INTRODUCTION TO ANALYSIS (MATH 420)

VERSION: December 4, 2023

ARMIN SCHIKORRA

Contents

References 4

Index 5

1. Review 7
2. Sequences review 13
3. Limit superior, limit inferior, Bolzano-Weierstrass 22
4. Cauchy sequences 29
5. Limits of functions 33
6. Continuous functions 43
7. Min-Max Theorem 53
8. Intermediate Value Theorem 58
9. Uniform continuity 61
10. Derivatives 67
11. Fermat's theorem 72
12. Mean Value Theorem 75
13. Continuous and differentiable function spaces 84
14. The Riemann integral 85
15. Fundamental Theorem of Calculus 104
16. Sequences of functions: Pointwise and uniform convergence 108
17. Series of functions – The Weierstrass M-Test 122
18. Power series 126

19. Taylor's Theorem 130
20. Picard-Lindelöf or Cauchy-Lipschitz Theorem for ordinary differential equations 136

In Analysis
there are no theorems
only proofs

These lecture notes are *substantially* based on the book [Leb], also several exercises are taken from there. Some exercises are also substantially inspired from [BS92].

For more exercises see also the standard reference [Rud76], which often is lovingly referred to as "Baby Rudin".

Pictures are taken from wikipedia or otherwise available sources. Self-made pictures are often made with geogebra.

If you find typos (most likely there are *many*) please email me: `armin@pitt.edu`.

References

[BS92] Robert G. Bartle and Donald R. Sherbert. *Introduction to real analysis*. John Wiley & Sons, Inc., New York, second edition, 1992.

[ht] Americo Tavares (https://math.stackexchange.com/users/752/amtavares). Proof of the L'hopital Rule for $\frac{\infty}{\infty}$. Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/147463 (version: 2012-05-21).

[KP02] Steven G. Krantz and Harold R. Parks. *A primer of real analytic functions*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Boston, Inc., Boston, MA, second edition, 2002.

[Leb] Jiri Lebl. *Basic Analysis*. free, http://calculus.math.pitt.edu/books/pittanal.pdf.

[Rud76] Walter Rudin. *Principles of mathematical analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, third edition, 1976.

Index

C^k, 119
$\mathcal{R}([a,b])$, 87
n-th Taylor polynomial, 131
Cauchy-Lipschitz-Theorem, 136
Fermat's theorem, 73
Picard-Lindelöf-Theorem, 136
Banach Fixed Point theorem, 50
Banach space, 111, 113
Cantor set, 97
Cauchy criterion, 31, 89
Cauchy sequence, 9, 29
Darboux sum, 85
Dedekind cuts, 9
Dirichlet function, 48
Euclidean metric, 11
Euler-Lagrange equation, 74
Fubini's theorem, 134
Hölder constant, 49
Hölder continuous, 49, 63
Hölder's inequality, 107
Hausdorff dimension, 97
Lipschitz constant, 49
Lipschitz continuous, 49, 62
Riemann integrable, 87
Riemann integral, 85
Riemann sum, 101
Sobolev-Poincaré inequality, 107
Taylor polynomial, 131
Thomae's function, 48
Thomae-function, 99

absolute value, 11
all but finitely many, 19
almost everywhere, 96, 99
analytic, 134

boundary, 34
bounded, 7, 12, 13
bounded from above, 7, 12, 53
bounded from below, 7, 12, 53
bounded function, 54
bump function, 84, 100

closed, 34
cluster point, 34
compact, 64
complete, 8, 9, 30, 52
completion, 8
complex numbers, 7
continuity, 19
continuous, 43
contraction, 33, 49, 52
convergence, 13
convergent subsequence, 20
converges pointwise, 109
critical points, 73

diagonal argument, 33
differentiable, 67
directional, 74
discontinuous, 47

energy, 74

finite dimensional, 29
first derivative test, 78
function space, 111
Fundamental Theorem of Calculus, 104
fundamental theorem of calculus, 104

improper integral, 89
infimum, 7, 12
integers numbers, 7
intermediate value property, 61
intermediate value theorem, 58, 60, 83
inverse function theorem, 51

limit inferior, 22
limit superior, 22
linear, 51
local extremum, 72
local maximum, 72
local minimum, 72
lower bound, 7
lower semicontinuos, 66
lower semicontinuous, 55, 56
lower Darboux sum, 86
lower Riemann integral, 86

maximum, 7
maximum point, 54
mean value theorem for integrals, 100
measure zero, 96
metric, 11, 111
metric completion, 30, 32
metric space, 11
minimum, 7

minimum point, 54
modulus of continuity, 50
mollifier, 84
monotone, 16
monotone decreasing, 16
monotone increasing, 16

natural numbers, 7
norm, 111

one-sided limits, 43
open, 34, 47

partition, 85
Picard iterates, 137
pointwise, 124
pointwise convergence, 109
popcorn function, 48

radius of convergence, 127
rational numbers, 7
real numbers, 7, 8
recovery sequence, 56
refinement, 87
relative open, 41
remainder term, 132

sequential continuity, 45
series, 122
strict contraction, 49
strictly monotone decreasing, 16
strictly monotone increasing, 16
subdifferential, 73
subsequence, 20
supremum, 7, 12

totally ordered set, 7

unbounded, 13
uniform convergence, 111
uniformly continuous, 61, 62
upper bound, 7
upper semicontinuos, 66
upper semicontinuous, 56
upper Darboux sum, 86
upper Riemann integral, 86

Wirtinger's inequality, 107

zero measure, 96

1. Review

1.1. Numbers.

- $\mathbb{N}$ denotes the *natural numbers* $\{1, 2, \ldots\}$[1]
- $\mathbb{Z}$ denotes the *integers numbers* $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
- $\mathbb{Q}$ denotes the *rational numbers* $\{\frac{p}{q} : \quad p \in \mathbb{Z}, q \in \mathbb{Z}\backslash\{0\}\}$
- We are going to discuss our main number field, the *real numbers* $\mathbb{R}$, below.
- We are not really going to work with *complex numbers* $\mathbb{C}$.

Recall the notion of an *upper bound* and *lower bound*:

Definition 1.1. Let X be a *totally ordered set* (i.e. there exist the operation $<$ with the usual reasonable properties and for any two $x, y \in X$ we have either $x = y$ or $x < y$ or $x > y$)[2]

- A set $A \subset X$ has an *upper bound* $c \in X$ if for any $a \in A$ we have $a \leq c$ (i.e. either $a < c$ or $a = c$).
- A set $A \subset X$ has a *lower bound* $c \in X$ if for any $a \in A$ we have $a \geq c$ (i.e. either $a < c$ or $a = c$).

A set $A \subset X$ with an upper bound is called *bounded from above*. A set $A \subset X$ with a lower bound is called *bounded from below*. A set A which is bounded from above and below is called *bounded*.

The *supremum* of a set is the smallest upper bound, the *infimum* is the largest lower bound – *if that exists* (because e.g. in $\mathbb{Q}$ it often doesn't).

Definition 1.2 (Supremum and infimum)**.** Let X be a *totally ordered set* and let $A \subset X$.

- A number $c \in X$ is called the *supremum* of A,
$$\sup A = c$$
if
  (1) c is an upper bound of A and
  (2) for any other upper bound b of A we have $c \leq b$.
  We call c the *maximum* of A, $c = \max A$, if $c = \sup A$ and additonally $c \in A$.
- A number $c \in X$ is called the *infimum* of A,
$$\inf A = c$$
if c is a lower bound of A and for any other lower bound b of A we have $c \geq b$.
  We call c the *minimum* of A, if $c = \inf A$ and $c \in A$.

[1] we do *not* consider 0 to be a natural number (this is not always the case in the literature)

[2] so $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}$ are clearly totally ordered sets – but e.g. for $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ it is a bit unclear how to define $\leq$ – or the set of powersets $(2^X, \subseteq)$ is often not a totally ordered set.

If $X = \mathbb{R}$ (as will be the case most of the time), then for notational convenience we often write

$$\sup A = +\infty \quad \text{if } A \text{ has no upper bound,}$$
$$\inf A = -\infty \quad \text{if } A \text{ has no lower bound.}$$

In the pathological case $A = \emptyset$ we write

$$\sup A = -\infty \quad \text{if } A = \emptyset$$
$$\inf A = +\infty \quad \text{if } A = \emptyset.$$

Example 1.3.

