0}{x} = \lim_{x \to 0^-} -x^2 = 0.$$

This concludes our first claim.
Similarly, one can prove that $f''(x)$ exists for all $x \in \mathbb{R}$ and

$$f''(x) = \begin{cases} 6x & \text{for } x \geq 0 \\ -6x & \text{for } x < 0 \end{cases}$$

With other words, $f''(x) = 6|x|$. However, as $|x|$ is not differentiable at $x = 0$, we obtain that $f'''(0)$ does not exist.

Definition 2.45. $f\colon E \to \mathbb{R}$ is called a function of class C^n if its first n derivatives $f', f'', \dots, f^{(n)}$ exists and are all continuous at all points $x_0 \in E$.

Notation 2.46. To denote that a function $f\colon E \to \mathbb{R}$ is a C^n function, we will use the notation $f \in C^n(E, \mathbb{R})$. We will write $f \in C^\infty(E, \mathbb{R})$ if $f \in C^n(E, \mathbb{R})$, $\forall n \in \mathbb{N}$, and we will say that f is a C^∞ function.

Example 2.47. (1) According to Example 2.5, x, x^2, etc. are C^n for all n, that is they are C^∞ function. More precisely, for $a \in \mathbb{N}$, defining $f(x) = x^a$ then

$$f^{(n)}(x) = \begin{cases} 0 & \text{if } n > a \\ a \cdot (a-1) \cdot \dots \cdot (a-n+1)x^{a-n} & \text{for } n \leq a \end{cases}$$

(2) We can repeat the same computation for $f\colon [0, +\infty) \to \mathbb{R}$, $f(x) := x^\alpha = e^{\log(x)\alpha}$, $\alpha \in \mathbb{R} \setminus \mathbb{N}$. Then $f^{(x)} = \alpha \cdot (\alpha - 1) \cdot (\alpha - 2) \cdot \dots \cdot (\alpha - n + 1)x^{\alpha - n}$, $x > 0$.

(3) $|x| : \mathbb{R} \to \mathbb{R}$ is not C^1, cf. Example 2.39.

(4) $|x^3| : \mathbb{R} \to \mathbb{R}$ is C^2 but not C^3, cf. Example 2.44.

2.4 Local and global extrema

Definition 2.48. Let $f\colon E \to \mathbb{R}$ be a function and let $x_0 \in E$.

(1) The function f admits a point of *local maximum* at x_0 if there is a real number $\delta > 0$ such that $]x_0 - \delta, x_0 + \delta[\subset E$ and for every $x \in E$ if $|x - x_0| < \delta$ then $f(x) \leq f(x_0)$.

(2) The function f has a point of *local minimum* at x_0 if there is a real number $\delta > 0$ such that $]x_0 - \delta, x_0 + \delta[\subset E$ and for every $x \in E$ if $|x - x_0| < \delta$ then $f(x) \geq f(x_0)$.

(3) We say that $x_0 \in E$ is a point of local extremum for f if it is either a point of local minimum or of local maximum.

(4) The function f has a point of global maximum at x_0 if $f(x_0) \geq f(x)$, for all $x \in E$.

(5) The function f has a point of global minimum at x_0 if $f(x_0) \leq f(x)$, for all $x \in E$.

Remark 2.49. We shall also say that f admits a local maximum (resp. local minimum, local extremum, global maximum, global minimum) at x_0 to indicate that property (1) (resp. (2), (3), (4), (5)) defined above is satisfied.

Remark 2.50. Let $f\colon E \to \mathbb{R}$ be a function and $x_0 \in E$. If x_0 is a point of global maximum (resp. global minimum) for f and E contains a neighborhood of x_0 of the form $]x_0 - \delta, x_0 + \delta[,\ \delta > 0$, then x_0 is also a point of local maximum (resp. local minimum) for f.

Example 2.51. Let us consider the function f

The following proposition shows that any point of local extremum for a function f coincides with a zero of the derivative f'.

Proposition 2.52. *If $f\colon E \to \mathbb{R}$ is differentiable at x_0, and f admits a local extremum at x_0, then $f'(x_0) = 0$.*

Proof. We present the local maximum case, as one just need to reverse a few signs, to modify the proof to obtain from it the case of local minimum.
Hence, let us assume that $x_0 \in E$ is a point of local by Definition 2.48, there is a real number $\delta > 0$ such that

$$|x - x_0| \leq \delta \Rightarrow f(x) \leq f(x_0). \tag{2.52.a}$$

However, then

$$\lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{(x - x_0)} \leq \underbrace{\lim_{x \to x_0^+} \frac{0}{(x - x_0)}}_{x > x_0,\ \text{and (2.52.a)}} = 0, \tag{2.52.b}$$

and

$$\lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{(x - x_0)} \geq \underbrace{\lim_{x \to x_0^-} \frac{0}{(x - x_0)}}_{x < x_0,\ \text{and (2.52.a)}} = 0. \tag{2.52.c}$$

As $f(x)$ is differentiable, at x_0 the two above limits agree (Proposition 2.40). Hence the following stream of inequalities have to be all equalities, which conclude our proof:

$$\underbrace{0 \leq \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{(x - x_0)}}_{(2.52.c)} = f'(x_0) = \underbrace{\lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{(x - x_0)} \leq 0.}_{(2.52.b)}$$

□

Images/x2_gr.png

Figure 5: $f(x) = x^2$ has a global minimum at $x = 0$.

Example 2.53. (1) For $f(x) = x^2$, we have $f'(x) = 2x$. Hence, $f'(x) = 0$ if and only if $x = 0$. So, $x = 0$ is the only option for stationary point of f, thus also for a point of local extremum, and, indeed, f admits a local (and global) minimum at $x = 0$.

(2) For $f(x) = x^3$, we have $f'(x) = 3x^2$. So $f'(x) = 0$ if and only if $x = 0$ (as in the previous case). However, $f(0) = 0$ is not a local extremum. This underlines that Proposition 2.52 yields only a necessary, but not a sufficient condition for having a local extremum.

Images/x3_gr.png

Figure 6: $f(x) = x^3$, $f'(0) = 0$, but 0 is not a point of local extremum for f.

Definition 2.54. Let $f\colon E \to \mathbb{R}$ be a function and let $x_0 \in E$. Assume that f is differentiable at x_0. Then, we say that x_0 a stationary point (for f) if $f'(x_0) = 0$.

We have just seen that when $f'(x_0) = 0$, we cannot necessarily conclude that x_0 is a point of local extremum for f. On the other hand, if the domain of f is a closed bounded interval $[a, b]$, then Theorem 1.63 implies that f admits both a global maximum and a global minimum in $[a, b]$. Therefore, using Proposition 2.52, a point of global maximum for f can only be:

- either a stationary point, $x \in (a, b)$ such that $f'(x) = 0$, or

- $x = a$, or $x = b$.

Similarly, a point of global minimum for f can only be:

- either a stationary point, $x \in (a, b)$ such that $f'(x) = 0$, or
- $x = a$, or $x = b$.

Hence, to find the value of the global maximum and the global minimum of f over the interval $[a, b]$, it suffices to compute:

$$\max_{x\in[a,b]} f = \sup\{f(x) \mid x = a \text{ or } x = b \text{ or } f'(x) = 0\},$$
$$\min_{x\in[a,b]} f = \inf\{f(x) \mid x = a \text{ or } x = b \text{ or } f'(x) = 0\}.$$

This procedure guves an algorithmic approach to finding the values of global extrema.

Example 2.55. We compute the global minimum and the global maximum of

$$f(x) = \frac{4}{3}x^3 + \frac{3}{2}x^2 - x + 2$$

on the closed bounded interval $\left[-2, \frac{1}{2}\right]$. By the discussion in the paragraph before the example we have to compute:

$$f'(x) = 4x^2 + 3x - 1,$$

and then find the solutions of the equation $f'(x) = 0$. These are:

$$x = \frac{-3 \pm \sqrt{25}}{8} = \frac{-3 \pm 5}{8} = -1, \text{ and } x = \frac{1}{4}.$$

Then, we have to compute the function values at these two points, and at the endpoints of our interval. The point, where the function value is the maximal yields the maximum and where the function value is minimal yields the minimum of $f(x)$ on $\left[-2, \frac{1}{2}\right]$:

value of x	$f(x)$
-2	$-\frac{4}{3}8 + \frac{3}{2}4 + 2 + 2 = \frac{36-64}{6} + 4 = 4 - \frac{28}{6} = -\frac{4}{6}$
-1	$-\frac{4}{3} + \frac{3}{2} + 1 + 2 = \frac{9-8}{6} + 3 = 3 + \frac{1}{6}$
$\frac{1}{4}$	$\frac{4}{3\cdot 64} + \frac{3}{2\cdot 16} - \frac{1}{4} + 2 = \frac{4+18-48}{192} + 2 = \frac{-26}{192} + 2 = 2 - \frac{13}{96}$
$\frac{1}{2}$	$\frac{4}{3\cdot 8} + \frac{3}{2\cdot 4} - \frac{1}{2} + 2 = \frac{4+9-12}{24} + 2 = 2 + \frac{1}{24}$

So, $f(x)$ on $\left[-2, \frac{1}{2}\right]$ takes its minimum at $x = -2$ and its maximum at $x = -1$.

2.5 Rolle's and Mean Value theorem

We are ready to state and prove the two main results of this chapter: Rolle's theorem (Theorem 2.56) and the Mean value theorem (Theorem 2.58).

Theorem 2.56 (Rolle's Theorem)**.** *Let $f: [a, b] \to \mathbb{R}$ is a continuous function, for $a, b \in \mathbb{R}$. Assume that f is differentiable on (a, b), and that $f(a) = f(b)$. Then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. If f is constant on $[a, b]$, then $f'(x) = 0$, $\forall x \in (a, b)$ and so we are done.
Hence, we can assume that f is not constant. Then, according to Theorem 1.63, f has both a maximum and a minimum on $[a, b]$. However, as f is non-constant, one of these values have to be not equal to $f(a) = f(b)$. Formally, this means that there is a $c \in [a, b]$, such that $f(c) \neq f(a) = f(b)$. In particular, we must have $a < c < b$. Then, f is differentiable at c, and as f has a (local) extremum at c, we have $f'(c) = 0$ according to Proposition 2.52. $\square$

Remark 2.57. The differentiability assumption is needed for Theorem 2.56 to hold. For example, considering $f(x) := |x|$ on $[-1, 1]$, then $f(-1) = f(1)$, but there is no point in $[-1, 1]$ at the derivative of f is 0.

Theorem 2.58 (Mean value theorem)**.** *Let $f\colon [a, b] \to \mathbb{R}$ be a continuous function, $a, b \in \mathbb{R}$. Assume that f is differentiable on (a, b). Then there exists $c \in (a, b)$ such that $f'(c)(b - a) = f(b) - f(a)$.*

Proof. Apply Theorem 2.56 to $g(x) = f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x - a)$. □

Example 2.59. Let f be a C^1 function (see Definition 2.45) on $[-1, 1]$ such that $f(-1) = 2$, $f(0) = 4$ and $f(1) = 3$. We show using Theorem 2.58 that there is a $c \in]-1, 1[$ such that $f'(c) = 1$:

(1) Applying Theorem 2.58 to $f|_{[-1,0]}$ we obtain that there is an $a \in]-1, 0[$ such that $f'(a) = \frac{4-2}{0-(-1)} = 2$.

(2) Applying Theorem 2.58 to $f|_{[0,1]}$ we obtain that there is a $b \in]0, 1[$ such that $f'(b) = \frac{3-4}{1-0} = -1$.

(3) As f is C^1, f' is continuous. Hence, Theorem 1.66 implies that there is a $c \in (a, b) \subseteq]-1, 1[$ such that $f'(c) = 1$.

Corollary 2.60. *Let $f, g\colon [a, b] \to \mathbb{R}$ be continuous functions, $a, b \in \mathbb{R}$. Assume that f, g are differentiable over (a, b) and that $f'(x) = g'(x)$ for each $x \in (a, b)$. Then there exists a real number C such that $f(x) = g(x) + C$.*

Proof. By taking $h(x) := f(x) - g(x)$, it suffice to apply Lemma 2.61. □

Lemma 2.61. *Let $h\colon [a, b] \to \mathbb{R}$ a continuous function which is differentiable on (a, b). Assume that $h'(x) = 0$, for all $x \in (a, b)$. Then, h is a constant function.*

Recall that by saying that h is a constant function we simply mean that $\forall x \in [a, b]$, $h(x) = C$ for some fixed real number $C \in \mathbb{R}$ (indipendent of x).

Proof. Assume that h is not constant. Then, there exists $c, d \in [a, b]$, $c < d$ such that $h(c) \neq h(d)$. Then, the Mean Value Theorem 2.58 implies that there exists $e \in (c, d)$, such that $h'(e) = \frac{h(d)-h(c)}{d-c} \neq 0$; nonetheless, this contradicts our assumption that $h'()$. □

2.5.1 Monotone functions and differentials

We can apply the Mean Value Theorem to characterize the derivative of monotone (differentiable) functions.

Corollary 2.62. *Let $f\colon [a, b] \to \mathbb{R}$ be a continuous function, $a, b \in \mathbb{R}$. Assume that f is differentiable on (a, b). Then,*

(1) f is increasing (resp. decreasing) if and only if $f'(x) \geq 0$ (resp. ≤ 0);

(2) if $f'(x) > 0$ (resp. < 0) for all $x \in (a, b)$, then f is strictly increasing (resp. strictly decreasing).

Proof. We only prove the increasing case of (1), as the others are similar.

Images/x3_gr.png

Figure 7: $f(x) = x^3$

- First we assume that f is increasing. Then,

$$x \geq x_0 \implies f(x) \geq f(x_0) \implies \frac{f(x) - f(x_0)}{x - x_0} \geq 0,$$
$$x \leq x_0 \implies f(x) \leq f(x_0) \implies \frac{f(x) - f(x_0)}{x - x_0} \geq 0.$$

  Thus,

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

- Second, let us assume by contradiction that $f'(x) \geq 0 \, \forall x \in (a, b)$ and that f is not increasing. Hence, there are $a \leq c < d \leq b$, such that $f(c) > f(d)$. However, then Theorem 2.58 tells us that then there esists $e \in \mathbb{R}$, $c < e < d$ such that $f'(e) = \frac{f(d)-f(c)}{d-c} < 0$.

□

Example 2.63. Let $f : E \to \mathbb{R}$ be a strictly increasing (resp. strictly decreasing) function. Then, it does not necessarily follow that $f'(x) > 0$ (resp. $f'(x) < 0$). For example, $f(x) = x^3$ is strictly increasing, but $f'(0) = 3 \cdot 0^2 = 0$.

Example 2.64. Let us consider the function $f_a : \mathbb{R} \to \mathbb{R}$, $f(x) := \sin(x) + ax$, where $a \in \mathbb{R}$ is a fixed real number. Let us compute for what value of a f_a is monotone. As $f_a(x)$ is differentiable on $\mathbb{R}$, then f_a is monotone if and only if either $f'(x) \geq 0$, $\forall x \in \mathbb{R}$ or $f'(x) \leq 0$, $\forall x \in \mathbb{R}$. Thus, let us compute $f'(x)$:

$$f'(x) = \cos(x) + a.$$

Thus,

- f is increasing if and only if $a \geq 1$;
- f is decreasing if and only if $a \leq -1$.

Example 2.65. Using Corollary 2.62 and Proposition 2.37 we obtain that the all the functions of Definition 2.36 are either monotone, or become monotone when restricted to $\mathbb{R}_+^*$ or to $\mathbb{R}_-^*$.

$f(x)$	$D(f)$	f'	monotonicity
$\sinh(x)$	$\mathbb{R}$	$\cosh(x)$	increasing over $\mathbb{R}$
$\cosh(x)$	$\mathbb{R}$	$\sinh(x)$	decreasing over $\mathbb{R}_-^*$ and increasing over $\mathbb{R}_+^*$
$\tanh(x)$	$\mathbb{R}$	$\frac{1}{\cosh(x)^2}$	increasing over $\mathbb{R}$
$\coth(x)$	$\mathbb{R}^*$	$\frac{-1}{\sinh(x)^2}$	decreasing over $\mathbb{R}_-^*$ and over $\mathbb{R}_+^*$
$x^a,\ a > 0$	$\mathbb{R}_+$	$ax^{a-1},\ x \neq 0$	increasing over $\mathbb{R}_+^*$
$x^a,\ a < 0$	$\mathbb{R}_+$	$ax^{a-1},\ x \neq 0$	decreasing over $\mathbb{R}_+^*$
$a^x, a > 1$	$\mathbb{R}$	$\log(a) \cdot a^x$	increasing over $\mathbb{R}$
$a^x, 0 < a < 1$	$\mathbb{R}$	$\log(a) \cdot a^x$	decreasing over $\mathbb{R}$
$\log_a(x),\ a > 1$	$\mathbb{R}_+^*$	$\frac{1}{\log(a)\cdot x}$	increasing over $\mathbb{R}_+^*$
$\log_a(x),\ 0 < a < 1$	$\mathbb{R}_+^*$	$\frac{1}{\log(a)\cdot x}$	decreasing over $\mathbb{R}_+^*$

2.5.2 L'Hôpital's rule

L'Hôpital rule gives a method to compute limits of fractions of function which are in the indeterminate forms

$$\frac{0}{0}, \frac{\infty}{\infty},$$

that is, either both values of the limit of the denominator and of the limit of the numerator approach 0, or they both approach $-\infty$ or $+\infty$ – in the latter case, the sign of ∞ does not really matter.

Example 2.66. How can we compute $\lim_{x\to+\infty} \frac{e^x}{x}$? In this case,

$$\lim_{x\to+\infty} e^x = +\infty = \lim_{x\to+\infty} x.$$

In this example, we cannot answer using the algebraic rules of Proposition 1.20.

Luckily, the following theorem provides us with new tools to carry out this kind of computations.

Theorem 2.67 (L'Hôpital rule)**.** *Let $f, g\colon (a, b) \to \mathbb{R}$ be differentiable functions, and let $a, b \in \overline{\mathbb{R}}$. Assume that the following conditions hold:*

(1) (exactly) one the following conditions hold for x_0:

(i) $x_0 \in (a, b)$;

(ii) $x_0 = a \in \mathbb{R}$;

(iii) $x_0 = b \in \mathbb{R}$;

(iv) $x_0 = a = -\infty$;

(v) $x_0 = b = +\infty$;

(2) $g(x) \neq 0$ *and* $g'(x) \neq 0$ *for all* $x \in (a, b) \setminus \{x_0\}$*;*

(3) $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = \alpha$ *for* $\alpha = 0$ *or* $\alpha = \pm\infty$*.*

Then, in the respective cases we have the following implications for any $\mu \in \overline{\mathbb{R}}$*:*

Cases (1|i), (1|iv) and (1|v)	*if* $\lim_{x\to x_0} \frac{f'(x)}{g'(x)} = \mu \quad \Rightarrow \quad \lim_{x\to x_0} \frac{f(x)}{g(x)} = \mu$
Case (1|ii)	*if* $\lim_{x\to x_0^+} \frac{f'(x)}{g'(x)} = \mu \quad \Rightarrow \quad \lim_{x\to x_0^+} \frac{f(x)}{g(x)} = \mu$
Case (1|iii)	*if* $\lim_{x\to x_0^-} \frac{f'(x)}{g'(x)} = \mu \quad \Rightarrow \quad \lim_{x\to x_0^-} \frac{f(x)}{g(x)} = \mu$

Proof. We prove only the $\alpha = 0$ and $x_0 \in (a, b)$ case, and we refer to page 121-122 of the book for the rest. As f and g are differentiable at x_0 they are also continuous there, and hence

$$f(x_0) = \lim_{x\to x_0} f(x) = \alpha = 0 \qquad \text{and} \qquad g(x_0) = \lim_{x\to x_0} g(x) = \alpha = 0. \tag{2.67.a}$$

So, by the mean value theorem for derivatives, there is a real number $c(x)$ between x and x_0 such that

$$f'(c(x)) = \frac{f(x) - f(x_0)}{x - x_0}. \tag{2.67.b}$$

In particular, $c(x)\colon E \setminus x_0 \to I \setminus x_0$ is a function such that $\lim_{x\to x_0} c(x) = x_0$. Then:

$$\mu = \underbrace{\lim_{x\to x_0} \frac{f'(x)}{g'(x)}}_{\text{definition of } \mu} = \underbrace{\lim_{x\to x_0} \frac{f'(c(x))}{g'(c(x))}}_{\text{Proposition 1.30}} = \underbrace{\lim_{x\to x_0} \frac{\frac{f(x)-f(x_0)}{x-x_0}}{\frac{g(x)-g(x_0)}{x-x_0}}}_{(2.67.b)} = \lim_{x\to x_0} \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \underbrace{\lim_{x\to x_0} \frac{f(x)}{g(x)}}_{(2.67.a)}$$

□

Remark 2.68. We show that the property (3) in the statement of Theorem 2.67 is a necessary one. Indeed, we show that if the limit $\lim_{x\to x_0} \frac{f'(x)}{g'(x)}$ does not exist, then we cannot conclude anything about the limit $\lim_{x\to x_0} \frac{f(x)}{g(x)}$.

(1) Let us take $f(x) = x + \sin(x)$, $g(x) = x$. Then f, g are differentiable over $\mathbb{R}$,

$$\lim_{x\to+\infty} f(x) = +\infty = \lim_{x\to 0} g(x), \quad g(x) \neq 0 \neq g'(x),\ \forall x \in \mathbb{R}^*.$$

Moreover,

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

On the other hand,

$$\lim_{x\to+\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to+\infty} \frac{1 + \cos(x)}{1}$$

which is not defined since the limit $\lim_{x\to+\infty} \cos(x)$ does not exist. Hence, since the limit of the quotient of the derivatives of f, g does not exist, a priori, we cannot conclude anything about the limit of the quotient of f, g. Nonetheless, in this case we are lucky and we can still carry out the computation.

(2) Consider $f(x) = \sqrt{x} + \sin(x)$, $g(x) = x$. Then f, g are differentiable over $\mathbb{R}_+^*$,

$$\lim_{x\to+\infty} f(x) = +\infty = \lim_{x\to 0} g(x), \quad g(x) \neq 0 \neq g'(x),\ \forall x \in \mathbb{R}^*.$$

Moreover,

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

On the other hand,

$$\lim_{x\to+\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to+\infty} \frac{\frac{1}{2\sqrt{x}} + \cos(x)}{1}$$

which is not defined since the limit $\lim_{x\to+\infty} \cos(x)$ does not exist.

(3) Consider $f(x) = x + \sin(x)$, $g(x) = x$. Then f, g are differentiable over $\mathbb{R}_+^*$,

$$\lim_{x\to+\infty} f(x) = +\infty = \lim_{x\to 0} g(x), \quad g(x) \neq 0 \neq g'(x),\ \forall x \in \mathbb{R}^*.$$

Moreover,

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

On the other hand,

$$\lim_{x\to+\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to+\infty} \frac{\frac{1}{2\sqrt{x}} + \cos(x)}{1}$$

which is not defined since the limit $\lim_{x\to+\infty} \cos(x)$ does not exist.

Hence, if the limit of the quotient of the derivatives of f, g does not exist, a priori, we cannot conclude anything about the limit of the quotient of f, g. Nonetheless, in some cases, such as () here, we are lucky and we can still carry out the computation.

Example 2.69. Let us consider the limit $\lim_{x\to 0} \frac{\arcsin(x)}{\sin(x)}$. Then, $\lim_{x\to 0} \arcsin(x) = 0 = \lim_{x\to 0} \sin(x)$ and $\sin(x)' = \cos(x)$, $\arcsin(x)' = \frac{1}{\sqrt{1-x^2}}$.

Moreover, both $\sin(x)$ and $\cos(x)$ are non-zero over the pointed neighborhood $]-\frac{\pi}{2}, \frac{\pi}{2}[\setminus\{0\}$ of 0. Hence, we can apply Theorem 2.67 to get

$$\lim_{x\to 0} \frac{\arcsin(x)}{\sin(x)} = \lim_{x\to 0} \frac{\frac{1}{\sqrt{1-x^2}}}{\cos(x)} = 1$$

Example 2.70. Let us consider the limit $\lim_{x\to+\infty} \frac{e^x}{x^n}$. Then, $f(x) = e^x$, $g(x) = x^n$, $n \in \mathbb{N}$. Let us start with the case $n = 1$. Then,

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

For $n = 2$,

$$\lim_{x\to+\infty} \frac{e^x}{x^2} = \lim_{x\to+\infty} \frac{e^x}{2x} = \lim_{x\to+\infty} \frac{e^x}{2} = +\infty.$$

Hence, inductively, one can prove that

$$\lim_{x\to+\infty} \frac{e^x}{x^n} = \lim_{x\to+\infty} \frac{e^x}{nx^{n-2}} = \lim_{x\to+\infty} \frac{e^x}{n(n-1)x^{n-2}} = \dots = \lim_{x\to+\infty} \frac{e^x}{n!} = +\infty.$$

Hence, the exponential function e^x goes to $+\infty$ – as x goes to $+\infty$ – faster than any monomial x^n; a similar argument shows that it goes faster than any polynomial.

Images/arcsin_gr.png

Figure 8: $f(x) = \arcsin(x)$

Example 2.71. Similarly to the previous example,

$$\lim_{x\to 0^+} x^n \log(x) = \lim_{x\to 0^+} \frac{\log(x)}{\frac{1}{x^n}} = \lim_{x\to 0^+} \frac{\frac{1}{x}}{\frac{-n}{x^{n+1}}} = \lim_{x\to 0^+} -\frac{x^n}{n} = 0,$$

while,

$$\lim_{x\to +\infty} \frac{\log(x)}{x^n} = \lim_{x\to +\infty} \frac{\frac{1}{x}}{nx^{n-1}} = \lim_{x\to +\infty} \frac{1}{nx^n} 0$$

So, log goes to $-\infty$ as x goes to 0 and to $+\infty$ as x goes to $+\infty$ slower than $\frac{1}{x}$ and x, respectively.

2.5.3 Taylor expansion

Definition 2.72. Let $f\colon E \to \mathbb{R}$ be a function and let $x_0 \in E$ Assume that there is a neighborhood of $a \in E$ which is contained in the domain (so in E). We say that f admits an expansion to the n-th order x_0 if there is an equality of the form

$$f(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_n(x-a)^n + (x-a)^n \epsilon(x), \tag{2.72.c}$$

where a_i are real number, and $\epsilon_n(x)\colon E \to \mathbb{R}$ satisfies $\lim\limits_{x\to x_0} \epsilon_n(x) = 0$.

Proposition 2.73. *In the hypotheses of Definition 2.72, if a function f admits an n-th order expansion around a point $x_0 \in D(f)$, then the coefficients a_i in (2.72.c) are uniquely determined.*

Proof. Let

$$\begin{aligned} f(x) &= a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n + (x-x_0)^n \epsilon_n(x), \\ f(x) &= a_0' + a_1'(x-x_0) + a_2'(x-x_0)^2 + \cdots + a_n'(x-x_0)^n + (x-x_0)^n \epsilon_n'(x), \end{aligned}$$

be two different expansions to order n of f around x_0. We show by induction on i that $a_i = a_i'$. For $i = 0$ this is given by passing to the limit as $x \to x_0$ of the two expansion:

$$\begin{aligned} a_0 &= \lim_{x\to x_0} a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n + (x-x_0)^n \epsilon_n(x) \\ &= \lim_{x\to x_0} f(x) \\ &= \lim_{x\to x_0} a_0' + a_1'(x-x_0) + a_2'(x-x_0)^2 + \cdots + a_n'(x-x_0)^n + (x-x_0)^n \epsilon_n'(x) = a_0' \end{aligned}$$

Let us prove the induction step. Thus, let us assume that we know that $a_j = a'_j$ for $j = 0, \ldots, i-1$. Then,

$$\begin{aligned}
f(x) =& a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots + a_n(x-x_0)^n + (x-x_0)^n\epsilon_n(x) \\
=& a'_0 + a'_1(x-x_0) + a'_2(x-x_0)^2 + \cdots + a'_n(x-x_0)^n + (x-x_0)^n\epsilon'_n(x) \\
=& a_0 + a_1(x-x_0) + \cdots + a_{i-1}(x-x_0)^{i-1} + a'_i(x-x_0)^i + \cdots + a'_n(x-x_0)^n + (x-x_0)^n\epsilon'_n(x)
\end{aligned}$$

Hence, taking the expansions on the 1st and the 3rd line of the previous chain of equalities, and subtracting from both expansions $a_0 + a_1(x-a) + \cdots + a_{i-1}(x-a)^{i-1}$ and then dividing both by $(x-a)^i$, we obtain

$$\begin{aligned}
& a_i + a_{i+1}(x-x_0) + \cdots + a_n(x-x_0)^{n-i} + (x-x_0)^{n-i}\epsilon_n(x) \\
=& a'_i + a'_{i+1}(x-x_0) + \cdots + a'_n(x-x_0)^{n-i} + (x-x_0)^{n-i}\epsilon'_n(x).
\end{aligned}$$

Taking limit of this equality as $x \to x_0$ yields that $a_i = a'_i$, which concludes the induction step. Hence, $a_i = a'_i$ for each i. In particular, it also follows that $\epsilon(x) = \epsilon'(x)$ for each $x \in E$. $\square$

When is it that we can find an expansion to order n for a function f around a point $x_0 \in D(f)$? The following theorem provides a first answer.

Theorem 2.74. *Let $n \geq 0$ be an integer. Let $f\colon E \to \mathbb{R}$ be a function defined on an open interval E, and let $x_0 \in E$. Assume that f is $n+1$ times differentiable over E. Then, for each $x \in E$ there exists $x' \in]x_0, x[$, if $x > x_0$ (resp. $x' \in]x, x_0[$, if $x < x_0$) and such that*

$$f(x) = \left(\sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i\right) + f^{(n+1)}(x')\frac{(x-a)^{n+1}}{(n+1)!}.$$

Remark 2.75. Theorem 2.74 not only tells us that, under the hypotheses posed in its statement, it is possible to find an order n expansion for a function f around a point x_0 but also that, when that is the case, we have a recipe to compute the coefficients which are given by the formula

$$a_j = \frac{f^{(j)}(x_0)}{j!}.$$

Moreover, we can also compute the error term in the

Proof. To understand the proof, note that the statement for $n = 0$ is just the Mean Value Theorem, cf. Theorem 2.58. Indeed, that results implies that there exists $x' \in]x_0, x[$, if $x > x_0$ (resp. $x' \in]x, x_0[$, if $x < x_0$) such that $f'(x') = \frac{f(x)-f(x_0)}{x-x_0}$. Multiply by $x - x_0$, then

$$f(x) = f(x_0) + f'(x_0 + \theta_{x,x_0}(x-x_0))(x-x_0),$$

where $\theta_{x,x_0}(x-x_0) \in [0,1]$ and $x' = x_0 + \theta_{x,x_0}(x-x_0)$ – which is possible exactly because x' is contained between x and x_0. Let us recall that the proof of Theorem 2.58 was an application of Rolle's theorem to the function $g(y) := f(y) - f(x_0) - \frac{f(x)-f(x_0)}{x-x_0}(y-x_0)$. Furthermore, this techique was working since $g(x_0) = g(x)$, and $g'(y) = f'(y) - \frac{f(x)-f(x_0)}{x-x_0}$, so that $g'(y)$ being 0 yielded exactly the above equation.
Let us now define

$$P_n(x) := \sum_{i=0}^{n} \frac{f^{(i)}(x_0)}{i!}(x-x_0)^i,$$

and let us consider

$$g(y) = f(y) - P_n(y) + \frac{P_n(x) - f(x)}{(x - x_0)^{n+1}}(y - x_0)^{n+1}.$$

Then,

$$0 = g(x) = g(a) = g'(a) = \cdots = g^{(n)}(a),$$

which means that there exists y_1 between x_0 and x such that $g'(y_1) = 0$ by Rolle's theorem. But then applying Rolle's theorem again we obtain a y_2 between x_0 and y_1 such that $g^{(2)}(y_2) = 0$. Iterating this process we obtain a point y_{n+1} between x_0 and x such that $g^{(n+1)}(y_{n+1}) = 0$. In particular, by setting $x' := y_{n+1}$, then

$$0 = g^{(n+1)}(x') = f(x') + \frac{P_n(x) - f(x)}{(x - x_0)^{n+1}}(n + 1)!.$$

Reorganizing the latter equation yields exactly the statement of the theorem. $\square$

Corollary 2.76. *Let $n \geq 0$ be a real number. Let $f : E \to \mathbb{R}$ be a function defined on an open interval E. Assume that $f \in C^n(E, \mathbb{R})$, and let $x_0 \in E$. Then, the n-th order expansion of f around x_0 exists and is given by the formula*

$$f(x) = \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!}(x - x_0)^j + (x - x_0)^n \epsilon_n(x).$$

The idea behind the proof of the corollary is that by the previous theorem the error term is $f^{(n+1)}(x') - f^{(n+1)}(x)$, which converges to zero as x goes to a as x' is between a and x, and $f^{(n)}$ is continuous. For the precise proof we refer to page 126 of the book.