- In $\mathbb{Q}$, the set
$$\{q \in \mathbb{Q}, -\infty < q < 2\} \equiv \mathbb{Q} \cap (-\infty, 2)$$
is bounded from above, not bounded from below
- In $\mathbb{Q}$, the set
$$\{q \in \mathbb{Q}, -\infty < q < 2\} \equiv \mathbb{Q} \cap (-\infty, 2)$$
has no infimum (i.e. $\inf = -\infty$), the supremum is 2. 2 is not a maximum, though.
- In $\mathbb{Q}$, the set
$$\{q \in \mathbb{Q}, -\infty < q \leq 2\} \equiv \mathbb{Q} \cap (-\infty, 2]$$
has no infimum, but the maximum is 2.
- In $\mathbb{Q}$, the set
$$\{q \in \mathbb{Q}, -1 < q < \sqrt{2}\} \equiv \mathbb{Q} \cap (-\infty, \sqrt{2}) \equiv \{q \in \mathbb{Q}, -1 < q < \infty \text{ and } q^2 \leq 2\}$$
is bounded from above and below. The infimum is -1. There is no supremum (it would be $\sqrt{2}$, but $\sqrt{2}$ does not belong to $\mathbb{Q}$).
- If a set $A \subset X$ has a supremum, it is necessarily bounded from above (similar statement for infimum)
- Any bounded set $A \subset \mathbb{Z}$ has a supremum and an infimum in $\mathbb{Z}$

Bounded sets in $\mathbb{Q}$ have always "*almost*" a supremum and an infimum – the only problem is this number may not belong to $\mathbb{Q}$. In other words, $\mathbb{Q}$ has infinitesiumal holes, it is not *complete*. This is why we defined $\mathbb{R}$, the real numbers, which are the *completion* of $\mathbb{Q}$.

- $\mathbb{R}$ denotes the *real numbers*. There are many different ways to define them:
  - Element of $\mathbb{R}$ correspond to the supremum of bounded sets $A \subset \mathbb{Q}$:
    Define
    $$\mathbb{R} := \{A \subset \mathbb{Q} : \quad \text{A bounded}\} / \sim$$
    where $\sim$ is an equivalence relation defined as
    $$A \sim B :\Leftrightarrow \text{every upper bound } a \in Q \text{ of } A \text{ is an upper bound of } B \text{ and vice versa.}$$
    Then $\mathbb{R}$ can be ordered just as $\mathbb{Q}$, and any element in $q \in \mathbb{Q}$ corresponds to the set
    $$\{q\} \sim \{r \in \mathbb{Q}, r \leq q\} \sim \{r \in \mathbb{Q}, 1 - q \leq r \leq q\} \sim \{r \in \mathbb{Q}, 1 - q \leq r < q\}.$$

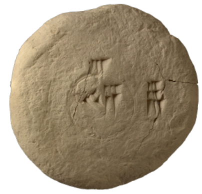

FIGURE 1.1. While some artefacts suggests that Babylonians simply used $\pi = 3$, like this one, there are also indications that people at the same time (not only in Babylon) knew that there was a more precise approximation. Source: Yale Babylonian Collection, 7302

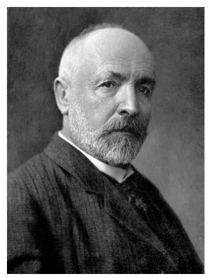

FIGURE 1.2. Georg Cantor, 1845-1918. German, one of the founders of modern set theory and the notion of cardinality.

This definition of the real numbers is related to the so-called *Dedekind cuts* (which had been considered already by Bertrand)

- From Analysis aspects, this is not such a great definition, since it requires an ordering $<$. Many generalized spaces (vector spaces, metric spaces, manifolds, function spaces) have no reasonable order. So instead, we will define (metric) "*complete*" and "completion" as plugging holes of limits (see *Cauchy sequence*s, Section 4). From this point of view $\mathbb{R}$ consist of all finite *limits* of sequences in $\mathbb{Q}$.

The history of "rational numbers are not everything" is very long – people around the world understood that e.g. $\sqrt{2}$ or π were not rational numbers thousands of years ago.[3] The modern understanding of $\mathbb{R}$ is due to Cantor who axiomatized set theory.

For now (until we get to Cauchy sequences, Section 4) we use the following property of $\mathbb{R}$[4]

Proposition 1.4. *For any bounded set $A \subset \mathbb{R}$ both $\sup A$ and $\inf A$ exist in $\mathbb{R}$.*

A useful classification of suprema and infima is the following

[3]Legend has it that Pythagoras, who lead some sort of number cult, had Hippasus murdered for figuring out that there were numbers not being able to be written as a ratio of two integers, namely $\sqrt{2}$. Early approximations of $\sqrt{2}$ are known e.g. from Shulva Sutras (India) or the Babylonian clay tablet YBC 7289

[4]indeed it is the definiting property of $\mathbb{R}$: $\mathbb{R}$ is the "smallest" set containing $\mathbb{Q}$ with these properties

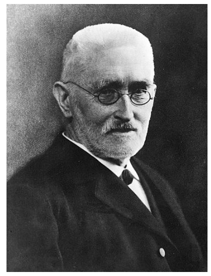

FIGURE 1.3. Richard Dedekind, 1831–1916. German, best known for his contributions to the definition of $\mathbb{R}$ via the notion of Dedekind cuts (which Bertrand actually defined before him).

FIGURE 1.4. Joseph Louis Francois Bertrand, 1822 – 1900 . French, did Dedekind cuts before Dedekind.

Lemma 1.5. *Let $S \subset \mathbb{R}$, $S \neq \emptyset$, and $x \in \mathbb{R}$.*

(1) The following are equivalent
  - *(a) $x = \sup S$*
  - *(b) x is an upper bound of S and for any $\varepsilon > 0$ there exists $s \in S$ with $s > x - \varepsilon$.*

(2) The following are equivalent
  - *(a) $x = \inf S$*
  - *(b) x is a lower bound of S and for any $\varepsilon > 0$ there exists $s \in S$ with $s < x + \varepsilon$.*

(3) The following are equivalent
  - *(a) $\sup S = \infty$*
  - *(b) For any $M > 0$ there exists $s \in S$ with $s > M$.*

(4) The following are equivalent
  - *(a) $\inf S = -\infty$*
  - *(b) For any $M > 0$ there exists $s \in S$ with $s < -M$.*

Proof. We only prove the first statement: Let $S \subset \mathbb{R}$ and $x \in \mathbb{R}$. The following are equivalent

(a) $x = \sup S$
(b) x is an upper bound of S and for any $\varepsilon > 0$ there exists $s \in S$ with $s > x - \varepsilon$.

$(a) \Rightarrow (b)$. Assume that $x = \sup S$, but assume that (b) is false. By definition of sup, x is an upper bound of S. If (b) is false, there thus must be $\varepsilon > 0$ such that for all $s \in S$ we have $s \leq x - \varepsilon$. This implies that $x - \varepsilon$ is an upper bound for S. Since $x - \varepsilon < x$, x cannot be the least upper bound of S. Contradiction. So (b) must have been true.

$(b) \Rightarrow (a)$. x is an upper bound of S, we only need to show that x is the *least* upper bound. So let $y \in \mathbb{R}$ be another upper bound of S, i.e. assume that $s \leq y$ for all $s \in S$. We need to show that $y \geq x$. Assume to the contrary that $y < x$. For $\varepsilon := \frac{|x-y|}{2}$ we then have $y < x - \varepsilon$. Since we assume that (b) holds, there exists an $s \in S$ with $s > x - \varepsilon$. But then $s > x - \varepsilon > y$ which means that y is not an upper bound of S. contradiction, so (a) must have been true.

□

Exercise. *Prove Proposition 1.4 (3) and (4).*

Exercise 1.6. [Leb, Exercise 1.2.10] *Let A and B be two nonempty bounded sets of non-negative real numbers. Define the set*

$$C := \{ab : \quad a \in A, \quad b \in B\}.$$

Show that C is a bounded set and that

$$\sup C = (\sup A)\,(\sup B)$$

and

$$\inf C = (\inf A)\,(\inf B)$$

1.2. **The Euclidean metric – absolute value.** For $x \in \mathbb{R}$ we define the *absolute value* $|x|$ as

$$|x| = \begin{cases} x & \text{if } x > 0 \\ -x & \text{if } x \leq 0. \end{cases}$$

The absolute value is incredibly important for the Analysis in $\mathbb{R}$, because it gives $\mathbb{R}$ a *metric*: we can use it to measure the (a reasonable) distance between to points $x, y \in \mathbb{R}$. Indeed, $d(x, y) := |x - y|$ is the so-called *Euclidean metric*.

Definition 1.7 (metric)**.** A map $d : X \times X \to \mathbb{R}$ is called a *metric* for a set X if

- $d(x, y) = d(y, x)$ for all $x, y \in X$ (symmetry)
- $d(x, y) \geq 0$ for all $x, y \in X$ (positivity)
- $d(x, y) = 0$ if and only if $x = y$ (non-degeneracy)
- $d(x, y) \leq d(x, z) + d(y, z)$ for all $x, y, z \in X$ (triangular inequality).

A set X with a metric d is called a *metric space*.

Almost everything[5] we do with respect to convergence, continuity has a metric space generalization. The proofs are the same, the theorem changes from $\mathbb{R}$ to a general metric space (X, d). Differentiability, however, becomes more tricky, then more structural assumptions on d are helpful (e.g. a "norm" structure).

Exercise 1.8. *Show the following*

(1) $d(x,y) = 2|x-y|$ *is a metric in* $\mathbb{R}$.
(2) $d(x,y) = \sqrt{|x-y|}$ *a metric in* $\mathbb{R}$
(3) $d(x,y) = |x-y|^2$ *is no metric in* $\mathbb{R}$
(4) $d(x,y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$ *is a metric in* $\mathbb{R}$

Exercise 1.9. [Leb, Exercise 1.3.1] *Let* $\varepsilon > 0$. *Show that* $|x-y| < \varepsilon$ *if and only if* $x - \varepsilon < y < x + \varepsilon$.

Exercise 1.10. [Leb, Exercise 1.3.2.]

(1) *Show that*
$$\max\{x,y\} = \frac{x+y+|x-y|}{2}$$
(2) *Show that*
$$\min\{x,y\} = \frac{x+y-|x-y|}{2}$$

1.3. **functions: boundedness, infimum, supremum.** We will mostly consider functions $f : D \subset \mathbb{R} \to \mathbb{R}$. But of course one can also consider more general sets D (like $D \subset \mathbb{R}^2$ etc.)

Definition 1.11. A function $f : D \to \mathbb{R}$ is

- *bounded from above* if there exists $M \in \mathbb{R}$ with $f(x) \leq M$ for all $x \in D$.
- *bounded from below* if there exists $M \in \mathbb{R}$ with $f(x) \geq M$ for all $x \in D$.
- *bounded* if it is bounded from above and below. In other terms: if there exists $M \in \mathbb{R}$ with $|f(x)| \leq M$ for all $x \in D$.

For a function $f : D \to \mathbb{R}$ we define (if existent)

- the *supremum* $\sup_D f := \sup f(D)$. If there exists $x \in D$ such that $f(x) = \sup_D f$ then $\max_D f := \sup_D f$ is called *the maximum (value)*.
- the *infimum* $\inf_D f := \inf f(D)$. If there exists $x \in D$ such that $f(x) = \inf_D f$ then $\min_D f := \inf_D f$ is called *the minimum (value)*.

[5]very importantly, *not* the Bolzano-Weierstrass theorem, Theorem 3.8, though

For notational convenience we write

$$\sup_D f = +\infty \quad \text{if } D \neq \emptyset \text{ and } f \text{ is not bounded from above}$$

$$\inf_D f = -\infty \quad \text{if } D \neq \emptyset \text{ and } f \text{ is not bounded from below}$$

and the pathological cases

$$\sup_D f = -\infty \quad \text{if } D = \emptyset$$

$$\inf_D f = +\infty \quad \text{if } D = \emptyset$$

Exercise 1.12. *Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Let*

$$g(x) := -f(x).$$

Show that for any $D \subset \mathbb{R}$ (including $D = \emptyset$)

$$\sup_D g = -\inf_D f$$

and

$$\inf_D g = -\sup_D f$$

2. Sequences review

(it is a fun exercise to try to translate the statements here into notions on metric spaces, cf. Definition 1.7)

A sequence, usually denoted by $(x_n)_{n=1}^{\infty} \subset X$, is a map $x : \mathbb{N} \mapsto X$. But instead of writing $x(n)$ we prefer to write x_n. Every sequence induces a set $x(\mathbb{N}) := \{x_n, n \in \mathbb{N}\}$ (but not the other way around, since we do not know which element of the set to take first). Thus we can use set operations on sequences, e.g.,

$$\sup(x_n)_{n\in\mathbb{N}} = \sup\{x_1, x_2, \ldots\}.$$

Definition 2.1. A sequence $(x_n)_{n=1}^{\infty}$

- is *bounded* if the set $\{x_1, \ldots, x_n, \ldots\} \subset \mathbb{R}$ is bounded.
- is *unbounded* if the set $\{x_1, \ldots, x_n, \ldots\} \subset \mathbb{R}$ is not bounded.
- *converges* to a number $x \in \mathbb{R}$ if

  $$\forall \varepsilon > 0 : \quad \exists N \in \mathbb{N} \quad \text{s.t. } |x_n - x| < \varepsilon \quad \forall n \geq N.$$

  In words: *all* sequence elements x_n with sufficiently large index $n \geq N$ are very close to the limit point x.

  In this case we say that x_n is convergent (to x).

  $$\lim_{n\to\infty} x_n = x.$$

  For a picture see Figure 2.1.

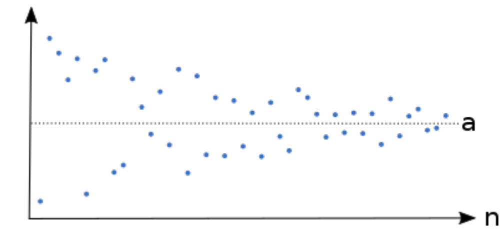

FIGURE 2.1. A sequence $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ which seems to converge to a, $\lim_{n\to\infty} a_n = a$

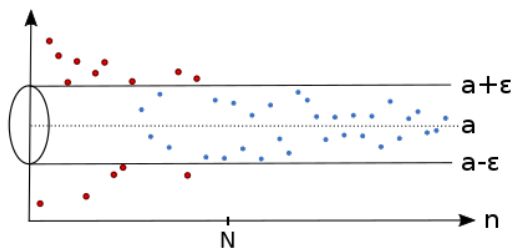

FIGURE 2.2. For a given ε, several sequence elements (red) are not close to a at the scale ε: $|a_n - a| \geq \varepsilon$ for the red a_n. But most of the sequence elements (blue) are close to a at the scale ε: $|a_n - a| < \varepsilon$. Indeed, we see that after some large enough number N, *all* sequence elements are blue, i.e. close to a, i.e.e $|a_n - a| < \varepsilon$ for all $n > N$.

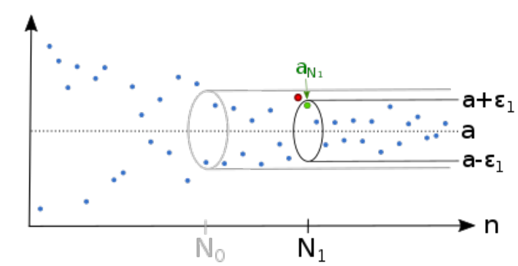

FIGURE 2.3. In order to show the convergence $\lim_n a_n = a$ we have to show this for $\varepsilon > 0$ we can find such an N from which on $|a_n - a| < \varepsilon$. The N is allowed to change with ε: for $\varepsilon_0 > 0$ we find some N_0, for ε_1 we find another N_1. In general, as $\varepsilon > 0$ is smaller N needs to be chosen larger.

all pictures: Ceranilo, wikipedia.

If the limit exists, then it is *unique* that is

Exercise 2.2. *Assume that $(x_n)_{n\in\mathbb{N}} \subset \mathbb{Q}$ is a sequence and for $x, y \in \mathbb{Q}$ we have*

$$\lim_{n\to\infty} x_n = x, \quad \lim_{n\to\infty} x_n = y$$

Show that $x = y$.

Exercise 2.3. *Show the following*

- *If* $x_n = 1 + \frac{1}{n}$*:* $\lim_{n\to\infty} x_n = 1$.
- *If* $x_n = (-1)^n$ *does not converge.*

Example 2.4. If $x_n = \frac{n^2}{2n^2+n}$ then $\lim_{n\to\infty} x_n = \frac{1}{2}$

Indeed: Let $\varepsilon > 0$ be given. We need to find $N \in \mathbb{N}$ such that

$$\left| x_n - \frac{1}{2} \right| < \varepsilon \quad \forall n \geq N.$$

Now observe that

$$\begin{aligned} x_n - \frac{1}{2} &= \frac{n^2}{2n^2+n} - \frac{1}{2} \\ &= \frac{2n^2 - (2n^2+n)}{4n^2+2n} \\ &= \frac{-n}{4n^2+2n} \\ &= \frac{-1}{4n+2} \end{aligned}$$

Thus,

$$\begin{aligned} \left| x_n - \frac{1}{2} \right| &= \frac{1}{4n+2} \\ &\leq \frac{1}{4n}. \end{aligned}$$

So if we choose $N \in \mathbb{N}$ such that $N > \frac{1}{4\varepsilon}$ then for any $n \geq N$

$$\left| x_n - \frac{1}{2} \right| \leq \frac{1}{4n} \leq \frac{1}{4N} < \varepsilon.$$

Lemma 2.5. *Every convergent sequence* $(x_n)_{n\in\mathbb{N}}$ *is bounded, i.e. there exists* $M \in \mathbb{R}$ *such that* $|x_n| \leq M$ *for all* $n \in \mathbb{N}$.

Proof.
- Since x_n is convergent, there exists $x \in \mathbb{R}$ and and $N \in \mathbb{N}$ such that
$$|x_n - x| \leq 1 \quad \forall n > N.$$
- Set $\tilde{M} := \max\{|x_1|, \dots, |x_N|\}$ – this maximum exists, because there are only finitely many points considered.
- Set $M := |x| + \tilde{M} + 1$. Then we have
$$|x_n| \leq \tilde{M} \leq M \quad \forall n \leq N$$
and
$$|x_n| \leq |x_n - x| + |x| \leq 1 + |x| \leq M \quad \forall n > N$$
That is $|x_n| \leq M$ for all $n \in \mathbb{N}$, and thus the sequence x_n is bounded.