Example 2.77. Applying Corollary 2.76 to $f(x) = \frac{1}{1-x}$ and $x_0 = 0$ yields that the order n expansion takes the form

$$\frac{1}{1 - x} = 1 + x + x^2 + \cdots + x^n + x^n \epsilon_n(x),$$

since

$$f^{(i)}(x) = \frac{i!}{(1 - x)^{i+1}} \qquad \Rightarrow \qquad f^{(i)}(0) = i! \qquad \Rightarrow \qquad \frac{f^{(i)}(0)}{i!} = 1.$$

Example 2.78. Applying Corollary 2.76 to $f(x) = e^x$ and $x_0 = 0$ yields that the order n expansion takes the form

$$e^x = \sum_{i=0}^{n} \frac{x^i}{i!} + x^n \epsilon_n(x),$$

since

$$f^{(i)}(x) = e^x \qquad \Rightarrow \qquad f^{(i)}(0) = 1 \qquad \Rightarrow \qquad \frac{f^{(i)}(0)}{i!} = \frac{1}{i!}.$$

Example 2.79. Similarly, the $(2n + 1)$-st order expansion of $\cos(x)$ around $x = 0$ is

$$\cos(x) = \sum_{j=0}^{n} (-1)^n \frac{x^{2j}}{(2j)!} + x^{2n+1}, \epsilon(x)$$

while the $(2n + 2)$-nd order expansion of $\sin(x)$ around $x = 0$ is

$$\sin(x) = \sum_{j=0}^{n} (-1)^n \frac{x^{2j+1}}{(2j + 1)!} + x^{2n+2} \epsilon_{2n+2}(x).$$

Example 2.80. One can also figure out expansions of products, sums, compositions, etc.

For example the 3-rd order expansion of $\sin(\cos(x))$ is as follows:

$$\begin{aligned}\cos(\sin(x)) =& \cos\left(x - \frac{x^3}{6} + x^3\epsilon_3(x)\right)\\ =&1 - \frac{\left(x - \frac{x^3}{6} + x^3\epsilon_3(x)\right)^2}{2} + \left(x - \frac{x^3}{6} + x^3\epsilon_3(x)\right)^3 \eta_3(\sin(x)) = 1 - \frac{x^2}{2} + x^3\tau_3(x),\end{aligned}$$

where $x^3\tau_3(x)$ is the sum of all terms of the form $x^3h(x)$, where $\lim_{x\to 0} h(x) = 0$. In particular, $\lim_{x\to 0} \tau_3(x) = 0$ and hence the above is indeed the 3-rd order expansion.

In general, we can compute the expansion to order n of a composition $(f \circ g)(x)$ around a point $f(x_0)$ by substituting the expension to order n of g around x_0 into the expansion of f around $f(x_0)$ to order n and then re-ordering all the terms thus obtained up to order n. Let us highlight how one should be careful that the base-point of the expansion of the function f should be the value $g(x_0)$ of the function g at the base-point x_0. So, for example, $\sin(\cos(x))$ at 0 is not easy to compute this way, because one would need the expansion of sin around $\cos(0) = 1$, for which there is no nice formula.

Another example is by taking $h(x) = \frac{1}{1-(e^x-1)}$, $x_0 = 0$. Then we can rewrite h as the composition $h = f \circ g$ of $f(y) = \frac{1}{1-y}$ and $g(x) = e^x - 1$. Then,

$$\begin{aligned}\frac{1}{1-(e^x-1)} =& \frac{1}{1 - \left(x + \frac{x^2}{2} + \frac{x^3}{6} + x^3\epsilon_3(x)\right)}\\ =&1 + \left(x + \frac{x^2}{2} + \frac{x^3}{6} + x^3\epsilon_3(x)\right) + \left(x + \frac{x^2}{2} + \frac{x^3}{6} + x^3\epsilon_3(x)\right)^2\\ &+ \left(x + \frac{x^2}{2} + \frac{x^3}{6} + x^3\epsilon_3(x)\right)^3 \eta_3(x + \frac{x^2}{2} + \frac{x^3}{6} + x^3\epsilon_3(x))\\ =&1 + x + \left(\frac{1}{2} + 1\right) x^2 + \left(\frac{1}{6} + 2 \cdot 1 \cdot \frac{1}{2} + 1\right) x^3 + x^3\tau_3(x)\\ =&1 + x + \frac{3}{2}x^2 + \frac{13}{6}x^3 + x^3\tau_3(x).\end{aligned}$$

Similarly, one can write the order 3 expansion of $\frac{1}{1-x} \cdot e^x$ around 0 as

$$\begin{aligned}\frac{1}{1-x} \cdot e^x =& (1 + x + x^2 + x^3 + x^3\epsilon_3(x)) \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + x^3\eta_3(x)\right)\\ =&1 + 2x + \frac{5}{2}x^2 + \frac{8}{3}x^2 + x^3\tau_3(x)\end{aligned}$$

You can find more examples in the book, pages 127-131.

Example 2.81. One can use expansions also to avoid using Theorem 2.67. For example, to compute

$$\lim_{x\to 0} \frac{(e^x - 1 - x) + x\sin(x)}{\cos(x) - 1},$$

then we can try to compute the 2-nd order expansions first:

$$
\begin{aligned}
(e^x - 1 - x) + x\sin(x) &= \left(1 + x + \frac{x^2}{2} + x^2\epsilon_2(x) - 1 - x\right) + x(x + x^2\eta_2(x)) \\
&= \frac{x^2}{2} + x^2 + x^2(\eta_2(x) + \epsilon_2(x)) = \frac{3}{2}x^2 + x^2 \underbrace{\gamma_2(x)}_{\gamma_2(x):=\eta_2(x)+\epsilon_2(x)}), \\
\cos(x) - 1 &= 1 - \frac{x^2}{2} + x^2\tau_2(x) - 1 = -\frac{x^2}{2} + x^2\tau_4(x)
\end{aligned}
$$

Then,

$$
\lim_{x\to 0} \frac{(e^x - 1 - x) + x\sin(x)}{\cos(x) - 1} = \lim_{x\to 0} \frac{\frac{3}{2}x^2 + x^2\epsilon_3(x)}{-\frac{x^2}{2} + x^2\epsilon_4(x)} = \lim_{x\to 0} \frac{\frac{3}{2} + \epsilon_3(x)}{-\frac{1}{2} + \epsilon_4(x)} = -3.
$$

2.5.4 Application of Taylor expansion to local extrema and inflection points

We have proven that if f has a point of local extremum at $x_0 \in D(f)$, then $f'(x_0) = 0$, cf. Proposition 2.52. However, we have also shown that the converse implication does not hold, cf. Example 2.53. Nevertheless, we would like to know whether, for example, by imposing suitable conditions on the higher derivatives of a function f, we can still characterize when a stationary point is a point of local extremum for a function.

Let $f\colon E \to \mathbb{R}$ be a function, and let $x_0 \in E$ be a stationary point for f. Moreover, let us assume that for some even natural number n, the first $n-1$ derivatives of f vanish at x_0

$$
f'(x_0) = 0 = f''(x_0) = \cdots = f^{(n-1)}(x_0),
$$

while the n-th derivative of f is non-zero and $f^{(n)}(x_0) > 0$. Then, writing the n-th order expansion of f around x_0,

$$
f(x) = f(x_0) + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + (x - x_0)^n\epsilon_n(x).
$$

Thus, for x sufficiently close to x_0 it holds that $|\varepsilon_n(x)| < \frac{1}{2} \cdot \frac{f^{(n)}(x_0)}{n!}$ holds. In particular, for such values of x, then

$$
f(x_0) < f(x_0) + \frac{1}{2} \cdot \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \leq f(x)
$$

This shows that x_0 is a point of local minimum for f. One can imitate this argument for the case where $f^{(n)}(x_0) < 0$ to yield that x_0 is a point of local maximum for f. Thus, we can summarize the results obtained so far in the following theorem.

Theorem 2.82. *Let $n \geq 2$ be an even integer. Let $f\colon E \to \mathbb{R}$ be a function on an open interval E and let $x_0 \in E$. Assume that f is differentiable n times on E and that $f^{(i)}(x_0) = 0$, $\forall i = 1, 2, 3, \ldots, n-1$.*

(1) If $f^{(n)} > 0$, then f has a point of local minimum at x_0.

(2) If $f^{(n)} < 0$, then f has a point of local maximum at x_0.

Example 2.83. Consider the function $f(x) = \sin(x) + \frac{1}{2}x$ over the interval $[0, 2\pi]$, cf. Figure 9. Then, $f'(x) = 0$ if and only if $\cos(x) = -\frac{1}{2}$, which is equivalent to $x = \frac{2\pi}{3}$ or $\frac{4\pi}{3}$. Whether or not we have a maximum or minimum at these points is decided by the sign of $f''(x) = -\sin(x)$.

- At $x = \frac{2\pi}{3}$, $f(x)'' < 0$, so $f(x)$ has a local maximum, and

Images/sinxplushalfx_gr.png

Figure 9: $f(x) = \sin(x) + \frac{1}{2}x$ over the interval $[0, 2\pi]$

- At $x = \frac{4\pi}{3}$, $f(x)'' > 0$, so $f(x)$ has a local minimum.

Question 2.84. What happens if we assume that n is an odd natural number in the statement of Theorem 2.82?

In that case, the expansion to order n for f around x_0 takes the same expression as before

$$f(x) = f(x_0) + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + (x - x_0)^n \epsilon_n(x),$$

but this time the first non-leading term will be of the form $(x - x_0)^3$, $(x - x_0)^5$, or $(x - x_0)^7$, etc., depending on the precise value of n. But then for $x > x_0$, $(x - x_0)^n > 0$, while $x < x_0$, $(x - x_0)^n < 0$. This type of behavior characterizes what is called an inflection. That is to say, that for a stationary point x_0 to be an inflection point for f, we require that, on one side of x_0 the graph of the function is above the tangent line to the graph of f through $(x_0, f(x_0))$, and on the other side it is below it, cf. Figure 10

Images/hor_flex.png

Figure 10: $f(x) = x^3$ has a stationary point at $x = 0$ which is a flex, as the graph goes through the tangent line $y = 0$.

We can actually generalize this tentative definition, as follows, to comprise not just the case of stationary points.

Definition 2.85. Let $f\colon E \to \mathbb{R}$ be a function. Assume that f is differentiable at $x_0 \in E$. We say that f has an *inflection point* at x_0 if there exists $\delta > 0$ such that either one of the following two conditions is satisfied:

(1) $\{x \in E | a < x < a + \delta\} \Rightarrow f(x) - f(a) - f'(a)(x - a) > 0$, and
$\{x \in E | a - \delta < x < a\} \Rightarrow f(x) - f(a) - f'(a)(x - a) < 0$; or,

(2) $\{x \in E | a < x < a + \delta\} \Rightarrow f(x) - f(a) - f'(a)(x - a) < 0$, and
$\{x \in E | a - \delta < x < a\} \Rightarrow f(x) - f(a) - f'(a)(x - a) > 0$.

The reasoning contained in the paragraph before Definition 2.85 immediately yields the following result.

Theorem 2.86. *Let $n \geq 3$ be an odd integer. Let $f\colon E \to \mathbb{R}$ be a function defined over an open interval E and let $x_0 \in E$. Assume that f is differentiable n times on E and that*

$$f''(x_0) = \cdots = f^{(n-1)}(x_0) = 0,$$

while $f^{(n)}(x_0) \neq 0$. Then, f has an inflection point at x_0.

Example 2.87. Let us consider the function $f(x) = 2\sin(x) - x$. Then $f'(x) = 2\cos(x) - 1$, $f''(x) = -2\sin(x)$ and $f'''(x) = -2\cos(x)$. Hence, $f'(0) = f''(0) = 0$, and $f'''(0) \neq 0$. Hence $f(x)$ has an inflection point at $x = 0$ according to Theorem 2.86.

Figure 11: The function $f(x) = 2\sin(x) - x$ has a flex at the point $x = 0$, which is non-stationary, as $f'(0) = 1$, through the tangent line $y = x$ to the graph of f at the point $(0, 0)$.

2.5.5 Convex and concave functions

Definition 2.88. Let $f\colon E \to \mathbb{R}$ be a function defined over an open interval E. We say that f is convex (resp. concave) if for every $a, b \in E$ and every $\lambda \in [0, 1]$ we have:

$$f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b).$$

(resp. $f(\lambda a + (1 - \lambda)b) \geq \lambda f(a) + (1 - \lambda)f(b)$).

Remark 2.89. Let us maintain the same notation as in the above definition. We may assume that $a < b$. Then, for $\lambda \in [0,1]$, $x := \lambda a + (1-\lambda)b$ is a point between a and b. Geometrically, the above definition means the following:

(1) f is convex, if for any choice of $a, b \in E$, then between a and b, the graph of f lies completely below the line segment connecting $(a, f(a))$ and $(b, f(b))$;

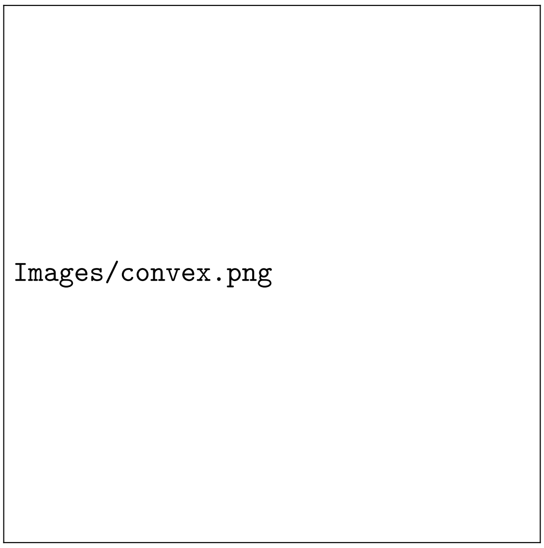

Figure 12: The graph of the function $f(x) = (x+1)^2 - 3$ lies below the line segment connecting the points $(-4, f(-4)) = -4, 6$ and $(1, f(1)) = (1, 1)$.

(2) if f is concave, then between a and b, the graph of f lies completely above the line segment connecting $(a, f(a))$ and $(b, f(b))$.

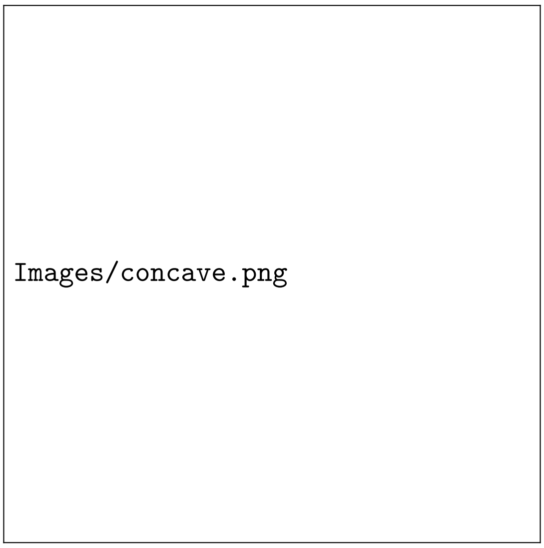

Figure 13: The graph of the function $f(x) = -(x-1)^2 + 5$ lies below the line segment connecting the points $(-2, f(-2)) = (-2, 4)$ and $(1, f(1)) = (1, 1)$.

We can characterize convexity and concavity of a function which is differentiable by means of the monotonicity of its first derivative.

Theorem 2.90. *Let $f\colon E \to \mathbb{R}$ be a function defined on an open interval E. Assume that f is differentiable. Then f is convex (resp. concave) if and only if $f'\colon E \to \mathbb{R}$ is an increasing (resp. decreasing) function.*

Proof. We prove only the statements about convexity, as f is convex if and only if $-f$ is concave.

(1) First, let us assume that f is convex. Let $a < b$ be points of I. We want to prove that $f'(a) \leq f'(b)$. By the above characterization of convexity we have

$$\frac{f(b) - f(\lambda a + (1-\lambda)b)}{b - (\lambda a + (1-\lambda)b)} \geq \frac{f(b) - f(a)}{b - a}, \text{ and } \frac{f(\lambda a + (1-\lambda)b) - f(a)}{(\lambda a + (1-\lambda)b) - a} \leq \frac{f(b) - f(a)}{b - a}.$$

Now, as λ goes to 0, the left side of the first inequality converges to $f'(b)$, and as λ goes to 1 the left side of the second inequality converges to $f'(a)$. This yields:

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

(2) For, the other direction let us assume that f' is increasing. Fix $a < b \in E$. Set $x := \lambda a + (1-\lambda)b$ for any $\lambda \in]0,1[$ (for $\lambda = 0$ and 1 the convexity inequality is automatic). Then, the mean value theorem tells us that there are $a < x_1 < x < x_2 < b$ such that $\frac{f(x)-f(a)}{x-a} = f'(x_1)$ and $\frac{f(b)-f(x)}{b-x} = f'(x_2)$. In particular, by our assumption that the derivative is increasing it follows that $\frac{f(x)-f(a)}{x-a} \leq \frac{f(b)-f(x)}{b-x}$. But this shows that f is convex by the above characterization of convexity in terms of slopes.