□

Corollary 2.6. *Any unbounded sequence is not convergent.*

Proof. This is just the logical equivalent of Lemma 2.5. Namely $A \Rightarrow B$ is equivalent to $\neg B \Rightarrow \neg A$, so

$$\begin{array}{llcl} & ((x_n)_{n\in\mathbb{N}} \text{ convergent}) & \Rightarrow & ((x_n)_{n\in\mathbb{N}} \text{ bounded}) \\ \Leftrightarrow & \neg((x_n)_{n\in\mathbb{N}} \text{ convergent}) & \Leftarrow & \neg((x_n)_{n\in\mathbb{N}} \text{ bounded}) \\ \Leftrightarrow & ((x_n)_{n\in\mathbb{N}} \text{ not convergent}) & \Leftarrow & ((x_n)_{n\in\mathbb{N}} \text{ not bounded}) \\ \Leftrightarrow & ((x_n)_{n\in\mathbb{N}} \text{ not bounded}) & \Rightarrow & ((x_n)_{n\in\mathbb{N}} \text{ not convergent}) \end{array}$$

□

Remark. In a *very common abuse of notation* we shall write

- "x_n converges to $+\infty$", in formulas
$$\lim_{n\to\infty} x_n = +\infty,$$
if
$$\forall M > 0\ \exists N \in \mathbb{N} \quad \text{such that} \quad x_n > M \quad \forall n > N,$$
that is all sequence elements are eventually very large.
- "x_n converges to $-\infty$", in formulas
$$\lim_{n\to\infty} x_n = -\infty,$$
if
$$\forall M > 0\ \exists N \in \mathbb{N} \quad \text{such that} \quad x_n < -M \quad \forall n > N.$$
that is all sequence elements are eventually very negative.

Definition 2.7. A sequence $(x_n)_{n\in\mathbb{N}}$ is[6]

- *monotone increasing* if $x_n \leq x_m$ holds for any $n, m \in \mathbb{N}$ with $n \leq m$
- *strictly monotone increasing* if $x_n < x_m$ holds for any $n, m \in \mathbb{N}$ with $n < m$
- *monotone decreasing* if $x_n \geq x_m$ holds for any $n, m \in \mathbb{N}$ with $n \leq m$
- *strictly monotone decreasing* if $x_n > x_m$ holds for any $n, m \in \mathbb{N}$ with $n < m$
- *monotone* if it is either monotone increasing or monotonce decreasing.

Theorem 2.8 (Bounded monotone sequences are convergent)**.** *Let $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ be a bounded monotone sequence. Then $x = \lim_{n\to\infty} x_n$ exists, and*

- $x = \sup_n x_n$ *(if $(x_n)_{n\in\mathbb{N}}$ is increasing), or*
- $x = \inf_n x_n$ *(if $(x_n)_{n\in\mathbb{N}}$ is decreasing).*

[6] of course, this doesnt make any sense in general metric spaces

Proof. Assume w.l.o.g. that x_n is monotone increasing (the other case goes exactly the same way).

Since $\{x_n, n \in \mathbb{N}\} \subset \mathbb{R}$ is bounded by assumption, and $\mathbb{R}$ is a complete space, Proposition 1.4, the supremum exists. We denote it by

$$x := \sup_{n\in\mathbb{N}} x_n.$$

We need to show that $\lim_{n\to\infty} x_n = x$. For this let $\varepsilon > 0$ be arbitrary. We need to find $N = N(\varepsilon) \in \mathbb{N}$ such that

$$|x_n - x| < \varepsilon \quad \forall n > N.$$

Equivalently we need to show that

$$x_n - x < \varepsilon \quad \forall n > N, \tag{2.1}$$

and

$$x - x_n < \varepsilon \quad \forall n > N. \tag{2.2}$$

Observe that (2.1) is true for any $n \in \mathbb{N}$ because x is the supremum of the x_n, and as such $x \geq x_n$ for all $n \in \mathbb{N}$.

So we only need to show (2.2). Assume to the contrary that for any N there exists an $M > N$ such that

$$x - x_M \geq \varepsilon \Leftrightarrow x_M \leq x - \varepsilon$$

But by monotonicity this implies

$$x_m \leq x_M \leq x - \varepsilon \quad \forall m \leq M.$$

That is we would have

$$\forall N \in \mathbb{N}\ \exists M > N: \quad x_m \leq x_M \leq x - \varepsilon \quad \forall m \leq M.$$

In particular we have

$$\forall N \in \mathbb{N}: \quad x_N \leq x - \varepsilon$$

Just relabelling this, we have

$$x_m \leq x - \varepsilon \quad \forall m \in \mathbb{N}$$

But this contradicts that x is the $\sup_n x_n$, indeed $x - \varepsilon$ is a smaller upper bound. Contradiction, so (2.2) must be true for some $N \in \mathbb{N}$. ☐

Exercise 2.9. *Show that the statement of Theorem 2.8 is false if $\mathbb{R}$ is replaced by $\mathbb{Q}$.*

For this give an example of a bounded monotone sequence in $\mathbb{Q}$, $(x_n)_{n\in\mathbb{N}} \subset \mathbb{Q}$, which does not converge in $\mathbb{Q}$. That is, show that there is no $x \in \mathbb{Q}$ with $\lim_{n\to\infty} x_n = x$.

We can also reformulate the supremum and infimum definition of Definition 2.1:

FIGURE 2.4. Lord Sandwich, 1715-1797. Inventor of the Sandwich lemma, Lemma 2.11. Image: Dall-E.

Exercise 2.10. [Leb, Exercise 2.1.12] *Show the following:*

Let $S \subset \mathbb{R}$ *be a nonempty bounded set. Then there exist monotone sequences* $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ *such that* $x_n, y_n \in S$ *for all* n *and*

$$\sup S = \lim_{n\to\infty} x_n$$

and

$$\inf S = \lim_{n\to\infty} y_n$$

Hint: Use the definition of supremum from Lemma 1.5 to find the sequence and Theorem 2.8 to ensure it converges.

The following lemma is also known as the sandwich theorem, cf. Figure 2.4.

Lemma 2.11 (Squeeze theorem)**.** *Assume that we have three real sequences*

$$(a_n)_{n\in\mathbb{N}}, \quad (x_n)_{n\in\mathbb{N}}, \quad (b_n)_{n\in\mathbb{N}}$$

such that

$$a_n \le x_n \le b_n \quad \forall n \in \mathbb{N}. \tag{2.3}$$

If there exists $x \in \mathbb{R}$ *with*

$$x = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$$

then

$$\lim_{n\to\infty} x_n = x.$$

Proof. Since by (2.3)

$$a_n - x \ge x_n - x \le b_n - x,$$

we have

$$|x_n - x| \le \max\{|a_n - x|, |b_n - x|\}. \tag{2.4}$$

Let now $\varepsilon > 0$. Since $a_n \to x$ and $b_n \to x$ there must be an $N(\varepsilon)$[7] such that

$$\max\{|a_n - x|, |b_n - x|\} < \varepsilon \quad \forall n \geq N.$$

Thus, by (2.4),

$$|x_n - x| \leq \max\{|a_n - x|, |b_n - x|\} < \varepsilon.$$

□

Proposition 2.12. *If $(x_n)_{n\in\mathbb{N}}$ is convergent, so is $(|x_n|)_{n\in\mathbb{N}}$, and we have*

$$\lim_{n\to\infty} |x_n| = |\lim_{n\to\infty} x_n|.$$

Proof. This is what we will later call the *continuity* of the absolute value $f(\cdot) := |\cdot|$.

Set

$$x := \lim_{n\to\infty} x_n.$$

The claim follows from the definition of a limit and the inverse triangle inequality which implies

$$||x_n| - |x|| \leq |x_n - x|. \tag{2.5}$$

Since $x_n \to x$, for any $\varepsilon > 0$ there must be $N \in \mathbb{N}$ such that

$$|x_n - x| < \varepsilon \quad \forall n \geq N$$

From (2.5) we conclude that then

$$||x_n| - |x|| < \varepsilon \quad \forall n \geq N$$

which implies by definition that $\lim_{n\to\infty} |x_n| = |x|$ □

Definition 2.13. We say that a property (A) holds for *all but finitely many* elements of a set $S \subset X$ if there exists a finite number K and elements $s_1, \ldots, s_K \in S$ such that property (A) holds for any $s \in S\backslash\{s_1, \ldots, s_K\}$.

It is an easy exercise to show that property (A) holds for all but finitely many elements of a sequence $(x_n)_{n\in\mathbb{N}}$ if and only if there exists a large number $N \in \mathbb{N}$ such that property (A) holds for all x_n, $n \geq N$. When talking about limits of sequences, we usually only care about all but finitely many elements of said sequence. For example:

Lemma 2.14. *Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be two sequences and assume that*

$$x_n \leq y_n \quad \textit{for all but finitely many } n \in \mathbb{N}$$

If $\lim_{n\to\infty} x_n$ and $\lim_{n\to\infty} y_n$ exist, then

$$\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n.$$

Exercise 2.15. *Prove Lemma 2.14*

[7]we take the maximum of the $N(\varepsilon) = \max\{N_1(\varepsilon), N_2(\varepsilon)\}$ where $N_1(\varepsilon)$ is such that the sequence (a_n) satisfies $|a_n - x| < \varepsilon$ for all $n \geq N_1(\varepsilon)$ and $N_2(\varepsilon)$ is such that the sequence (b_n) satisfies $|b_n - x| < \varepsilon$ for all $n \geq N_2(\varepsilon)$

We will also discuss a strengthened version of Lemma 2.14 in Exercise 3.3.

Exercise 2.16. [Leb, Ex. 2.1.3] *Is the sequence* $\left(\frac{(-1)^n}{2n}\right)_{n\in\mathbb{N}}$ *convergent? If so, what is the limit?*

2.1. Subsequences.

Definition 2.17. Suppose $(x_n)_{n\in\mathbb{N}}$ is a sequence. Let $(n_i)_{i\in\mathbb{N}}$ be a strictly increasing sequence of natural numbers (i.e., $n_i < n_{i+1}$ for all i). The sequence

$$(x_{n_i})_{i\in\mathbb{N}}$$

is then called a *subsequence* of $(x_n)_{n\in\mathbb{N}}$.

As sequence $(x_n)_{n\in\mathbb{N}}$ has a *convergent subsequence* if there exists a subsequence $(x_{n_i})_{i\in\mathbb{N}}$ which is convergent.

Example 2.18.
- Let
$$(x_n)_{n\in\mathbb{N}} = (1, 5, 7, 8, 9, 10, 20, 33, \dots)$$
then
$$(y_n)_n = (1, 7, 33, \dots)$$
is a subsequence, whereas
$$(z_n)_n = (1, 7, 5, 33, \dots)$$
is (most likely) not a subsequence.
- Let
$$x_n := (-1)^{n+1}$$
Then
$$x_{2n} = -1$$
and
$$x_{2n+1} = 1.$$
Both subsequences are clearly convergent, but $(x_n)_{n\in\mathbb{N}}$ is clearly not convergent.

Exercise. *Let* $x_1 = 8$ *and* $x_{n+1} := \frac{1}{2}x_n + 2$ *for* $n \in \mathbb{N}$. *Show that* $(x_n)_{n\in\mathbb{N}}$ *is convergent and compute the limit.*

Hint: Use Theorem 2.8.

Lemma 2.19. *If* $(x_n)_{n\in\mathbb{N}}$ *is a convergent sequence, then every subsequence of* $(x_n)_{n\in\mathbb{N}}$ *is also convergent. Moreover if*

$$x := \lim_{n\to\infty} x_n$$

then for any subsequence $(x_{n_i})_{i\in\mathbb{N}}$,

$$x = \lim_{i\to\infty} x_{n_i}$$

Exercise 2.20. *Prove Lemma 2.19.*

Exercise 2.21. *Let* $(x_n)_{n\in\mathbb{N}}$ *be a sequence and assume one of the following property:*

(1) *there is some* x *such that any subsequence* $(x_{n_i})_{i\in\mathbb{N}}$ *contains another subsequence* $\left(x_{n_{i_j}}\right)_{j\in\mathbb{N}}$ *which is convergent to* x.
(2) *any subsequence* $(x_{n_i})_{i\in\mathbb{N}}$ *contains another subsequence* $\left(x_{n_{i_j}}\right)_{j\in\mathbb{N}}$ *which is convergent (a priori not necessarily to the same* x*)*

Show that in one of the cases the sequence x_n *is convergent. Give a counterexample for the other case.*

Exercise 2.22. [Leb, Exercise 2.1.15] *Let* $(x_n)_{n\in\mathbb{N}}$ *be a sequence defined by*

$$x_n := \begin{cases} n & \text{if } n \text{ is odd,} \\ 1/n & \text{if } n \text{ is even.} \end{cases}$$

a) Is the sequence bounded? (prove or disprove)

b) Is there a convergent subsequence? If so, find it.

Exercise 2.23. [Leb, Exercise 2.2.7] *True or false, prove or find a counterexample. If* $(x_n)_{n\in\mathbb{N}}$ *is a sequence such that* $(x_n^2)_{n\in\mathbb{N}}$ *converges, then* (x_n) *converges as well.*

Exercise 2.24. *Let* $(x_n)_{n\in\mathbb{N}}$ *be a sequence and assume one of the following properties:*

(1) *there is some* x *such that any subsequence* $(x_{n_i})_i$ *contains another subsequence* $(x_{n_{i_j}})_j$ *which is convergent to* x.
(2) *any subsequence* $(x_{n_i})_i$ *contains another subsequence* $(x_{n_{i_j}})_j$ *which is convergent (a priori not necessarily to the same* x*)*

Show in which cases $(x_n)_n$ *is convergent. Give a counterexample for the other case.*

Exercise 2.25. *Find the following limit. Show all work.*

$$\lim_{n\to\infty}\left(\frac{1}{\sqrt{n^2+1}}+\frac{1}{\sqrt{n^2+2}}+\ldots+\frac{1}{\sqrt{n^2+2n}}\right)$$

2.2. **Further exercises for limits.** Sequences are very important, so here we collect some (option) ε-N-type exercises

Exercise. *Use the precise* ε, N *definition of limit to prove the following statements.*

(1) $\lim_{n\to\infty}\frac{3n^2+2}{2n^2-5}=\frac{3}{2}$.
(2) $\lim_{n\to\infty}\left|\frac{-n+5}{\sqrt{n}+28}\right|=+\infty$
(3) $\lim_{n\to\infty}\frac{n}{n^2+1}=0$
(4) $\lim_{n\to\infty}\frac{2n}{n+1}=1$

FIGURE 3.1. Karl Weierstrass, 1815-1897. German, "father of modern analysis".

(5) $\lim_{n\to\infty} \frac{n^2+3}{n+5} = +\infty$.
(6) $\lim_{n\to\infty} \frac{1}{\sqrt{n+7}} = 0$
(7) $\lim_{n\to\infty} \frac{(-1)^n n}{n^2+1} = 0$
(8) *More abstractly show that whenever* $a, b \neq 0$ *we have* $\lim_{n\to\infty} \frac{an^2+2n+7}{bn^2+5n-5} = \frac{a}{b}$.

Exercise. *Assume that* $(x_n)_{n\in\mathbb{N}}$ *is a sequence and* $\lim_{n\to\infty} x_n = 5$. *Show that there exists some* $N \in \mathbb{N}$ *such that* $x_n \geq 4$ *for all* $n \geq N$.

Exercise. *Show that*

$$\lim(\frac{2^n}{n!}) = 0.$$

Hint: You can use without proof that for $n \geq 3$ *we have* $\frac{2^n}{n!} \leq 2\left(\frac{2}{3}\right)^{n-2}$

Exercise. *Give an example of an unbounded sequence that has a convergent subsequence.*

Exercise. *Prove that the following sequences are divergent:*

(1) $x_n := 1 + (-1)^n + 1/n$
(2) $y_n := \sin\left(\frac{n\pi}{4}\right)$

Hint: subsequences, Lemma 2.19

Exercise 2.26. *Assume that* $(x_n)_{n\in\mathbb{N}}$ *satisfies* $x_n \geq 0$ *for all* $n \in \mathbb{N}$ *and assume* $\lim_{n\to\infty}(-1)^n x_n$ *exists. Show that* $(x_n)_{n\in\mathbb{N}}$ *is convergent.*

3. Limit superior, limit inferior, Bolzano-Weierstrass

Sequences can be subdivided into subsequences as discussed above, Section 2.1. The *limit superior*, lim sup is the largest possible limit (or $+\infty$) of any subsequence, the *limit inferior*. lim inf is the smallest possible limit (or $-\infty$) of any subsequence. More precisely,

Definition 3.1. Let $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ be any sequence.

Then $\limsup_{n\to\infty} x_n, \liminf_{n\to\infty} x_n \in \mathbb{R}\cup\{-\infty, +\infty\}$ are defined as follows (cf. Figure 3.3)

FIGURE 3.2. Bernard Bolzano, 1781 - 1848. Italian-German-Czech; Bohemian mathematician, philosopher, Catholic priest, antimilitarist.