□

If f' is differentiable – or, equivalently, f is twice differentiable – then f' being increasing (resp. decreasing) is equivalent to $f'' \geq 0$ (resp. $f'' \leq 0$). Thus, we can characterize convexity and concavity of a function which is twice differentiable by means of the sign of its second derivative.

Corollary 2.91. *Let $f\colon E \to \mathbb{R}$ be a two times differentiable function on an open interval. Then f is convex (resp. concave) if and only if $f''(x) \geq 0$ (resp. $f''(x) \leq 0$) for all $x \in E$.*

Example 2.92. (1) Let us consider $f(x) = e^x$. Then, $f''(x) = e^x$, thus $f\colon \mathbb{R} \to \mathbb{R}$ is convex, since the second derivative $e^x > 0$, $\forall x \in \mathbb{R}$.

(2) Let us consider $f(x) = \log(x)$. Then, $f''(x) = \left(\frac{1}{x}\right)' = \frac{-1}{x^2}$, so the function $\log(x)$, which is defined over $\mathbb{R}_+^*$ is concave over its entire domain.

Example 2.93. Here we explain why the graph of a differentiable convex function f must lie above the tangent line to the graph through a point $(a, f(a))$. That is to say, we show that for a differentiable convex function f,

$$f(x) \geq f(a) + f'(a)(x - a).$$

We take $x > a$, and leave the case $x < a$ to the reader. Thus, assuming that $x > a$, we wish to show that $f(x) \geq f(a) + f'(a)(x - a)$, or equivalently that

$$\frac{f(x) - f(a)}{x - a} \geq f'(a). \tag{2.93.d}$$

Indeed, by the Mean value theorem (Theorem 2.58) there exists a $x' \in \mathbb{R}$, $a < x' < x$ such that

$$\frac{f(x) - f(a)}{x - a} = f'(x'). \tag{2.93.e}$$

As f is convex, f' is increasing by Theorem 2.90, hence, $f'(x') \geq f'(a)$. Adding this observation the equality in (2.93.e), we have shown that (2.93.d) must hold.

2.6 Asymptotes

Definition 2.94. (1) If for some $c \in \mathbb{R}$, $\lim_{x \to c^-} f(x) = \pm\infty$ or $\lim_{x \to c^+} f(x) = \pm\infty$, then we say that the function f has a vertical asymptote at $x = c$.

(2) If for some $c \in \mathbb{R}$, $\lim_{x \to +\infty} f(x) = c$ (resp. $\lim_{x \to -\infty} f(x) = c$), then we say that the function f has a horizontal asymptote at $+\infty$ (resp. $-\infty$) at $y = c$.

(3) If for some $a \neq 0, b \in \mathbb{R}$, $\lim_{x \to +\infty} f(x) - ax = b$ (resp. $\lim_{x \to -\infty} f(x) - ax = b$), then we say that f has an oblique (or slant) asymptote at $+\infty$ (resp. $-\infty$) along the line $y = ax + b$.

Example 2.95. Here we give a few examples of the different notions introduce in the above definition.

(1) Vertical asymptote: $f(x) = \frac{1}{1-x}$ has a vertical asymptote at at $x = 1$, cf. Figure 14;

(2) Horizontal asymptote: the function $f(x) = 2 - e^{-x}$ has a horizontal asymptote at $+\infty$ of value $y = 2$, cf. Figure 15;

(3) Slant asymptote: the function $f(x) = 2 + 3x + \frac{1}{x^2}$ has a slant asymptote both at $+\infty$ and $-\infty$ along the line $y = 3x + 2$, cf. Figure 16.

Images/1suxmeno1.png

Figure 14: The function $f(x) = \frac{1}{x-1}$, and the line $x = 1$.

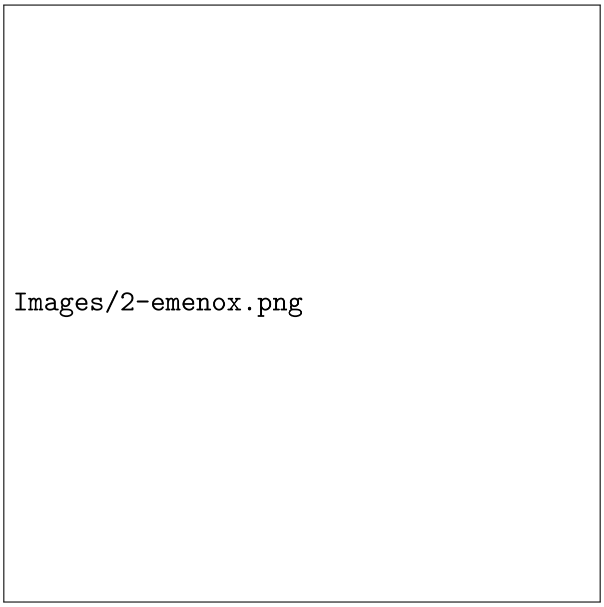

Figure 15: The function $f(x) = 2 - e^{-x}$, and the line $y = 2$.

3 Integration

3.1 Definition

The idea behind integration is that the integral $\int_a^b f(x)dx$ of a bounded function f on a closed interval $[a, b]$ should be the area under the graph of f. However, it is not that easy to say what this area means and when it is computable at all. If it is computable, we say that the function is integrable (Definition 3.11), and the value of this area is then called the integral $\int_a^b f(x)dx$ of f.

Now, the idea of trying to define the area under the graph of f is simple. We start with the only area that we can compute trustably, that is of rectangles, and then we try to approximate the area under the graph of f from above and from below using rectangles. These approximations are called upper and lower Darboux sums (Definition 3.3). We say that the area under the graph of f is computable, which as above means that the function is integrable, if these two approximations meet in the limit. This is spelled out in precise mathematical terms below.

Definition 3.1. (1) A *partition* $\sigma = (x_i)$ of a bounded interval $[a, b]$ is an ordered collection $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ of points of $[a, b]$.

(2) The *norm* or *mesh* of σ is

$$\max\{x_i - x_{i-1} | 1 \leq i \leq n\}.$$

(3) A *refinement* $\sigma' = (x'_i)$ of σ is a partition such that each value of x_i shows up amongst x'_i. we indicate that σ' is a refinement of σ by writing $\sigma \succeq \sigma$.

(4) The *regular partition* of length n is $x_i := a + i\frac{b-a}{n}$, $i = 0, 1, 2, \ldots, n$.

Proposition 3.2. *Given a bounded interval $[a, b]$, any two partitions σ, σ' have a common refinement σ''. Moreover, each partition can be refined to a new one with arbitrarily small norm.*

Images/2piu3xpiu1sux2.png

Figure 16: The function $f(x) = 2 + 3x + \frac{1}{x^2}$, and the line $y = 3x + 2$.

Definition 3.3. Let $f\colon [a,b] \to \mathbb{R}$ be a bounded function and $\sigma = (x_i)$ a partition of $[a,b]$. Then, the *upper Darboux sum* of f with respect to σ is

$$\overline{S}_\sigma = \sum_{i=1}^{n} M_i(x_i - x_{i-1}),$$

where $M_i := \sup_{x\in[x_{i-1},x_i]} f(x)$. The *lower Darboux sum* of f with respect to σ is

$$\underline{S}_\sigma = \sum_{i=1}^{n} m_i(x_i - x_{i-1}),$$

where $m_i := \inf_{x\in[x_{i-1},x_i]} f(x)$.

Example 3.4. Let us consider a constant function $f(x) = c$, $c \in \mathbb{R}$ over a closed bounded interval $[a, b]$. Then for any partition σ of $[a, b]$,

$$\underline{S}_\sigma = \overline{S}_\sigma = \sum_{i=1}^{n} c(x_i - x_{i-1}) = \underbrace{cx_n - cx_0}_{\text{telescopic sum}} = c(b-a).$$

Example 3.5. Let us consider the function $f(x) = x$ over a closed bounded interval $[a, b]$, and let $\sigma_n = \left(a + \frac{i(b-a)}{n}\right)$ be the regular partition of length n. Then,

$$\overline{S}_{\sigma_n} = \sum_{i=1}^{n} \left(a + i\frac{b-a}{n}\right) \frac{b-a}{n} = a(b-a) + \frac{n(n+1)}{2} \frac{(b-a)^2}{n^2}$$

and

$$\underline{S}_{\sigma_n} = \sum_{i=1}^{n} \left(a + (i-1)\frac{b-a}{n}\right) \frac{b-a}{n} = a(b-a) + \frac{(n-1)n}{2} \frac{(b-a)^2}{n^2},$$

where in both cases we used the following identity

$$\sum_{i=1}^{n} i = \frac{(n+1)n}{2}.$$

Note that $\lim_{n\to\infty} \overline{S}_{\sigma_n} = \lim_{n\to\infty} \underline{S}_{\sigma_n} = a(b-a) + \frac{(b-a)^2}{2} = \frac{b^2}{2} - \frac{a^2}{2}$.

Proposition 3.6. *Let $f : [a, b] \to \mathbb{R}$ be a function (not necessarily a continuous one). Assume that f admits an upper bound M (resp. a lower bound m) for its range $R(f)$. Then, for any partition σ of $[a, b]$, $m(b-a) \le \overline{S}_\sigma, \underline{S}_\sigma \le M(b-a)$. In particular, the sets*

$$\{\overline{S}_\sigma | \sigma \text{ is a partition of } [a, b]\} \quad \text{and} \quad \{\underline{S}_\sigma | \sigma \text{ is a partition of } [a, b]\}$$

are bounded.

Proof. This follows immediately from the definitions, since for any interval $[x_i, x_{i+1}] \subset [a, b]$ then

$$m \le \inf_{[x_i, x_{i+1}]} f \le \sup_{[x_i, x_{i+1}]} f \le M.$$

Hence, for a partition $\sigma = \{x_i\}$ of $[a, b]$,

$$\begin{aligned} m(b-a) =& m \sum_{i=0}^{n-1} m(x_{i+1} - x_i) \le \sum_{i=0}^{n-1} \left(\inf_{[x_i, x_{i+1}]} f\right)(x_{i+1} - x_i) = \overline{S}_\sigma \\ \le& \sum_{i=0}^{n-1} \left(\sup_{[x_i, x_{i+1}]} f\right)(x_{i+1} - x_i) = \underline{S}_\sigma \le \sum_{i=0}^{n-1} M(x_{i+1} - x_i) = M(b-a) \end{aligned}$$

$\square$

Definition 3.7. Let $f : [a, b] \to \mathbb{R}$ be a bounded function.

(1) The upper Darboux integral of f on $[a, b]$ is defined as

$$\overline{S} := \inf\{\overline{S}_\sigma | \sigma \text{ is a partition of } [a, b]\}.$$

(2) The lower Darboux integral of f on $[a, b]$ is defined as

$$\underline{S} := \sup\{\underline{S}_\sigma | \sigma \text{ is a partition of } [a, b]\}$$

Example 3.8. Using the above computation for the constant function, cf. Example 3.4, we see that if f is the constant function on $[a, b]$, then $\overline{S} = \underline{S} = (b - a)c$.

Proposition 3.9. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function.*

(1) If σ is a partition of $[a, b]$ and σ' is a refinement of σ, then:

$$\underline{S}_\sigma \leq \underline{S}_{\sigma'}, \text{ and } \overline{S}_\sigma \geq \overline{S}_{\sigma'}.$$

(2) If σ is a partition of $[a, b]$, then:

$$\underline{S}_\sigma \leq \overline{S}_\sigma.$$

Corollary 3.10. *If $f : [a, b] \to \mathbb{R}$ is a bounded function, then $\underline{S} \leq \overline{S}$.*

Proof. It is enough to prove that $\underline{S}_{\sigma_1} \leq \overline{S}_{\sigma_2}$ for any partitions σ_1 and σ_2 of $[a, b]$. However, this follows straight from Proposition 3.9. Indeed, if σ is a common refinement of σ_1 and σ_2, then Proposition 3.9 yields that

$$\underline{S}_{\sigma_1} \leq \underbrace{\underline{S}_\sigma}_{\text{Proposition 3.9.(1)}} \leq \underbrace{\overline{S}_\sigma}_{\text{Proposition 3.9.(2)}} \leq \underbrace{\overline{S}_{\sigma_2}}_{\text{Proposition 3.9.(1)}}.$$

$\square$

Definition 3.11. Let $f : [a, b] \to \mathbb{R}$ be a bounded function. We say that f is *integrable*, if $\overline{S} = \underline{S}$, in which case this common value is called the *integral of f between a and b*, and it is denoted by

$$\int_a^b f(x)dx.$$

Remark 3.12. Using Corollary 3.10, f is integrable over a closed bounded interval $[a, b]$ if one exhibits a sequence (σ_n) of partitions such that $\lim_{n\to\infty} \overline{S}_{\sigma_n} = \lim_{n\to\infty} \underline{S}_{\sigma_n}$. Indeed, this follows immediately by the following chain of inequalities

$$\lim_{n\to\infty} \underline{S}_{\sigma_n} \leq \underline{S} \leq \overline{S} \leq \lim_{n\to\infty} \overline{S}_{\sigma_n}, \tag{3.12.a}$$

passing to the limit for $n \to \infty$.

Example 3.13. Using Example 3.4, the constant functions are integrable on $[a, b]$, and

$$\int_a^b c\ dx = (b - a)c$$

Example 3.14. Using Remark 3.12 and the computation of Example 3.5 for $f(x) := x$ over a closed bounded interval $[a, b]$ then f is integrable, and

$$\int_a^b x dx = \frac{b^2}{2} - \frac{a^2}{2}$$

Example 3.15. Consider the function $[0, 2] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{Q} \\ 3 & \text{if } x \notin \mathbb{Q} \end{cases}$$

Then, for all partition σ, $\overline{S}_\sigma = 6$, and $\underline{S}_\sigma = 0$. So, $\overline{S} = 6$, $\underline{S} = 0$, and hence f is not integrable.

Proposition 3.16. *If $f : [a, b] \to \mathbb{R}$ is continuous then it is integrable.*

Proof. As f is continuous over a closed bounded interval $[a, b]$, Theorem 1.60 implies the uniform continuity of f. Let us fix $\varepsilon > 0$. Let $\delta > 0$ be the constant in the definition of uniform continuity associated to $\frac{\varepsilon}{b-a}$ – that is,

$$|x - y| \leq \delta \Rightarrow |f(x) - f(y)| \leq \frac{\varepsilon}{b-a}.$$

Claim Let σ be a partition of $[a, b]$ with norm at most δ. Then $\overline{S}_\sigma - \underline{S}_\sigma \leq \varepsilon$.

Proof. In fact,

$$\begin{aligned}
\overline{S}_\sigma - \underline{S}_\sigma &= \sum_{i=1}^{n} (\max_{[x_i, x_{i+1}]} f - \min_{[x_i, x_{i+1}]} f)(x_i - x_{i-1}) \\
&\leq \sum_{i=1}^{n} \frac{\varepsilon}{b-a}(x_i - x_{i-1}) = \frac{\varepsilon}{b-a} \underbrace{(b-a)}_{\sum_{i=1}^{n}(x_i - x_{i-1}) = b-a} = \varepsilon
\end{aligned}$$

□

□

3.2 Basic properties

Proposition 3.17. *Let $f, g \colon [a, b] \to \mathbb{R}$ be integrable functions. Then,*

(1) If f extends over $[b, c]$ for some $c \in \mathbb{R}$, $c > b$ and f is also integrable over $[b, c]$, then it is integrable over $[a, c]$, and

$$\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx.$$

(2) Given $\alpha, \beta \in \mathbb{R}$, $\alpha f + \beta g$ is integrable on $[a, b]$, and

$$\int_a^b (\alpha f + \beta g)(x)dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx$$

(3) If $f \leq g$, then

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx$$

(4) The function $|f|$ is integrable on $[a, b]$, and

$$\int_a^b |f(x)|dx \geq \left| \int_a^b f(x)dx \right|$$

Proof. The proofs of all these statements follow the same pattern: one writes up the inequalities for lower (resp. upper) Darboux sums for a fixed partition σ. Then these inequalities remain valid when taking sup (resp. inf) of the lower (resp. upper) Darboux sums along all possible partitions of an interval $[a, b]$. This gives inequalities in both direction, which then implies equalities.
For example, let us show how this strategy works in the case of point (1) – we leave the other

cases to the reader. Let σ, and τ be partitions for $[a, b]$ and $[b, c]$, respectively. Then the union of σ and τ gives a partition ρ for $[a, c]$ and by definition we have

$$\overline{S}_\rho = \overline{S}_\sigma + \overline{S}_\tau, \tag{3.17.a}$$

$$\underline{S}_\rho = \underline{S}_\sigma + \underline{S}_\tau. \tag{3.17.b}$$

As both these equalities are true for all choice of partitions σ of $[a, b]$ and τ of $[b, c]$, by taking the inf (resp. the sup) of (3.17.a) (resp. of (3.17.b)) along all possible choices of partitions of $[a, b], [b, c], [a, c]$, then

$$\overline{S}^{[a,c]} \leq \overline{S}^{[a,b]} + \overline{S}^{[b,c]}, \text{ and } \underline{S}^{[a,c]} \geq \underline{S}^{[a,b]} + \underline{S}^{[b,c]}, \tag{3.17.c}$$

where $\overline{S}^{[a,c]}$ (resp. $\underline{S}^{[a,c]}$) denotes the upper (resp. lower) Darboux integral of partitions of $[a, c]$, and similarly for the other cases. As f is integrable both on $[a, b]$ and on $[b, c]$, we have $\int_a^b f(x)dx = \overline{S}^{[a,b]} = \underline{S}^{[a,b]}$ and $\int_b^c f(x)dx = \overline{S}^{[b,c]} = \underline{S}^{[b,c]}$. Thus, (3.17.c) yields

$$\int_a^b f(x)dx + \int_b^c f(x)dx \leq \underline{S}^{[a,c]} \leq \overline{S}^{[a,c]} \leq \int_a^b f(x)dx + \int_b^c f(x)dx. \tag{3.17.d}$$

As the two ends of (3.17.d) are the same, we have everywhere equalities. This concludes both the integrability of f over $[a, c]$ as well as the statement of (1). $\square$

Example 3.18.

$$\int_a^b (1+x)dx = \underbrace{\int_a^b 1dx + \int_a^b xdx}_{\text{point (2) of Proposition 3.17}} = \underbrace{(b-a)}_{\text{Example 3.13}} + \underbrace{\frac{b^2-a^2}{2}}_{\text{Example 3.14}}.$$

3.3 Fundamental theorem of calculus

In this section, we learn how to compute integrals using the anti-derivative, cf. Theorem 3.24. We start by giving the definition of an anti-derivative.

Definition 3.19. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. A function $G : [a, b] \to \mathbb{R}$ is called an *anti-derivative* of f if

(1) G is continuous on $[a, b]$,

(2) G is differentiable on $]a, b[$, and

(3) $G'(x) = f(x)$ for all $x \in]a, b[$.

Remark 3.20. Given a continuous function $f\colon [a, b] \to \mathbb{R}$ admitting an anti-derivative $G\colon [a, b] \to \mathbb{R}$, then for any $C \in \mathbb{R}$, also $G_C\colon [a, b] \to \mathbb{R}$, $G_C(x) := G(x) + C$ is an anti-derivative for f. According to Corollary 2.60, the vice versa is also true: namely, if G, H are anti-derivatives of f, then there exists $C \in \mathbb{R}$ such that $G(x) = H(x) + C$, $\forall x \in \mathbb{R}$.

Notation 3.21. The anti-derivatives of f are at times denoted by $\int f(x)dx + c$, where $c \in \mathbb{R}$ is a constant that is free to vary in $\mathbb{R}$. Also, sometimes the expressions $\int f(x)dx + c$ is also called the indefinite integral, while what we defined as the integral of f over $[a, b]$, $\int_a^b f(x)dx$ is then called the definite integral – where the definitiveness comes from the fact that we computed the integral over the closed bounded integral $[a, b]$. We use the integral/anti-derivative naming in this course.

Example 3.22. We collect here a few important functions together with their anti-dervatives.

function	e^x	$\cos(x)$	$\sin(x)$	$\frac{1}{x}$	$x^n,\ n \in \mathbb{N}$	...
anti-derivative	e^x	$\sin(x)$	$-\cos(x)$	$\log\|x\|$	$\frac{x^{n+1}}{n+1}$	...

Now that we have defined the notion of anti-derivative of a continuous function, there are two questions that arise spontaneously:

- Does an anti-derivative for a continuous function $f\colon [a,b] \to \mathbb{R}$ over a closed bounded interval $[a,b]$ always exist?
- If an anti-derivative exists for a continous function $f\colon [a,b] \to \mathbb{R}$, does it help us in any way in computing the value of the integral $\int_a^b f(x)dx$?

These two questions have some simple but very powerful answers provided by the following two theorems, that are usually called the first and second fundamental theorems of calculus.

Theorem 3.23. $\boxed{\text{Fundamental theorem of calculus I}}$
Let $f : [a,b] \to \mathbb{R}$ be continuous. Then,

$$F(x) := \int_a^x f(t)dt$$

is an anti-derivative of f.

Theorem 3.24. $\boxed{\text{Fundamental theorem of calculus II}}$
Let $f : [a,b] \to \mathbb{R}$ be continuous and let G be an anti-derivative of f. Then,

$$\int_a^b f(x)dx = G(b) - G(a).$$

Notation 3.25. In order for the statement of Theorem 3.23 to make full sense, we need to introduce some further notation: so far we defined $\int_a^b f(x)dx$ only for $a < b$. If $a = b$, then we define

$$\int_a^a f(x)dx := 0.$$

If $a > b$, then we also define

$$\int_a^b f(x) := -\int_b^a f(x)dx.$$

With these notations our previously proven rules give that if $f : [a,b] \to \mathbb{R}$ is continuous, and $c, d \in [a,b]$ are any points, then

$$\int_a^c f(x)dx + \int_c^d f(x)dx = \int_a^d f(x)dx.$$

The next statement is not too interesting in itself but it is needed in the proof of Theorem 3.24.

Theorem 3.26. $\boxed{\text{Mean Value Theorem for Integrals}}$
If $f : [a,b] \to \mathbb{R}$ is continuous, then there is a $c \in [a,b]$, such that

$$\int_a^b f(x)dx = f(c)(b-a).$$

Proof. As $[a,b]$ is closed and f is continuous, by Theorem 1.63, f admits global maximum and minimum over $[a,b]$. Set $M := \max_{x\in[a,b]} f(x)$ and $m := \min_{x\in[a,b]} f(x)$. By Theorem 1.66, f takes all values in $[m,M]$. However, by Proposition 3.6, then

$$m \leq \frac{\int_a^b f(x)dx}{b-a} \leq M,$$

so there is a $c \in [a,b]$ such that $f(c)$ equals the above fraction, which is exactly the statement of the theorem. □

Let us now give the proofs of the two fundamental theorems of calculus.

Proof of Theorem 3.23. Fix $x_0 \in]a,b[$. Then, for any $x_0 \neq x \in]a,b[$:

$$\frac{F(x)-F(x_0)}{x-x_0} = \frac{1}{x-x_0}\int_{x_0}^{x} f(t)dt = \underbrace{f(c(x))}_{\text{Theorem 3.26}},$$

for a real number $c(x)$ between x and x_0. Hence:

$$\lim_{x\to x_0}\frac{F(x)-F(x_0)}{x-x_0} = \lim_{x\to x_0} f(c(x)) = \underbrace{\lim_{x\to x_0} f(x)}_{\lim_{x\to x_0} c(x)=x_0} = \underbrace{f(x_0)}_{f \text{ is continuous}}.$$

□

Proof of Theorem 3.24. We have already shown in Theorem 3.23 that $F(x) = \int_a^x f(t)dt$ is an anti-derivative of f. As both F and G are anti-derivatives, they differ by a constant $C \in \mathbb{R}$, that is, $(F-G)(x) = C$, $\forall x \in [a,b]$. Then:

$$\begin{aligned} G(b)-G(a) =& (G(b)+c)-(G(a)+c) = F(b)-F(a) \\ =& \int_a^b f(x)dx - \int_a^a f(x)dx = \int_a^b f(x)dx. \end{aligned}$$

□

Notation 3.27. The expression $G(b)-G(a)$ appearing in the statement of Theorem 3.24 is usually denoted by

$$G(x)|_a^b \qquad \text{or} \qquad G(x)|_{x=a}^{x=b}.$$

Example 3.28.

$$\int_{-5}^{-1}\frac{1}{x} = \big(\log|x|\big)\big|_{x=-5}^{x=-1} = \log 1 - \log 5 = -\log 5$$

3.4 Substitution

Theorem 3.29. *Let $f : [a,b] \to \mathbb{R}$ be a continuous function, and let $\phi : [\alpha,\beta] \to [a,b]$ be a C^1 function. Then,*

$$\int_{\phi(\alpha)}^{\phi(\beta)} f(x)dx = \int_\alpha^\beta f(\phi(t))\phi'(t)dt. \tag{3.29.a}$$

Proof. Define $G(x) := \int_a^x f(u)du$. By Theorem 3.23, G is an anti-derivative of f, so that Theorem 3.24 tells us

$$\int_{\phi(\alpha)}^{\phi(\beta)} f(x)dx = G(\phi(\beta)) - G(\phi(\alpha)).$$

So, it is enough to show that the value of the right side of (3.29.a) is the same. To show that, let us just note that by the chain rule $G(\phi(t))' = G'(\phi(t))\phi'(t) = f(\phi(t))\phi'(t)$. Then applying Theorem 3.24 to the integral $\int_\alpha^\beta f(\phi(t))\phi'(t)dt$ implies that

$$\int_\alpha^\beta f(\phi(t))\phi'(t)dt = G(\phi(\beta)) - G(\phi(\alpha)),$$

which concludes the proof. $\square$

Example 3.30. In this example, we go from the right side of (3.29.a) to the left side.

$$\int_0^1 \sqrt{e^x}e^x dx = \underbrace{\int_1^e \sqrt{u}du}_{u=e^x \quad (e^x)'=e^x} = \left.\frac{u^{\frac{3}{2}}}{\frac{3}{2}}\right|_{u=1}^{u=e} = \frac{2}{3}\left(e^{\frac{3}{2}} - 1\right)$$

Example 3.31. Let us consider the function $f\colon [0,1] \to \mathbb{R}$, $f(x) := \sqrt{1-x^2}$.
We want to compute the integral $\int_0^1 f(x)dx$. This integral computes the area of a quarter

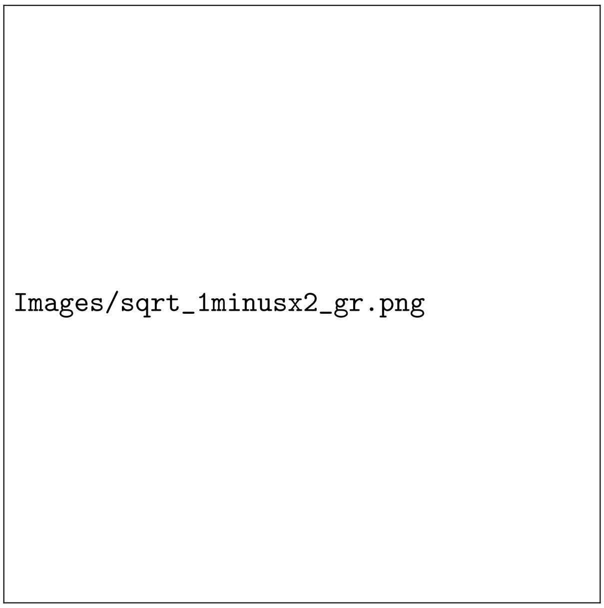

Figure 17: $f(x) := \sqrt{1-x^2}$.

of a circle of radius 1, as shown by Figure 17 so the result should be $\frac{\pi}{4}$. Indeed, the above computation shows that are train of thought is correct. Note that, opposite to the previous example, in this argument at our first substitution we go from the left side of (3.29.a) to the right side.

$$\int_0^1 \sqrt{1-x^2}dx = \underbrace{\int_0^{\frac{\pi}{2}} \sqrt{1-(\sin(t))^2}\cos(t)dt}_{x=\sin(t) \qquad \sin(t)'=\cos(t)} = \int_0^{\frac{\pi}{2}} \sqrt{\cos(t)^2}\cos(t)dt$$

$$= \int_0^{\frac{\pi}{2}} |\cos(t)|\cos(t)dt = \underbrace{\int_0^{\frac{\pi}{2}} \cos(t)\cos(t)dt}_{t\in\left[0,\frac{\pi}{2}\right] \Rightarrow \cos(t)\geq 0 \Rightarrow |\cos(t)|=\cos(t)}$$

$$= \int_0^{\frac{\pi}{2}} \frac{\cos(2t)+1}{2}dt = \underbrace{\int_0^{\pi} \left(\frac{\cos(u)+1}{2}\right)\frac{1}{2}du}_{t=\frac{u}{2}}$$

$$= \frac{1}{4}\int_0^{\pi} (\cos(u)+1)du = \frac{1}{4}\left(\sin(u)+u\right)|_{u=0}^{u=\pi}$$

$$= \frac{1}{4}(\sin(\pi)+\pi-\sin(0)-0) = \frac{\pi}{4}$$

Example 3.32. Recall that $\sinh(x) : \mathbb{R} \to \mathbb{R}$ is an odd function and it is strictly increasing (indeed, $\sinh(x)' = \cosh(x) > 0$). In particular, it has an inverse, which we denote by $\sinh^{-1}(x)\colon \mathbb{R} \to \mathbb{R}$. With this we may compute similarly :

$$\int_0^1 \sqrt{1+x^2}dx = \underbrace{\int_0^{\sinh^{-1}(1)} \sqrt{1+(\sinh(t))^2}\cosh(t)dt}_{x=\sinh(t) \qquad \sinh(t)'=\cosh(t)} = \int_0^{\sinh^{-1}(1)} \sqrt{\cosh(t)^2}\cosh(t)dt$$

$$= \int_0^{\sinh^{-1}(1)} |\cosh(t)^2|\cosh(t)dt = \underbrace{\int_0^{\sinh^{-1}(1)} \cosh(t)\cosh(t)dt}_{\cosh(t)>0 \Rightarrow |\cosh(t)|=\cosh(t)} = \int_0^{\sinh^{-1}(1)} \frac{\cosh(2t)+1}{2}dt$$

$$= \underbrace{\int_0^{2\sinh^{-1}(1)} \frac{\cosh(u)+1}{2}\frac{1}{2}du}_{t=\frac{u}{2}} = \frac{1}{4}\int_0^{2\sinh^{-1}(1)} \cosh(u)+1du = \frac{1}{4}\left(\sinh(u)+u\right)|_{u=0}^{u=2\sinh^{-1}(1)}$$

$$= \frac{\sinh(2\sinh^{-1}(1))+2\sinh^{-1}(1)}{4} = \frac{2\sinh(\sinh^{-1}(1))\cosh(\sinh^{-1}(1))+2\sinh^{-1}(1)}{4}$$

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

$$= \frac{2\sqrt{2}+2\sinh^{-1}(1)}{4}$$

Example 3.33. Substitution can be used the generally integrate $\cos(x)^n$ and $\sin(x)^n$ when n is a positive integer.

(1) The simplest case is when n odd. Here is an example of that:

$$\int_0^{\frac{\pi}{2}} \cos(x)^5 dx = \int_0^{\frac{\pi}{2}} \cos(x)(1-\sin(x)^2)^2 dx = \underbrace{\int_0^1 (1-u^2)^2 du}_{u(x)=\sin(x) \quad u(x)'=\cos(x)}$$
$$= \int_0^1 (1-2u^2+u^4)du = \left(u - \frac{2u^3}{3} + \frac{u^5}{5}\right)\Big|_{u=0}^{u=1}$$
$$=1 - \frac{2}{3} + \frac{1}{5} = \frac{6}{15} = \frac{2}{5}.$$

(2) On the other hand, when n is even, by reverse-engineering duplication formulas for sine and cosine, we can reduce again to the case of an odd power:

$$\begin{aligned}\int_0^{\frac{\pi}{2}} \sin^4(x)dx &= \int_0^{\frac{\pi}{2}} \left(\frac{1-\cos(2x)}{2}\right)^2 dx = \int_0^{\frac{\pi}{2}} \frac{1}{4} - \frac{\cos(2x)}{2} + \frac{\cos(2x)^2}{4} dx \\ &= \int_0^{\frac{\pi}{2}} \frac{1}{4}dx - \int_0^{\frac{\pi}{2}} \frac{\cos(2x)}{2}dx + \int_0^{\frac{\pi}{2}} \frac{\cos(2x)^2}{4}dx \\ &= \left(\frac{\pi}{8} - \frac{\sin(2x)}{4}\right)\Big|_{x=0}^{x=\frac{\pi}{2}} + \int_0^{\frac{\pi}{2}} \frac{\cos(4x)+1}{8}dx \\ &= \left(\frac{\pi}{8} + \frac{\pi}{16} + \frac{\sin(4x)}{32}\right)\Big|_{x=0}^{x=\frac{\pi}{2}} = \frac{\pi}{8} + \frac{\pi}{16} = \frac{3\pi}{16}\end{aligned} \tag{3.33.b}$$

3.5 Integration by parts

Theorem 3.34. *If $f, g\colon : E \to \mathbb{R}$ be two C^1 functions on an open interval E, and let $a < b$ be elements of E. Then,*

$$\int_a^b f(x)g'(x)dx = f(x)g(x)|_a^b - \int_a^b f'(x)g(x)dx, \tag{3.34.a}$$

Proof. It is enough to show that

$$\int_a^b (f(x)g'(x) + f'(x)g(x))dx = f(x)g(x)|_a^b.$$

This follows immediately from the Leibniz formula, Theorem 3.24,

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

□

Generally integration by parts are useful for products. The main question in applying (3.34.a) is how one chooses f and g. There is a rule which works in most cases (but not always!). Using the list below, when you encounter a product of two functions both of which belong to one of the categories in the list, the idea is that you should assign g' in the formula (3.34.a) to the function that belongs to the category appearing earlier in the list.