- If $(x_n)_{n\in\mathbb{N}}$ is bounded from above we set[8]
$$\limsup_{n\to\infty} x_n := \inf_{n\in\mathbb{N}} \sup_{k\geq n} x_k \in \mathbb{R}\cup\{-\infty\}$$
Observe that since $n\mapsto \sup_{k\geq n} x_k$ is monotonce decreasing we have equivalently
$$\limsup_{n\to\infty} x_n := \lim_{n\to\infty} \sup_{k\geq n} x_k \in \mathbb{R}\cup\{-\infty\}$$
so the $\limsup_{n\to\infty} x_n$ computes the largest sequence element "at infinity".
- If $(x_n)_{n\in\mathbb{N}}$ is *not* bounded from above we set $\limsup_{n\to\infty} x_n := +\infty$
- If $(x_n)_{n\in\mathbb{N}}$ is bounded from below we set
$$\liminf_{n\to\infty} x_n := \sup_{n\in\mathbb{N}} \inf_{k\geq n} x_k \equiv \lim_{n\to\infty} \inf_{k\geq n} x_k.$$
so the $\limsup_{n\to\infty} x_n$ computes the smallest sequence element "at infinity".
- If $(x_n)_{n\in\mathbb{N}}$ is not bounded from below we set $\liminf_{n\to\infty} x_n := -\infty$

Exercise 3.2. *Let*
$$x_n := \begin{cases} \frac{1}{n} & n \text{ even} \\ -n & n \text{ odd} \end{cases}$$
Show that
$$\limsup_{n\to\infty} x_n = 0$$
and
$$\liminf_{n\to\infty} x_n = -\infty.$$

Exercise 3.3. *Show the following version of Lemma 2.14:*

Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be two sequences such that
$$x_n \leq y_n \quad \text{for all but finitely many } n\in\mathbb{N}$$
Then we have
$$\liminf_{n\to\infty} x_n \leq \liminf_{n\to\infty} y_n$$

[8]observe that this number exists: $(x_n)_{n\in\mathbb{N}}$ is bounded from above, so $a_n := \sup_{k\geq n} x_k$ is finite number for each n. So the infimum $\inf_n a_n$ is defined by the properties of $\mathbb{R}$, Proposition 1.4.

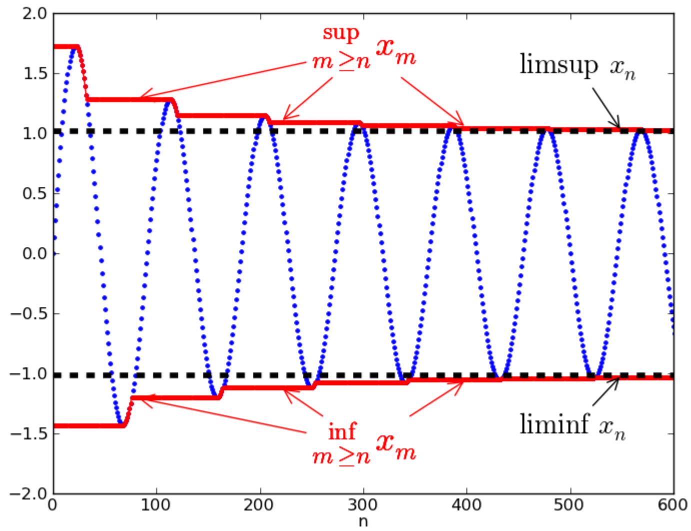

FIGURE 3.3. An illustration of limit superior and limit inferior. The sequence x_n is shown in blue. The two red curves approach the limit superior and limit inferior of x_n, shown as dashed black lines. In this case, the sequence accumulates around the two limits. The superior limit is the larger of the two, and the inferior limit is the smaller of the two. The inferior and superior limits agree if and only if the sequence is convergent (i.e., when there is a single limit). (text and image: Eigenjohnson, Wikipedia)

and

$$\limsup_{n\to\infty} x_n \leq \limsup_{n\to\infty} y_n$$

Exercise 3.4. [Leb, Ex. 2.3.7] *Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be bounded sequences.*

(1) Show that $(x_n + y_n)_{n\in\mathbb{N}}$ is bounded.
(2) Show that[9]

$$(\liminf_{n\to\infty} x_n) + (\liminf_{n\to\infty} y_n) \leq \liminf_{n\to\infty}(x_n + y_n).$$

(3) Find explicit $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ such that

$$(\liminf_{n\to\infty} x_n) + (\liminf_{n\to\infty} y_n) < \liminf_{n\to\infty}(x_n + y_n).$$

To match lim sup and lim inf with our intuition as computing "smallest subsequence" and "largest subsequence", we observe

Lemma 3.5. *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence.*

[9] this stays true if (x_n) and (y_n) are *not* assumed to be unbounded, as long as we avoid $\infty - \infty$ on the left-hand side.

(1) *Set* $a_n := \sup_{k\geq n} x_k$*, then*
$$\limsup_{n\to\infty} x = \lim_{n\to\infty} a_n$$
in the sense that either both sides are finite or both sides are $\pm\infty$*.*

(2) *Set* $b_n := \inf_{k\geq n} x_k$*, then*
$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} b_n$$
in the sense that either both sides are finite or both sides are $\pm\infty$*.*

(3) *Let* $(x_{n_i})_{i\in\mathbb{N}}$ *be any convergent subsequence. Then*
$$\liminf_{n\to\infty} x_n \leq \lim_{i\to\infty} x_{n_i} \leq \limsup_{n\to\infty} x_n.$$

(4) *If* $\limsup_{n\to\infty} x_n \in (-\infty,\infty)$ *then there exists a convergent subsequence* $(x_{n_i})_{i\in\mathbb{N}}$ *with*
$$\lim_{i\to\infty} x_{n_i} = \limsup_{n\to\infty} x_n.$$

(5) *If* $\liminf_{n\to\infty} x_n \in (-\infty,\infty)$ *then there exists a convergent subsequence* $(x_{n_i})_{i\in\mathbb{N}}$ *with*
$$\lim_{i\to\infty} x_{n_i} = \liminf_{n\to\infty} x_n.$$

(6) *If* $\limsup_{n\to\infty} x_n = \infty$ *then there exists a subsequence* $(x_{n_i})_{i\in\mathbb{N}}$ *with* $\lim_{i\to\infty} x_{n_i} = \infty$*. If* $\limsup_{n\to\infty} x_n = -\infty$ *then all subsequences* $(x_{n_i})_{i\in\mathbb{N}}$ *satisfy* $\lim_{i\to\infty} x_{n_i} = -\infty$*.*

(7) *If* $\liminf_{n\to\infty} x_n = -\infty$ *then there exists a subsequence* $(x_{n_i})_{i\in\mathbb{N}}$ *with* $\lim_{i\to\infty} x_{n_i} = -\infty$*. If* $\liminf_{n\to\infty} x_n = +\infty$ *then all subsequences* $(x_{n_i})_{i\in\mathbb{N}}$ *satisfy* $\lim_{i\to\infty} x_{n_i} = +\infty$*.*

Proof. (1) If $(a_n)_{n\in\mathbb{N}}$ is not bounded from above, $(x_n)_{n\in\mathbb{N}}$ is not bounded from above, and so $\lim_{n\to\infty} a_n = \limsup_{n\to\infty} x_n = \infty$.

If a_n is bounded from above then it is a monotonce decreasing, bounded, sequence. From Theorem 2.8 we find that a_n is convergent and
$$\lim_{n\to\infty} a_n = \inf_n \sup_{k\geq n} x_k = \limsup_{n\to\infty} x_n.$$

(2) exercise! (almost the same argument as as above)

(3) We only show
$$\lim_{i\to\infty} x_{n_i} \leq \limsup_{n\to\infty} x_n. \tag{3.1}$$
The other inequality follows the same way.

If $\limsup_n x_n = \infty$ then (3.1) is trivially satisfied. So let us assume $\limsup_n x_n < \infty$. Then
$$x_{n_i} \leq \sup_{k\geq n_i} x_k =: a_i \quad \forall i\in\mathbb{N}.$$

We observe that $(a_i)_{i\in\mathbb{N}}$ is a monotone increasing sequence. Since $\limsup_n x_n < \infty$ we have that a_i is bounded from above. So by Theorem 2.8 a_i is convergent and

$$\lim_{i\to\infty} a_i = \inf_i \sup_{k\geq n_i} x_k \overset{n_i\geq i}{\leq} \inf_i \sup_{k\geq i} x_k = \limsup_{i\to\infty} x_i.$$

By monotonicity of the limit, Lemma 2.14,

$$\lim_{i\to\infty} x_{n_i} \leq \lim_{i\to\infty} a_i = \limsup_{n\to\infty} x_n.$$

(4) Set

$$a_n := \sup_{k\geq n} x_k.$$

Since $\limsup_{n\to\infty} x_n < \infty$, by the definition of supremum as lowest upper bound (cf. Lemma 1.5), for any $n \in \mathbb{N}$ there must be a number $K = K(n) \geq n$ such that

$$a_n - \frac{1}{n} \leq x_K \leq a_n.$$

Now we build our subsequence as follows. $n_1 := K(1)$, $n_2 := K(n_1+1)$, $n_i := K(n_{i-1}+1)$. This is an strictly increasing sequence, and we have

$$a_{n_i} - \frac{1}{n_i} \leq x_{n_{i+1}} \leq a_{n_i} \quad \forall i.$$

Since in particular $n_i \geq i$ we find

$$a_{n_i} - \frac{1}{i} \leq x_{n_{i+1}} \leq a_{n_i} \quad \forall i.$$

By the squeeze theorem, Lemma 2.11, we have that

$$\lim_{i\to\infty} x_{n_i} = \lim_{i\to\infty} a_{n_i} = \limsup_{n\to\infty} x_n.$$

(5) same as above

(6) If $\limsup_{n\to\infty} x_n = \infty$ then $\inf_{n\in\mathbb{N}} a_n = \infty$ where $a_n = \sup_{k\geq n} x_k$. That means that for any $M \in \mathbb{N}$ and for any $n \in \mathbb{N}$ there exists $k = k(n) \geq n$ with $x_k > M$. From this we can build a subsequence. Take $x_{n_1} := x_{k(1)}$, $x_{n_2} := x_{k(k(1)+1)}$ etc. This subsequence goes to infinity.