- E(xponential)
- T(rigonometric)
- A(lgebraic, that is, polynomial)
- L(ogarithm)

∘ I(nverse trigonometric).

We present now a few examples.

Example 3.35.

$$
\begin{aligned}
\underbrace{\int_a^b xe^x dx}_{g'(x)=e^x \quad g(x)=e^x \quad f(x)=x \quad f'(x)=1} &= (xe^x)|_{x=a}^b - \int_a^b e^x dx \\
&= (xe^x - e^x)|_{x=a}^b = ((x-1)e^x)|_{x=a}^b
\end{aligned}
$$

Example 3.36.

$$
\begin{aligned}
\underbrace{\int_a^b \sin(x)e^x dx}_{g'(x)=e^x, g(x)=e^x, f(x)=\sin(x), f'(x)=\cos(x)} &= (\sin(x)e^x)|_{x=a}^b - \underbrace{\int_a^b \cos(x)e^x dx}_{g'(x)=e^x, g(x)=e^x, f(x)=\cos(x), f'(x)=-\sin(x)} \\
&= (\sin(x)e^x - \cos(x)e^x)|_{x=a}^b + \int_a^b (-\sin(x))e^x dx
\end{aligned}
$$

Thus, we can rewrite the equality between the LHS of the first line and the second line as

$$
\int_a^b \sin(x)e^x dx = \frac{(e^x(\sin(x) - \cos(x)))|_{x=a}^b}{2}.
$$

Example 3.37.

$$
\underbrace{\int_a^b \log(x) dx}_{f(x)=\log(x), f'(x)=\frac{1}{x}, g'(x)=1, g(x)=x} = (x\log(x))|_{x=a}^b - \int_a^b 1dx = (x\log(x) - x)|_{x=a}^b
$$

Example 3.38.

$$
\begin{aligned}
\underbrace{\int_a^b \arctan(x) dx}_{f(x)=\arctan(x), f'(x)=\frac{1}{1+x^2}, g'(x)=1, g(x)=x} &= (x\arctan(x))|_{x=a}^b - \int_a^b \frac{x}{1+x^2} dx \\
&= (x\arctan(x))|_{x=a}^b - \underbrace{\frac{1}{2}\int \frac{1}{u} du}_{u(x)=1+x^2, u(x)'=2x} \\
&= \left(x\arctan(x) - \frac{1}{2}\log|u| \right)\Big|_{x=a}^b \\
&= \left(x\arctan(x) - \frac{1}{2}\log|1+x^2| \right)\Big|_{x=a}^b
\end{aligned}
$$

3.6 Integrating rational functions

A *rational function*, is a function of the form $\frac{P(x)}{Q(x)}$, where $P(x)$ and $Q(x)$ are polynomials with real coefficients.

We start by recalling how polynomials with real coefficients can be factorized.

Theorem 3.39 (Fundamental theorem of algebra over $\mathbb{R}$)**.** *Let $Q(x)$ be a polynomial with real coefficients. Then, $Q(x)$ can be factored as*

$$
Q(x) = (x-a_1)^{k_1} \dots (x-a_n)^{k_n} (x^2 + 2b_1 x + c_1)^{l_1} \dots (x^2 + 2b_m x + c_m)^{l_m}, \tag{3.39.a}
$$

where the a_i, b_i and c_i are real numbers, $k_i, l_i > 0$ are positive integers, and the quadratic polynomials $x^2 + 2b_i x + c_i$, $1 \le i \le m$ are irreducible, that is, there is no real number x_0 such that $x_0^2 + 2b_i x_0 + c_i = 0$, $1 \le i \le m$.

Remark 3.40. The Fundamental Theorem of Algebra is originally for polynomials $R(x)$ with complex coefficients. For those polynomial, the statement is even better: namely, polynomials with complex coefficients can be factored into linear terms. That is,

$$R(x) = (x - d_1)^{s_1} \dots \left(x - d_n'\right)^{s_n}, \quad d_i \in \mathbb{C}, \; s_i \in \mathbb{N}^*. \tag{3.40.b}$$

This does work also for $R(x) := Q(x)$ a real polynomial,as $\mathbb{R} \subset \mathbb{C}$, however it may happen that some of the d_i that are complex and not real. Then, the expression cannot be used for integration because we did not learn integration of complex valued functions. Hence, the idea is to collect the d_i that are real numbers. These numbers provide the a_i in (3.39.a). Furthermore, as we are working with a polynomial $R(x)$ with real coefficients,then (3.40.b) is invariant under conjugation, since $R(x)$ is. Hence, whenever d_i is not real, then also the conjugate of $(x - d_i)$ has to show up in (3.40.b) with the same power. That is, we have a factor of the right hand side of (3.40.b) of the form:

$$\begin{aligned}(x - d_i)^{s_i} \left(x - \overline{d_i}\right)^{s_i} &= \left((x - d_i)\left(x - \overline{d_i}\right)\right)^{s_i} \\ &= \left(x - 2\left(d_i + \overline{d_i}\right) + d_i\overline{d_i}\right)^{s_i} = \left(x - 2\operatorname{Re}(d_i) + |d_i|^2\right)^{s_i}\end{aligned}$$

Then, setting $b_j := -\operatorname{Re}(d_i)$, $l_j := s_i$ and $c_j := |b_i|^2$, we obtain one of the terms of the form $(x^2 + 2b_jx + c_j)^{l_j}$ in (3.39.a).

Example 3.41. Take $Q(x) = x^3 + x^2 - 2$, and consider the factorization as in (3.39.a). As the degree of Q is three, there must be a linear term (the product of the non-linear terms has even degree). This correspond to a real root of $Q(x)$, so let us search for it.

(1) *Finding the real root.*

1st try: $Q(0) = -2$. As $\lim_{x \to +\infty} Q(x) = +\infty$, the Intermediate value theorem, Theorem 1.66, implies that Q has a root greater than 0.

2nd try: $Q(1) = 0$. We found the root, great.

(2) *Factoring out the linear term.*

$$\begin{array}{ccccccc|c|c}
x^3 & + & x^2 & & & - & 2 & x-1 & x^2+2x+2 \\ \cline{8-9}
x^3 & - & x^2 & & & & & & \\ \cline{1-7}
 & & 2x^2 & & & - & 2 & & \\
 & & 2x^2 & - & 2x & & & & \\ \cline{1-7}
 & & & & 2x & - & 2 & & \\
 & & & & 2x & - & 2 & & \\ \cline{1-7}
 & & & & & & 0 & &
\end{array}$$

Hence, we have

$$(x^2 + 2x + 2)(x - 1) = x^3 + x^2 - 2$$

Remark 3.42. Unfortunately, for polynomials of degree ≥ 5 there is no algorithm for finding the roots; one just has to try to use the Intermediate Value theorem, hoping that the given polynomial yields a nice root.

Using (3.39.a), we have the following nice factorization of rational functions.

Proposition 3.43. *Any rational function* $\frac{P(x)}{Q(x)}$ *can be written as*

$$\frac{P(x)}{Q(x)} = \alpha_1 R_1(x) + \dots + \alpha_t R_t(x),$$

where the α_i *are real numbers, and* $R_i(x)$ *are of the form:*

(1) polynomial, or

(2) $\frac{1}{(x-r)^p}$*, or*

(3) $\frac{x+c}{(x^2+2rx+s)^p}$.

Instead of giving a proof, we explain the idea behind Proposition 3.43 in the following example.

Example 3.44. Given the rational function $\frac{4x^3+9x^2+11x+8}{(x^2+x+1)^2}$, let us try to factorize it as follows:

$$\begin{aligned}\frac{4x^3+9x^2+11x+8}{(x^2+x+1)^2} &= \frac{Ax+B}{(x^2+x+1)^2} + \frac{Cx+D}{x^2+x+1} \\ &= \frac{Ax+B+(Cx+D)(x^2+x+1)}{(x^2+x+1)^2} \\ &= \frac{Cx^3+(C+D)x^2+(A+C+D)x+(B+D)}{(x^2+x+1)^2},\end{aligned}$$

which yields the following linear system

$$\begin{cases} C &= 4 \\ C+D &= 9 \\ A+C+D &= 11 \\ B+D &= 8 \end{cases}$$

for which the solutions are

$$\begin{aligned} C = 4 \Rightarrow 4 + D = 9 \Rightarrow D = 5 \\ \Rightarrow A + 4 + 5 = 11; B + 5 = 8 \\ \Rightarrow A = 2; B = 3 \end{aligned}$$

Thus,

$$\frac{4x^3+9x^2+11x+8}{(x^2+x+1)^2} = \frac{2x+3}{(x^2+x+1)^2} + \frac{4x+5}{x^2+x+1}.$$

Having the decomposition stated in Proposition 3.43, the question is how we integrate these terms separately.

Example 3.45.

- For $p > 1$, then

$$\int \frac{1}{(x-r)^p}dx = \frac{(x-r)^{1-p}}{1-p}.$$

- For $p = 1$, then

$$\int \frac{1}{(x-r)}dx = \log|x-r|.$$

Example 3.46.

$$\begin{aligned}\int \frac{x+c}{(x^2+2rx+s)^p} &= \frac{1}{2}\int \frac{2(x+r)}{(x^2+2rx+s)^p}dx + \int \frac{c-r}{(x^2+2rx+s)^p}dx \\ &= \frac{1}{2}\int \frac{2(x+r)}{(x^2+2rx+s)^p}dx + (c-r)\int \frac{1}{(x^2+2rx+s)^p}dx.\end{aligned}$$

Hence, we need to compute the two integrals

$$\int \frac{2(x+r)}{(x^2+2rx+s)^p}dx, \quad \int \frac{1}{(x^2+2rx+s)^p}dx,$$

individually.

- Using the substitution $u = x^2 + 2rx + s$, then

$$\int \frac{2(x+r)}{(x^2+2rx+s)^p}dx = \begin{cases} \log|x^2+2rx+s| & \text{if } p = 1 \\ \frac{(x^2+2rx+s)^{1-p}}{1-p} & \text{if } p > 1 \end{cases}$$

- Hence, we now know how to integrate all the terms in Proposition 3.43, except for the integral

$$\int \frac{1}{(x^2+2rx+s)^p}dx, \quad p > 0.$$

  Hence,

$$\begin{aligned}\int \frac{1}{(x^2+2rx+s)^p}dx &= \int \frac{1}{((x+r)^2+(s-r^2))^p}dx \\ &= \underbrace{\frac{1}{(s-r^2)^p}\int \frac{1}{\left(\left(\frac{x+r}{\sqrt{s-r^2}}\right)^2+1\right)^p}dx}_{s-r^2>0, \text{ as } x^2+2rx+s \text{ has no real roots}} \\ &= \underbrace{\frac{1}{(s-r^2)^{p-\frac{1}{2}}}\int \frac{1}{(u^2+1)^p}du,}_{u=\frac{x+r}{\sqrt{s-r^2}}}\end{aligned}$$

  and we are left to compute the integral

$$\int \frac{1}{(u^2+1)^p}du, \quad \text{for } p > 0.$$

  Setting

$$I_p := \int \frac{1}{(u^2+1)^p}du,$$

  then $I_1 := \arctan(u)$ and furthermore, if $p \geq 1$, then we obtain a recursive formula as follows:

$$I_p = \int \frac{1}{(u^2+1)^p}du = \underbrace{\frac{u}{(u^2+1)^p} - \int \frac{(-p)u \cdot 2u}{(u^2+1)^{p+1}}du}_{\text{integrating by parts with } f(u)=\frac{1}{(u^2+1)^p}, g'(u)=1}$$

$$= \frac{u}{(u^2+1)^p} + 2p\int \frac{u^2+1-1}{(u^2+1)^{p+1}}du = \frac{u}{(u^2+1)^p} + 2pI_p - 2pI_{p+1}$$

  So, by looking at the two ends of the equation, we obtain the recursive equality:

$$I_{p+1} = \frac{\frac{u}{(u^2+1)^p} + (2p-1)I_p}{2p}.$$

Remark 3.47. Be careful, there is an error on page 201 of the book where they prove the formulas above: intead of $2p-1$, they wrote $2(p-1)$!!

Example 3.48. Let us compute I_2 for example:

$$I_2 = \frac{\frac{u}{u^2+1} + I_1}{2} = \frac{1}{2}\left(\frac{u}{u^2+1} + \arctan(u)\right)$$

Example 3.49. Let us get back to Example 3.44:

$$\begin{aligned}
\int \frac{4x^3+9x^2+11x+8}{(x^2+x+1)^2}dx &= \underbrace{\int \frac{2x+3}{(x^2+x+1)^2}dx + \int \frac{4x+5}{x^2+x+1}dx}_{\text{by Example 3.44}} \\
&= \underbrace{\int \frac{2x+3}{\left(\left(x+\frac{1}{2}\right)^2+\frac{3}{4}\right)^2}dx + \int \frac{4x+5}{\left(x+\frac{1}{2}\right)^2+\frac{3}{4}}dx}_{\text{completing the square in the denominators}} \\
&= \underbrace{\int \frac{(2x+1)+2}{\left(\left(x+\frac{1}{2}\right)^2+\frac{3}{4}\right)^2}dx + \int \frac{(4x+2)+3}{\left(x+\frac{1}{2}\right)^2+\frac{3}{4}}dx}_{\text{writing the numerators in terms of a multiple of } x+\frac{1}{2}}
\end{aligned}$$

$$= \underbrace{\int \left(\frac{2}{\sqrt{3}}\right)^3 \frac{\frac{2}{\sqrt{3}}(2x+1)+\frac{4}{\sqrt{3}}}{\left(\left(\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)\right)^2+1\right)^2}dx + \int \frac{2}{\sqrt{3}} \frac{\frac{2}{\sqrt{3}}(4x+2)+\frac{6}{\sqrt{3}}}{\left(\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)\right)^2+1}dx}_{\substack{\text{We multiply the numerators and the denominators by adequate multiples of } \frac{2}{\sqrt{3}}, \\ \text{to make them of the form } u^2+1 \text{ or } (u^2+1)^2}}$$

$$= \underbrace{\frac{4}{3}\int \frac{2u+\frac{4}{\sqrt{3}}}{(u^2+1)^2}du + \int \frac{4u+\frac{6}{\sqrt{3}}}{u^2+1}du}_{u=\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right) \quad \Rightarrow \quad x=\frac{\sqrt{3}}{2}u-\frac{1}{2} \quad \Rightarrow \quad x(u)'=\frac{\sqrt{3}}{2}}$$

$$= \frac{4}{3}\int \frac{2u}{(u^2+1)^2}du + \frac{16}{3\sqrt{3}}\int \frac{1}{(u^2+1)^2}du$$

$$+2\int \frac{2u}{u^2+1}dx + \frac{6}{\sqrt{3}}\int \frac{1}{u^2+1}du$$

$$= \frac{4}{3}\frac{-1}{u^2+1} + \frac{8}{3\sqrt{3}}\left(\frac{u}{u^2+1}+\arctan(u)\right) + 2\log|u^2+1| + \frac{6}{\sqrt{3}}\arctan(u)$$

$$= \frac{4}{3}\frac{-1}{\left(\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)\right)^2+1} + \frac{8}{3\sqrt{3}}\left(\frac{\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)}{\left(\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)\right)^2+1} + \arctan\left(\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)\right)\right)$$

$$+2\log\left|\left(\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)\right)^2+1\right| + \frac{6}{\sqrt{3}}\arctan\left(\frac{2}{\sqrt{3}}\left(x+\frac{1}{2}\right)\right)$$

3.6.1 Rational functions in exponentials

There is a method of integrating functions obtained by plugging in e^x into a rational function. We explain it via the next example:

Example 3.50.

$$\begin{aligned}\int \frac{1}{e^x+1}dx = \int \frac{1}{(e^x+1)e^x}e^x dx = &\underbrace{\int \frac{1}{(t+1)t}dt}_{t(x)=e^x \quad t(x)'=e^x} \\ &= \int \left(\frac{1}{t}-\frac{1}{t+1}\right)dt = \log|t| - \log|t+1| \\ &= \log|e^x| - \log|e^x+1| = x - \log|e^x+1|\end{aligned}$$

3.6.2 Rational functions in roots

There is a method of integrating functions obtained by plugging in $\sqrt{x}$ into a rational function. We explain it via the next example.

Example 3.51.

$$\begin{aligned}\int \frac{1}{\sqrt{x}+1}dx &= \int \left(\frac{1}{\sqrt{x}+1}2\sqrt{x}\right)\frac{1}{2}\frac{1}{\sqrt{x}}dx \\ &= \underbrace{\int \frac{2t}{t+1}dt}_{t(x)=\sqrt{x} \quad t(x)'=\frac{1}{2}\frac{1}{\sqrt{x}}} \quad = \int 2dt - \int \frac{2}{t+1}dt \\ &= 2t - 2\log|t+1| = 2\sqrt{x} - 2\log|\sqrt{x}+1|.\end{aligned}$$

3.7 Improper integrals

So far we have studied integrals of functions (mostly continuous ones) that are defined over a closed bounded interval.

But how can we make sense of integrating a function over an unbounded interval, for example, $\int_1^{+\infty} \frac{1}{x^2}dx$? Or more generally, we have a real valued function f that is continuous on an interval I of the form $[a,b[$, $]a,b]$ or $]a,b[$, where $a,b \in \overline{\mathbb{R}}$, but either f does not extend continuously to $[a,b]$, or the interval $[a,b]$ does not exist at all – such as in the case when a or b are $\pm\infty$.

Definition 3.52. Let $f: I \to \mathbb{R}$ be a continuous function.

(1) If $I = [a,b[$, $a \in \mathbb{R}$, and either $b \in \mathbb{R}$ or $b = +\infty$, then we define the *improper integral* of f on I to be the limit

$$\int_a^{b-} f(t)dt := \lim_{x \to b-} \left(\int_a^x f(t)dt\right),$$

provided that the above limits exist.

(2) If $I =]a,b]$, $b \in \mathbb{R}$, and either $a \in \mathbb{R}$ or $a = -\infty\infty$, then we define the *improper integral* of f on I to be the limit

$$\int_{a^+}^{b} f(t)dt := \lim_{x \to a+} \left(\int_x^b f(t)dt\right),$$

provided that the above limits exist.

(3) If $I =]a,b[$, $a,b \in \overline{\mathbb{R}}$, then we define the *improper integral* of f on I to be the limit

$$\int_{a^+}^{b^-} f(t)dt := \int_{a^+}^{c} f(t)dt + \int_c^{b-} f(t)dt$$

for any chosen $c \in I$, provided that the improper integrals

$$\int_{a^+}^{c} f(t)dt, \quad \int_c^{b-} f(t)dt$$

exist.

If the limits above exist and are finite, then we say that the improper integrals converge. If the above limits diverge, we say that the corresponding improper integral is divergent.

Remark 3.53. (1) It is an easy exercise to verify that part (3) of the above the definition does not depend on the choice of $c \in I$.

(2) By abuse of notation many times the $+$ and the $-$ is forgotten from the lower and upper limits.

Example 3.54.

$$\int_{0^+}^{1} \frac{1}{\sqrt{t}} dt = \lim_{x \to 0^+} 2t^{\frac{1}{2}}\Big|_{t=x}^{t=1} = \lim_{x \to 0^+} 2 - 2\sqrt{x} = 2$$

Example 3.55.

$$\int_{0^+}^{1} \frac{1}{t} dt = \lim_{x \to 0^+} \log(t)|_{t=x}^{t=1} = \lim_{x \to 0^+} -\log(x) = +\infty$$

So, $\int_{0^+}^{1} \frac{1}{t} dt$ is divergent.

Example 3.56.

$$\int_{0^+}^{1} \log(t) dt = \lim_{x \to 0^+} (\log(t)t - t)|_{t=x}^{t=1} = -1 - \lim_{x \to 0^+} (\log(x)x - x) = -1 - \lim_{x \to 0^+} (\log(x)x)$$

Here, we may compute $\lim_{x \to 0^+}(\log(x)x)$ using L'Hospital's rule:

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

Hence, $\int_{0^+}^{1} \log(t) dt = -1$.

Improper integrals enjoy many of the basic features of standard integrals.

Proposition 3.57. *Let $f, g \colon I \to \mathbb{R}$ be continuous functions defined over an interval I, where I is either one of the following intervals*

$$[a, b'[,]a', b],]a', b'[, \qquad a, b \in \mathbb{R}, \quad a', b' \in \overline{\mathbb{R}}.$$

Then,

(1) If $I = [a, b[$ (resp. $I =]a, b]$, $I =]a, b[$), $a, b \in \mathbb{R}$ and f extends to a continuous function defined over the interval $[a, b]$, then the improper interval

$$\int_{a}^{b-} f(x)dx \text{ (resp. } \int_{a+}^{b} f(x)dx, \ \int_{a+}^{b-} f(x)dx).$$

converges and it is equal to $\displaystyle\int_a^b f(x)dx$.

(2) If $c \in I$ and $c \neq \sup I, \inf I$ then if the improper interval of f on I converges, we have that

$$\int_a^b f(x)dx + \int_a^c f(x)dx = \int_c^b f(x)dx.$$

(3) Given $\alpha, \beta \in \mathbb{R}$, if the improper integral of f, g over I converge, then also $\alpha f + \beta g$ is integrable on I, and

$$\int_a^b (\alpha f + \beta g)(x)dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx$$

(4) If $0 \le f \le g$, then

(i) *if the improper integral $\int_a^b g(x)dx$ converges then also the improper integral $\int_a^b f(x)dx$ does;*

(ii) *if the improper integral $\int_a^b f(x)dx$ diverges then also the improper integral $\int_a^b g(x)dx$ does.*

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx$$

Remark 3.58. In part (2-4) of the previous proposition, we have used the simplified notation for improper integrals that was introduced in Remark 3.53(2).

Definition 3.59. In the hypotheses of Definition 3.52, we say that an improper integral is absolutely convergent if the improper integral defined by integrating the function $|f|$ instead of f is convergent.

The following is an an analogue, for improper integrals, of **??**.

Proposition 3.60. *If an improper integral is absolutely convergent, then it is also convergent.*

Example 3.61. The backwards implication of Proposition 3.60 does not hold, as shown by the next example.

$$\int_{\frac{\pi}{4}}^{+\infty} \frac{\sin(t)}{t} dt = \underbrace{\lim_{x\to+\infty} \frac{-\cos(x)}{x} - \frac{-\cos\left(\frac{\pi}{4}\right)}{\frac{\pi}{4}} - \int_{\frac{\pi}{4}}^{+\infty} \frac{-\cos(t)}{-t^2} dt}_{g'=\sin(t) \quad f=\frac{1}{t} \quad g=-\cos(t) \quad f'=\frac{1}{-t^2}} = \frac{\sqrt{2}}{\pi} - \int_{\frac{\pi}{4}}^{+\infty} \frac{\cos(t)}{t^2} dt. \tag{3.61.a}$$

So, $\displaystyle\int_{\frac{\pi}{4}}^{+\infty} \frac{\sin(t)}{t} dt$ is convergent if so is $\displaystyle\int_{\frac{\pi}{4}}^{+\infty} \frac{\cos(t)}{t^2} dt$. However, the latter is convergent because it is absolute convergent:

$$\int_{\frac{\pi}{4}}^{+\infty} \left|\frac{\cos(t)}{t^2}\right| dt \leq \int_{\frac{\pi}{4}}^{+\infty} \frac{1}{t^2} dt = \lim_{x\to+\infty} \left(\frac{-1}{t} \Big|_{\frac{\pi}{4}}^{x} \right) = \frac{4}{\pi} + \lim_{x\to+\infty} \frac{1}{x} = \frac{4}{\pi}.$$

This yields that $\displaystyle\int_{\frac{\pi}{4}}^{+\infty} \frac{\sin(t)}{t} dt$ is convergent.

However, be careful, the fact that $\displaystyle\int_{\frac{\pi}{4}}^{+\infty} \frac{\cos(t)}{t^2} dt$ is absolute convergent, does not mean that so is $\displaystyle\int_{\frac{\pi}{4}}^{+\infty} \frac{\sin(t)}{t} dt$. That is, equation (3.61.a) does not work for $\frac{\sin(t)}{t}$ replaced by $\left|\frac{\sin(t)}{t}\right|$. And, in fact, $\displaystyle\int_{\frac{\pi}{4}}^{+\infty} \frac{\sin(t)}{t} dt$ is not absolute convergent, because

$$\int_{\frac{\pi}{4}}^{n\pi} \left|\frac{\sin(t)}{t}\right| dt \geq \sum_{k=1}^{n} \int_{k\pi-\frac{3\pi}{4}}^{k\pi-\frac{\pi}{4}} \left|\frac{\sin(t)}{t}\right| dt \geq \sum_{k=1}^{n} \frac{\pi}{2} \left(\min_{t\in\left[k\pi-\frac{3\pi}{4}, k\pi-\frac{\pi}{4}\right]} \frac{|\sin(t)|}{t} \right)$$

$$\geq \sum_{k=1}^{n} \frac{\pi}{2} \left(\frac{\min_{t\in\left[k\pi-\frac{3\pi}{4}, k\pi-\frac{\pi}{4}\right]} |\sin(t)|}{\max_{t\in\left[k\pi-\frac{3\pi}{4}, k\pi-\frac{\pi}{4}\right]} t} \right) = \sum_{k=1}^{n} \frac{\pi}{2} \frac{\frac{1}{\sqrt{2}}}{k-\frac{\pi}{4}} = \frac{\pi}{2\sqrt{2}} \sum_{k=1}^{n} \frac{1}{k}.$$

As $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k}$ is divergent, $\displaystyle\lim_{n\to\infty} \int_{\frac{\pi}{4}}^{n\pi} \left|\frac{\sin(t)}{t}\right| dt$ does not exist. Therefore, $\displaystyle\int_{\frac{\pi}{4}}^{\infty} \left|\frac{\sin(t)}{t}\right| dt$ is divergent.

Example 3.62. A typical application of improper integral is to give an upper bound on infinite sums. For example

$$\sum_{k=10}^{\infty} \frac{1}{k^2} \leq \int_9^{+\infty} \frac{1}{x^2} dx = \left.\frac{-1}{x}\right|_{x=9}^{x\to+\infty} = \left(\lim_{n\to\infty x\to+\infty} \frac{-1}{x}\right) - \frac{-1}{9} = \frac{1}{9}$$

4 Power series

Definition 4.1. Let $(a_k) \subset \mathbb{R}$ be a sequence of real numbers and let $x_0 \in \mathbb{R}$ be a fixed real number.
A *power series* centered at x_0 is an expression of the form

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k.$$

Remark 4.2. In the expression above, the coefficients a_k and the center of the power series x_0 are fixed real numbers (the ones we fixed at the start of the definition), while the variable x is free to vary.

As the value of the series $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ varies when $x \in \mathbb{R}$ varies, in particular, the series will converge for certain values of x and diverge for others. For example, if we take $x = x_0$, then the value of the series above is a_0. For those values of x for which $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ converges, we can then discuss the value of the series, and we will see below that we can construct a function out of it. Hence, in view of this observation, we are compelled to introduce the following definition.

Definition 4.3. The *domain of convergence* $D \subset \mathbb{R}$ of the power series $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ is the set

$$D := \left\{ x \in \mathbb{R} \,\middle|\, \sum_{k=0}^{\infty} a_k (x - x_0)^k \text{ is convergent} \right\}$$

As already noted above, for a power series $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ centered at $x_0 \in \mathbb{R}$, then $x_0 \in D$ always, as the only non-zero term in $\sum_{k=0}^{\infty} a_k (x_0 - x_0)^k$ is a_0. The following theorem illustrates that the domain of convergence of a power series $\sum_{k=0}^{\infty} a_k (x_0 - x_0)^k$ has a relatively simple form: namely, it is an interval.

Theorem 4.4. *Let* $\sum_{k=0}^{\infty} a_k (x_0 - x_0)^k$ *be a power series centered at* $x_0 \in \mathbb{R}$. *Then, there is a real number* $R \in \mathbb{R}_+$ *such that if*
$$\begin{cases} |x - x_0| < R, \\ \\ |x - x_0| > R, \end{cases} \quad \text{then} \quad \begin{cases} \sum_{k=0}^{\infty} a_k (x - x_0)^k \text{ is convergent} \\ \sum_{k=0}^{\infty} a_k (x - x_0)^k \text{ is divergent} \end{cases}.$$
The number R is called the radius of convergence *of the power series* $\sum_{k=0}^{\infty} a_k (x - x_0)^k$.

Proof. Let D be the domain of convergence of $\sum_{k=0}^{\infty} a_k (x - x_0)^k$. We have to show that if $x' \in D$, then for all x with $0 < |x - x_0| < |x' - x_0|$, $x \in D$.

So, let us assume that for a given $x' \neq x_0$, $\sum_{k=0}^{\infty} a_k(x'-x_0)^k$ is convergent. In particular, $\lim_{k\to\infty} a_k(x'-x_0)^k \to 0$, and hence the sequence $b_k := a_k(x'-x_0)^k$ is bounded, that is, there exists a non-negative real number B, such that $|b_k| \leq B$. However, setting $y := \frac{x-x_0}{x'-x_0}$, then $|y| < 1$, by our assumption, and

$$0 \leq \left|a_k(x-x_0)^k\right| = \left|b_k y^k\right| \leq B|y|^k.$$

Since $\sum_{k=0}^{\infty} B|y|^k$ is a geometric series and $|y| < 1$, then $\sum_{k=0}^{\infty} B|y|^k$ is convergent. The Squeeze Theorem for series, (**??**), then implies that $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ is absolutely convergent, and hence also convergent. We then define

$$R := \sup\{r \in \mathbb{R} \mid \forall x \in]x_0 - r, x_0 + r[\ , \sum_{k=0}^{\infty} a_k(x-x_0)^k \text{ converges absolutely}\}.$$

A similar argument to the one used in the first part of the the proof can be used to show the other part of the statement about the non-convergence of the power series for $x \in \mathbb{R}$, $|x - x_0| > R$. □

Remark 4.5. Let $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ be a power series and let R be its radius of convergence. If $R = +\infty$, then the domain of convergence D is $\mathbb{R}$. If $R = 0$, then $D = \{x_0\}$. If $R > 0, R \neq +\infty$, then Theorem 4.4 implies that $D \supseteq]x_0 - R, x_0 + R[$ and for any $x \in \mathbb{R}$ such that $|x - x_0| > R$ then $x \notin D$. Thus, $D\setminus\ ]x_0 - R, x_0 + R[\ \subseteq \{x_0 - R, x_0 + R\}$. One can prove that it cannot happen that both $x_0 - R, x_0 + R$ belong to the domain of convergence, cf. Example 4.7 for some examples.

Now that we know that the radius of convergence exists, the question is how we can compute it. In **??** we saw several criteria to prove the convergence of a series. In particular, it is not too hard to adapt Cauchy's and D'Alembert criteria, **??** and **??**, to compute the radius of convergence of a series.

Theorem 4.6. *Let $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ be a power series, with radius of convergence R.*

(1) If $l_1 := \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$ exists for some $l_1 \in \overline{\mathbb{R}}$, then $R = \frac{1}{l_1}$.

(2) If $l_2 := \lim_{n\to\infty} \sqrt[n]{|a_n|}$ exists for some $l_2 \in \overline{\mathbb{R}}$, then $R = \frac{1}{l_2}$.

The above formulas are to be understood with the following notation when $l_i = 0, +\infty$:

$$\text{if } \begin{cases} l_i = 0 \\ l_i = +\infty \end{cases} \quad \text{then } \begin{cases} R = +\infty \\ R = 0 \end{cases}.$$

Proof. Let $x \neq x_0$. We want to understand when $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ is convergent using the existence of the limits in the hypotheses of parts (1)-(2) of the statement.

(1) Let us use the quotient criterion:

$$\lim_{n\to\infty}\left|\frac{a_{n+1}(x-x_0)^{n+1}}{a_n(x-x_0)^n}\right| = \lim_{n\to\infty}\left|\frac{a_{n+1}(x-x_0)^{n+1}}{a_n(x-x_0)^n}\right| = |x-x_0|\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = |x-x_0|l$$

So, by the quotient criterion, that is, **??**, $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ is convergent if $|x-x_0| < \frac{1}{l}$ and it is divergent if $|x-x_0| > \frac{1}{l}$.