Assume now that $\limsup_{n\to\infty} x_n = -\infty$ and take $(x_{n_i})_{i\in\mathbb{N}}$ any subsequence.

Then $\inf_{n\in\mathbb{N}} a_n = -\infty$ where $a_n = \sup_{k\geq n} x_k$. That is, for any $M > 0$ there must be some $N \in \mathbb{N}$ such that $a_N < -M$. But since $a_N = \sup_{k\geq N} x_k$, this implies $x_k \leq -M$ for *all* $k \geq N$. That is, for all $M > 0$ we have that $x_n < -M$ for all but finitely many $n \in \mathbb{N}$. In particular, for all $M > 0$ we have that $x_{n_i} < -M$ for all but finitely many $i \in \mathbb{N}$. This means that $\lim_{i\to\infty} x_{n_i} = -\infty$.

(7) analogue argument to above.

□

Lemma 3.6. *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$*

(1) $\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n$

(2) *For any subsequence* (x_{n_i}),

$$\liminf_{n\to\infty} x_n \le \liminf_{i\to\infty} x_{n_i} \le \limsup_{i\to\infty} x_{n_i} \le \limsup_{n\to\infty} x_n$$

(3) *If* $\lim_{n\to\infty} x_n = x$ *then* $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = x$.

(4) *If* $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$ *and the value is finite then* $(x_n)_{n\in\mathbb{N}}$ *converges and we have* $\lim_{n\to\infty} x_n = \liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$.

We collect this in a smashy corollary for emphasis:

Corollary 3.7. *Let* $(x_n)_{n\in\mathbb{N}}$. *Then*

- $(x_n)_{n\in\mathbb{N}}$ *is a convergent sequence if and only if*
- $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$ *and this number is finite.*

Also

- $\lim_{n\to\infty} x_n = \pm\infty$ *if and only if*
- $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = \pm\infty$

Proof of Lemma 3.6. (1) obvious from the definition

(2) Obvious from the definition of lim sup, and monotonicity of the supremum/infimum.

(3) From Lemma 3.5 we have that there exists a subsequence $(x_{n_i})_{i\in\mathbb{N}}$ such that

$$\lim_{i\to\infty} x_{n_i} = \limsup_{n\to\infty} x_n.$$

On the other hand, since x_n converges, so does any of its subsequences, so

$$\lim_{i\to\infty} x_{n_i} = \lim_{n\to\infty} x_n.$$

Together we find

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} x_n.$$

The same argument works for the lim inf.

(4) Let $a_n := \inf_{k\ge n} x_k$ and $b_n := \sup_{k\ge n} x_k$. Then

$$a_n \le x_n \le b_n \quad \forall n \in \mathbb{N}.$$

Since by assumption and Lemma 3.5,

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \limsup_{n\to\infty} x_n$$

We conclude by the squeeze theorem, Lemma 2.11 that

$$\lim_{n\to\infty} x_n = \liminf_{n\to\infty} x_n = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \limsup_{n\to\infty} x_n$$

□

FIGURE 4.1. Augustin-Louis Cauchy, 1789 – 1857. French, mathematician, engineer, and physicist. Almost singlehandedly founded complex analysis (thats why almost every theorem in complex analysis is Cauchy's theorem).

A very useful theorem (indeed a consequence of Lemma 3.5) is that in $\mathbb{R}$ every bounded sequence has a convergent subsequence[10]

Theorem 3.8 (Bolzano-Weierstrass)**.** *Suppose that $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence in $\mathbb{R}$. Then there exists a convergent subsequence.*

Proof. Since $(x_n)_{n\in\mathbb{N}}$ is bounded, $x := \limsup_{n\to\infty} x_n$ is a well-defined (finite!) number $x \in \mathbb{R}$. From Lemma 3.5 we thus know that there must be a subsequence $(x_{n_i})_{i\in\mathbb{N}}$ with $\lim_{i\to\infty} x_{n_i} = x$. □

Exercise 3.9. *Prove Corollary 3.10.*

As a consequence of Theorem 3.8 and Exercise 2.21 one obtains the following statement.

Corollary 3.10. *Assume that $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence in $\mathbb{R}$ and that there exists $x \in \mathbb{R}$ such that any convergent subsequence $(x_{n_i})_{i\in\mathbb{N}}$ converges to x. Then x_n converges to x.*

3.1. **Further (optional) exercises.**

Exercise. *Assume $(x_n)_{n\in\mathbb{N}}$ is a sequence with $x_n \neq 0$ for all but finitely many $n \in \mathbb{N}$, and such that*

$$\limsup_{n\to\infty} \left|\frac{x_n}{x_{n+1}}\right| < 1.$$

Show that $\lim_{n\to\infty} x_n = 0$.

4. Cauchy sequences

A Cauchy sequence is a sequence where *all* sequence elements eventually lie arbitrarily close to each other. This is almost as good as converging – unless there is a whole in our underlying space.

Here is the formal definition.

Definition 4.1. A sequence $(x_n)_{n\in\mathbb{N}}$ is called a *Cauchy sequence* if for any $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that

$$|x_n - x_m| < \varepsilon \quad \forall n, m > N.$$

Example 4.2. (1) The sequence

$$x_n = \text{first } n \text{ digits of } \pi$$

is a Cauchy sequence. Indeed, fix $\varepsilon > 0$ arbitrary. Let $N \in \mathbb{N}$ such that $10^{1-N} < \varepsilon$. Let $n, m \geq N$ with w.l.g. $n \leq m$. Then

$$x_n - x_m = \underbrace{0.0\ldots0}_{\text{n digits}} \quad \underbrace{\ldots}_{\text{remaining } (m-n) \text{ digits of } \pi}.$$

That is

$$|x_n - x_m| \leq 10^{1-n} \leq 10^{1-N} < \varepsilon.$$

That is, x_n is a Cauchy sequence.

Observe that the sequence $(x_n)_{n\in\mathbb{N}}$ is convergent in $\mathbb{R}$ ($\lim_{n\to\infty} x_n = \pi$) but not in $\mathbb{Q}$ (because $\pi \notin \mathbb{Q}$).

(2) *Warning:* The following is *not* an equivalent definition for a Cauchy sequence: for any $\varepsilon > 0$ there exists $N = N(\varepsilon) \in \mathbb{N}$ such that

$$|x_n - x_{n+1}| < \varepsilon \quad \forall n > N.$$

Indeed, take

$$x_n := \sum_{\ell=1}^{n} \frac{1}{\ell}.$$

We have that

$$|x_n - x_{n+1}| = \frac{1}{n+1} \xrightarrow{n\to\infty} 0.$$

However, we know from Calculus 2 that

$$\lim_{n\to\infty} \sum_{\ell=1}^{n} \frac{1}{\ell} = \sum_{\ell=1}^{\infty} \frac{1}{\ell} = \infty.$$

So $\lim_{n\to\infty} x_n$ does not exist, so by Theorem 4.4 below, $(x_n)_{n\in\mathbb{N}}$ is not a Cauchy sequence.

[10] This remains true in *finite dimensional* metric spaces (whatever that means), but becomes *false* in *infinite dimensional* spaces. Since many important spaces are infinite dimensional (e.g. function spaces), for *some* function spaces a replacement is known: weak convergence, and the theorem by Banach-Alaoglu. This generalization is part of Functional Analysis and is one of the most crucial results in Analysis.

This also can be seen explicitely: Relabelling this implies that for any $n \in \mathbb{N}$

$$\sum_{\ell=n}^{\infty} \frac{1}{\ell} = \infty.$$

This in turn (by a contradiction argument) implies that for any $n \in \mathbb{N}$ there must be an $m \in \mathbb{N}$, $m > n$ such that

$$\sum_{\ell=n}^{m} \frac{1}{\ell} \geq 1.$$

That is, for any $N \in \mathbb{N}$ and any $n \geq N$ there exists $m \geq n \geq N$ such that

$$|x_n - x_m| \not< 1.$$

That is, $(x_n)_{n\in\mathbb{N}}$ is not a Cauchy sequence.

As we said before, Cauchy sequences are almost as good as converging sequences if the underlying space is *complete* (i.e. has no holes).

First we observe that any converging sequence is necessarily Cauchy.