(2) We use the Alembert's criterion:

$$\lim_{n\to\infty}\sqrt[n]{|a_n(x-x_0)^n|} = \lim_{n\to\infty}|\sqrt[n]{a_n}|\,|x-x_0| = L|x-x_0|$$

So, by Alembert's criterion, that is, **??**, $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ is convergent if $|x-x_0| < \frac{1}{L}$ and it is divergent if $|x-x_0| > \frac{1}{L}$.

□

Example 4.7. (1) We have seen earlier that

$$e^x = \sum_{n=0}^{\infty}\frac{x^n}{n!}$$

is convergent for all $x \in \mathbb{R}$, which we can now verify at once:

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n} = \lim_{n\to\infty}\frac{\frac{1}{(n+1)!}}{\frac{1}{n!}} = \lim_{n\to\infty}\frac{1}{n+1} = 0$$

So, indeed, the radius of convergence of the above series is ∞ and the domain of converge is $\mathbb{R}$.

(2) According to Theorem 4.6, the radius of convergence of

$$\sum_{n=0}^{\infty} nx^n$$

is $\frac{1}{l}$, where

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

So, the radius of convergence is 1. In particular the domain of convergence D of the series satisfies

$$]-1,1[\subseteq D \subseteq [-1,1].$$

To determine D fully, we need to understand whether $1, -1$ belong to D. For $x = 1$, the power series becomes

$$\sum_{k=0}^{\infty} n$$

which is clearly divergent as the sequence $c_n := n$ does not converge to 0; while, for $x = -1$, the power series becomes

$$\sum_{k=0}^{\infty}(-1)^n n$$

which is divergent as the sequence of the partial sums $s_j := \sum_{k=0}^{j}(-1)^n$ satisfies

$$s_j = \begin{cases} -1 & \text{for } n \text{ odd,} \\ 1 & \text{for } n \text{ even,} \end{cases}$$

thus the series does not converge. Thus, the domain of convergence is $D =]-1, 1[$.

(3) According to Theorem 4.6, the radius of convergence of

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

is $\frac{1}{l}$, where

$$l = \lim_{n\to\infty} \sqrt[n]{\frac{1}{n}} = 1.$$

So, the radius of convergence is 1. In particular the domain of convergence D of the series satisfies

$$]-1, 1[\subseteq D \subseteq [-1, 1].$$

To determine D fully, we need to understand whether $1, -1$ belong to D. For $x = 1$, the power series becomes

$$\sum_{k=1}^{\infty} \frac{1}{n}$$

which is clearly divergent, being the harmonic series, cf. **??**; while, for $x = -1$, the power series becomes

$$\sum_{k=0}^{\infty} (-1)^n \frac{1}{n}$$

which is convergent by Leibniz's criterion **??**. Thus, the domain of convergence is $D = [-1, 1[$.

(4) According to Theorem 4.6, the radius of convergence of the series

$$\sum_{n=0}^{\infty} e^{n+1} x^n,$$

centered at 0, is $\frac{1}{l}$, where

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

So, the radius of convergence is $\frac{1}{e}$. In particular the domain of convergence D of the series satisfies

$$]-\frac{1}{e}, \frac{1}{e}[\subseteq D \subseteq [-\frac{1}{e}, \frac{1}{e}].$$

To determine D fully, we need to understand whether $\frac{1}{e}, -\frac{1}{e}$ belong to D. For $x = \frac{1}{e}$, the power series becomes

$$\sum_{k=0}^{\infty} e^{n+1} \left(\frac{1}{e}\right)^n = \sum_{k=0}^{\infty} e$$

which is clearly divergent as the constant sequence $c_n := e$ does not converge to 0. For $x = -\frac{1}{e}$, the power series becomes

$$\sum_{k=0}^{\infty}(-1)^n e$$

which is divergent for the same reason as above. Thus, the domain of convergence is $D =]-\frac{1}{e}, \frac{1}{e}[$.

Let $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ be a power series and assume that the radius of convergence $R > 0$. This means that the domain of convergence of the power series contains the open interval $]x_0 - R, x_0 + R[$, hence we can define a function

$$\begin{aligned} f\colon\ ]x_0 - R, x_0 + R[&\to \mathbb{R}, \\ x &\mapsto f(x) := \sum_{k=0}^{\infty} a_k(x-x_0)^k. \end{aligned}$$

We are interested in understanding what kind of regularity properties the function f possesses. Is it continuous? Is it differentiable? How many derivatives does it have? Are those continuous over the domain of convergence of the series? All these questions are answered by the following result.

Theorem 4.8. *Let $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ be a power series with radius of convergence $R > 0$. Let $f(x)\colon\ ]x_0 - R, x_0 + R[\to \mathbb{R}$ the function $f(x) := \sum_{k=0}^{\infty} a_k(x-x_0)^k$. Then, for every $x \in]x_0 - R, x_0 + R[$,*

$$f'(x) = \sum_{k=1}^{\infty} k a_k (x-x_0)^{k-1}, \tag{4.8.a}$$

$$\int_{x_0}^{x} f(t)dt = \sum_{k=0}^{\infty} \frac{a_k}{k+1}(x-x_0)^{k+1}. \tag{4.8.b}$$

Remark 4.9. Given a power series $\sum_{k=0}^{\infty} a_k(x-x_0)^k$ with radius of convergence $R > 0$, then one can show, using the same argument as in the proof of Theorem 4.6, that both power series

$$\sum_{k=1}^{\infty} k a_k (x-x_0)^{k-1}, \text{ and } \sum_{k=0}^{\infty} \frac{a_k}{k+1}(x-x_0)^{k+1}$$

have the same radius of convergence R. In particular, this implies that we can iteratively use Theorem 4.8 to compute the n-th derivative of f and this will be given by the power series

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1)(k-2)\cdots(k-n+2)(k-n+1)a_k(x-x_0)^{k-n}.$$

In particular, $f \in C^{\infty}(]x_0 - R, x_0 + R, \mathbb{R})$.

The proof of Theorem 4.8 is omitted.

Example 4.10. We consider the power series $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k$. Then, using either one of the criteria from Theorem 4.6, it is not hard to see that the radius of convergence of the series is 1. Thus, we have a well-defined associated function $f\colon]-1,1[\to \mathbb{R}$, $f(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k$ and using the above theorem,

$$f'(x) = \sum_{k=1}^{\infty} (-1)^{k+1} x^{k-1},$$
$$\int_0^x f(t)dt = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)} (x)^{k+1}.$$

4.1 Taylor series

In Section 2.5.3, we saw how to approximate a function $f\colon E \to \mathbb{R}$ around a point $x_0 \in E$ by means of the values of f and its derivatives at x_0. When the function $f \in C^{\infty}(E,\mathbb{R})$, then we may be tempted to take expansions to larger and larger order of f around x_0 and wonder how they may compare with each other when we let $n \to \infty$. To this end, we introduce the following definition.

Definition 4.11. Let $f\colon E \to \mathbb{R}$ be a function defined on an open interval E and let $x_0 \in E$. Assume that $f \in C^{\infty}(E,\mathbb{R})$. Then the Taylor series of f is

$$T_{f,x_0}(x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k.$$

Remark 4.12. The terms of the Taylor series $T_{f,x_0}(x)$ of f centered at x_0 are the same as those of the Taylor expansions of order n, for any $n \in \mathbb{N}$, of f centered at x_0, cf. Theorem 2.74.

Hence, given a function $f\colon E \to \mathbb{R}$, $f \in C^{\infty}(E,\mathbb{R})$, and $x_0 \in E$, we can consider the Taylor series $T_{f,x_0}(x)$ of f centered at x_0

$$T_{f,x_0}(x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k.$$

At this point, we may be tempted to wonder whether $f(x) = T_{f,x_0}(x)$, for x sufficiently close to x_0. Let us remind the reader that in Theorem 2.74 we saw how if f is differentiable n times on E, then

$$f(x) = \left(\sum_{i=0}^{n-1} \frac{f^{(i)}(x_0)}{i!}(x-x_0)^i\right) + f^{(n)}(c)\frac{(x-x_0)^n}{n!}, \tag{4.12.a}$$

for some $c \in]x_0, x[$ (resp. $c \in]x, x_0[$). Hence, the Taylor series $T_{f,x_0}(x)$ equals $f(x)$ if and only if the error term $f^{(n)}(c)\frac{(x-x_0)^n}{n!}$ included in (4.12.a) converges to 0 as $n \to \infty$. This yields the following proposition.

Proposition 4.13. *In the situation of Definition 4.11, if $f \in C^{\infty}(E,\mathbb{R})$, then $f(x) = T_f(x)$ for a value $x \in E$, $x > x_0$ (resp. $x < x_0$) if*

$$\lim_{n\to\infty}\left(\sup_{y\in]x_0,x[}\left|f^{(n)}(y)\right|\right)\frac{|x-x_0|^n}{n!} = 0$$
$$(\textit{resp.} \lim_{n\to\infty}\left(\sup_{y\in]x,x_0[}\left|f^{(n)}(y)\right|\right)\frac{|x-x_0|^n}{n!} = 0).$$

Remark 4.14. Sometimes the Taylor series $T_{f,x_0}(x)$ of a a function f equals the original function f only in the center point x_0. A famous example is the following: let $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{for } x \neq 0 \\ 0 & x = 0 \end{cases}$$

We claim that $f \in C^\infty(\mathbb{R}, \mathbb{R})$, and $f^{(n)}(0) = 0$, $\forall n \in \mathbb{N}$. We only give the idea of how to prove this claim: one proves by induction that outside 0, $f^{(n)}$ is a function which is a sum of terms of the form $\frac{e^{-\frac{1}{x^2}}}{x^j}$, for some $j \geq 0$, and furthermore $f^{(n)}(0) = 0$. In view of this claim, the Taylor series of f around 0 is the constant 0 function. So, the radius of convergence is $+\infty$, but apart from 0 there is no positive x at which the Taylor series $T_{f,0}(x) = f(x)$.

Example 4.15. (1) The Taylor series of $f(x) = e^x$ centered at $x = 0$ is given by

$$T_{f,0}(x) := \sum_{k=0}^{\infty} \frac{1}{k!} x^k.$$

In fact, $f(0) = e^0 = 1$,

$$f^{(n)}(0) = (e^x)(0) = 1$$

We have already seen that the radius of convergence of $T_{f,0}(x)$ is $+\infty$. In particular, according to Theorem 4.6, $T_{f,0}(x)$ $\forall x \in \mathbb{R}$. Hence, we have to understand whether for $x \in \mathbb{R}$

$$e^x = T_{f,0}(x). \tag{4.15.b}$$

Let us use Proposition 4.13. For $x < 0$, then

$$\sup_{y \in]x,0[} |f^{(n)}(y)| = \sup_{y \in]x,0[} |e^y| = 1,$$

and

$$\lim_{n\to\infty} \left(\sup_{y \in]x,0[} \left| f^{(n)}(y) \right| \right) \frac{|x|^n}{n!} = \lim_{n\to\infty} \frac{|x|^n}{n!} = 0,$$

since x belongs to the domain of convergence of the power series. Similarly, for $x > 0$, then

$$\sup_{y \in]0,x[} |f^{(n)}(y)| = \sup_{y \in]x,0[} |e^y| = e^x,$$

and

$$\lim_{n\to\infty} \left(\sup_{y \in]0,x[} \left| f^{(n)}(y) \right| \right) \frac{|x|^n}{n!} = \lim_{n\to\infty} \frac{e^x |x|^n}{n!} = 0,$$

by the same argument as above. Then, for $x \in \mathbb{R}$,

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k.$$

(2) The Taylor series of $f(x) = \log(1 + x)$ centered at $x = 0$ is given by

$$T_{f,0}(x) := \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k.$$

In fact, $f(0) = \log(1) = 0$,

$$f'(0) = (\frac{1}{1+x})(0) = \frac{1}{1+0} = 1, \quad f^{(n)}(0) = \left(\frac{(-1)^{n-1}(n-1)!}{(1+x)^n} \right)(0) = (-1)^{n-1}(n-1)!.$$

Let us determine the radius of convergence of $T_{f,0}(x)$:

$$\lim_{n\to\infty}\left|\frac{\frac{(-1)^n}{n+1}}{\frac{(-1)^{n-1}}{n}}\right| = 1,$$

so the radius of convergence of $T_{f,0}(x)$ is 1. In particular, according to Theorem 4.6, $\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}x^k$ converges whenever $|x|<1$ and diverges whenever $|x|>1$. Hence, we have to understand whether for $x\in]-1,1[$, then

$$\log(1+x) = T_{f,0}(x). \tag{4.15.c}$$

Let us use Proposition 4.13. For $-1<x<0$, then

$$\sup_{y\in]x,0[}|f^{(n)}(y)| = \sup_{y\in]x,0[}|\frac{(n-1)!}{(1+y)^n}| = \frac{(n-1)!}{(1+x)^n},$$

and

$$\lim_{n\to\infty}\left(\sup_{y\in]x,0[}\left|f^{(n)}(y)\right|\right)\frac{|x|^n}{n!} = \lim_{n\to\infty}\frac{|x|^n}{n(1+x)^n} = 0,$$

since $\lim\limits_{n\to\infty}|x|^n = \lim\limits_{n\to\infty}(1+x)^n = 0$. Similarly, for $0<x<1$, then

$$\sup_{y\in]0,x[}|f^{(n)}(y)| = \sup_{y\in]x,0[}|\frac{(n-1)!}{(1+y)^n}| = (n-1)!,$$

and

$$\lim_{n\to\infty}\left(\sup_{y\in]0,x[}\left|f^{(n)}(y)\right|\right)\frac{|x|^n}{n!} = \lim_{n\to\infty}\frac{|x|^n}{n} = 0,$$

since $|x|<1$. Then, for $x\in]-1,1[$,

$$\log(1+x) = \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k}x^k.$$

Example 4.16. Here we collect a list of important functions with their Taylor series centered at 0.

$f(x)$	$T_{f,0}(x)$	radius of convergence of $T_{f,0}(x)$	$f(x)=T_{f,0}(x)$ over the interval
$\frac{1}{1-x}$	$\sum_{k=0}^{\infty}x^k$	1	$]-1,1[$
$\frac{1}{1+x}$	$\sum_{k=0}^{\infty}(-1)^k x^k$	1	$]-1,1[$
e^x	$\sum_{k=0}^{\infty}\frac{1}{k!}x^k$	$+\infty$	$\mathbb{R}$
$\log(1+x)$	$\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}x^k$	1	$]-1,1[$
$\cos(x)$	$\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k)!}x^{2k}$	$+\infty$	$\mathbb{R}$
$\sin(x)$	$\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)!}x^{2k+1}$	$+\infty$	$\mathbb{R}$
$\arctan(x)$	$\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}x^{2k+1}$	1	$]-1,1[$
$\cosh(x)$	$\sum_{k=0}^{\infty}\frac{1}{(2k)!}x^{2k}$	$+\infty$	$\mathbb{R}$
$\sinh(x)$	$\sum_{k=0}^{\infty}\frac{1}{(2k+1)!}x^{2k+1}$	$+\infty$	$\mathbb{R}$

Example 4.17. Similar example as above is the Taylor series for $\frac{1}{x+1}$. Using the Taylor expansion, computed in Example 2.77, the Taylor series is

$$\sum_{k=0}^{\infty}(-1)^k x^k$$

As in Example 4.15, using Theorem 4.6, the radius of convergence is 1. Furthermore, using Proposition 4.13 as in Example 4.15, for $x \in]-1,1[$,

$$f(x) = \frac{1}{x+1} = \sum_{k=0}^{\infty}(-1)^k x^k. \tag{4.17.d}$$

Example 4.18. If we want to compute the derivative and integral of $f(x) = \log(x+1)$, then, over $]-1,1[$, we can do so using Theorem 4.8 and deriving/integrating the power series term by term. So, for $x \in]-1,1[$:

$$\begin{aligned} f'(x) =& \sum_{k=1}^{\infty} \frac{(-1)^{k+1}k}{k}(x)^{k-1} = \sum_{k=1}^{\infty}(-1)^{k-1}(x)^{k-1} \\ =& \sum_{k=0}^{\infty}(-1)^k(x)^k = \frac{1}{1+x}, \end{aligned}$$

and

$$\begin{aligned} \int_0^x f(t)dt =& \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)}(x)^{k+1} = \sum_{k=1}^{\infty}(-1)^{k+1}\left(\frac{1}{k} - \frac{1}{k+1}\right)(x-1)^{k+1} \\ =&(x-1)\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}x^k + \sum_{k=1}^{\infty}(-1)^{k+2}\frac{1}{k+1}x^{k+1} \\ =&(x-1)\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}x^k + \sum_{k=2}^{\infty}(-1)^{k+1}\frac{1}{k}x^k \\ =&x\log(1+x) + \log(x+1) - x = (x+1)\log(x+1) - x \end{aligned}$$

where the result of the second computation is exactly what we obtained in Example 3.37