Lemma 4.3 (Converging sequences are Cauchy)**.** *Let $(x_n)_{n\in\mathbb{N}}$ be a converging series. Then $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence*[11].

Proof. Set $x := \lim_{n\to\infty} x_n$ (exists, because $(x_n)_{n\in\mathbb{N}}$ is converging). That is, for any $\varepsilon > 0$ there exist $N = N(\varepsilon)$ such that

$$|x_n - x| < \frac{\varepsilon}{2} \quad \forall n \geq N.$$

But then also

$$|x_n - x_m| \leq |x_n - x| + |x_m - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \forall n, m \geq N.$$

That is, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. □

As many things of this course, the notion of a Cauchy sequences lives up to its full potential in metric spaces (X, d). Metric spaces are *complete* if any Cauchy sequence has a limit (in the same space). If the metric space is not complete it has essentially an infinitesimal hole. Plugging these holes is called *metric completion*. For our purposes: $\mathbb{Q}$ is not complete, and $\mathbb{R}$ is the *metric completion* of $\mathbb{Q}$.

Theorem 4.4. *Any Cauchy sequence in $\mathbb{R}$ is convergent, and any convergent sequence is a Cauchy sequence.*

Before proving Theorem 4.4 we first show the following property (which holds in general metric spaces)

[11] can be in $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{R} \setminus \{\sqrt{2}\}$, it does not matter

Lemma 4.5. *Any Cauchy sequence is bounded*[12].

Proof. The argument is very similar to the proof of Lemma 2.5.

Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence. Then there exists an $N \in \mathbb{N}$ such that

$$|x_{N+1} - x_n| < 1 \quad \forall n > N. \tag{4.1}$$

Set

$$M := \max\{|x_1|, \dots, |x_{N+1}|\} + 1$$

Then we have

$$|x_n| \le M \quad \forall n = 1, \dots, N+1.$$

On the other hand by (4.1) we have that

$$|x_n| \le |x_n - x_{N+1}| + |x_{N+1}| < 1 + |x_{N+1}| \le M \quad \forall n > N.$$

That is, we have shown that $|x_n| \le M$ for all $n \in \mathbb{N}$. □

Now we can give

Proof of Theorem 4.4. Any converging sequence is Cauchy: This is Lemma 4.3.

Any Cauchy sequence is convergent. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence, we need to show it converges in $\mathbb{R}$. In view of Lemma 4.5 $(x_n)_{n\in\mathbb{N}}$ is bounded. By Bolzano-Weierstrass, Theorem 3.8, there exist a convergent subsequence $(x_{n_i})_{i\in\mathbb{N}}$ with

$$\lim_{i\to\infty} x_{n_i} = x. \tag{4.2}$$

Now we show that the Cauchy sequence property implies that $\lim_{n\to\infty} x_n = x$. For this let $\varepsilon > 0$ be given. By the limit property for $(x_{n_i})_{i\in\mathbb{N}}$, (4.2), there must be $N_1 \in \mathbb{N}$ such that

$$|x_{n_i} - x| < \frac{\varepsilon}{2} \quad \forall i > N_1.$$

By the Cauchy property of $(x_n)_{n\in\mathbb{N}}$ there must be another $N_2 \in \mathbb{N}$ such that

$$|x_n - x_m| < \frac{\varepsilon}{2} \quad \forall n, m > N_2.$$

Now choose $i > N_1$ such that $n_i > N_2$. Then by the above estimates,

$$|x_n - x| \le |x_{n_i} - x| + |x_n - x_{n_i}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \forall n > N_2.$$

This proves that $\lim_{n\to\infty} x_n = x$ and the proof of Theorem 4.4 is finished. □

Remark 4.6. Theorem 4.4 is sometimes called the *Cauchy criterion*: A sequence $(x_n)_{n\in\mathbb{N}}$ (in $\mathbb{R}$) is convergent if and only if $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence.

[12] again, this can be in $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{R} \setminus \{\sqrt{2}\}$, it does not matter

Exercise 4.7. *In Theorem 4.4 we have shown*

$$\text{any Cauchy sequence } (x_n)_{n\in\mathbb{N}} \subset \mathbb{R} \text{ has a limit in } \mathbb{R},$$
$$\text{i.e. there exists } x \in \mathbb{R} \text{ with } \lim_{n\to\infty} x_n = x. \tag{1}$$

The same statement is false in $\mathbb{Q}$. Namely, the following is false:

$$\text{any Cauchy sequence } (x_n)_{n\in\mathbb{N}} \subset \mathbb{Q} \text{ has a limit in } \mathbb{Q},$$
$$\text{i.e. there exists } x \in \mathbb{Q} \text{ with } \lim_{n\to\infty} x_n = x. \tag{2}$$

a) Give a counterexample to (2).
b) Which part of the proof of (1) (from Theorem 4.4) fails when we attempt to prove (2)?

Exercise 4.8. *Only using the definition of Cauchy sequence, in particular without using Theorem 4.4, show the following*

Assume the $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, and there exists a subsequence $(x_{n_i})_{i\in\mathbb{N}}$ such that $(x_{n_i})_{i\in\mathbb{N}}$ is convergent. Set

$$z := \lim_{i\to\infty} x_{n_i}.$$

Show that $(x_n)_{n\in\mathbb{N}}$ converges to z, i.e. show that

$$z = \lim_{n\to\infty} x_n.$$

Hint: You are not allowed to use the (statement) of Theorem 4.4, but you can look at the proof of Theorem 4.4, where we have essentially shown this.

The notion of Cauchy sequence is very useful to

- Check if a sequence converges in $\mathbb{R}$ (more generally complete metric spaces): Sometimes it is easier to check if a sequence is Cauchy than to guess the limit
- to "complete spaces" (this is called: *metric completion*). Let us illustrate this for the metric completion of $\mathbb{Q}$ to $\mathbb{R}$ (but this works for any metric space).

  We define

  $$\mathbb{R} = \{(x_n)_{n\in\mathbb{N}} \subset \mathbb{Q} : \quad (x_n)_{n\in\mathbb{N}} \quad \text{is Cauchy sequence}\} / \sim$$

  where $\sim$ is the equivalence relation

  $$(x_n)_{n\in\mathbb{N}} \sim (y_n)_n \quad :\Leftrightarrow \lim_{n\to\infty} |x_n - y_n| = 0.$$

  That is, we consider two sequences to be the same if their distance converges to zero.

  – $\mathbb{Q} \subset \mathbb{R}$ in the following sense: We identify an element $q \in \mathbb{Q}$ with

  $$[q] \sim \left\{(x_n)_{n\in\mathbb{N}} \subset \mathbb{Q} : \lim_{n\to\infty} x_n = q\right\}.$$

- $\mathbb{R}$ corresponds to the usual definition of $\mathbb{R}$ since any element $r \in \mathbb{R}$ can be approximated by a sequence in $(q_n)_n \subset \mathbb{Q}$, $\lim_{n\to\infty} q_n = r$. If we have two such sequences, $(q_n)_n, (s_n)_n \subset \mathbb{Q}$, $\lim_{n\to\infty} q_n = r = \lim_{n\to\infty} s_n$ then $\lim_{n\to\infty} |q_n - s_n| = 0$, so $(q_n)_n \sim (s_n)_n$.
- The fact that our new definition of $\mathbb{R}$ as Cauchy sequences of $\mathbb{Q}$ is indeed complete follows with a *diagonal argument*, we will not treat it here.
- This method is general for metric spaces (X, d) (any metric space (X, d) can be made complete by considering its Cauchy sequences). One advantage of this method is that uniformly continuous functions on (X, d) will extend uniquely to uniform continuous functions on the larger space, cf. Exercise 9.10.

Exercise 4.9. [Leb, Ex. 2.4.1] *Prove that $\left(\frac{n^2-1}{n^2}\right)$ is Cauchy using directly the definition of Cauchy sequences.*

The following result is very useful for *contraction* arguments, such as the Banach Fixed point theorem (which we will not treat in this course, but see Theorem 6.19).

Exercise 4.10. [Leb, Ex. 2.4.2] *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence such that there exists a $0 < \lambda < 1$ such that*

$$|x_{n+1} - x_n| \le \lambda |x_n - x_{n-1}|.$$

(1) Prove that $(x_n)_{n\in\mathbb{N}}$ is Cauchy.
(2) Why doesn't this contradict Example 4.2(2)?

Hint: You can freely use the formula (for $\lambda \neq 1$)

$$1 + \lambda + \lambda^2 + \cdots + \lambda^n = \frac{1 - \lambda^{n+1}}{1 - \lambda}.$$

4.1. Optional exercises.

Exercise. *Give an example of a bounded sequence that is not a Cauchy sequence.*

Exercise. *Show that*

$$\sum_{i=1}^{n} \frac{1}{n!}$$

is a Cauchy sequence (directly from the definition of Cauchy sequence)

5. Limits of functions

Remark 5.1 (Further reading)**.**
- Interactive picture example of ε-δ limit operation https://www.desmos.com/calculator/4efsywgvtg

We want to describe how a function f behaves near to a point $c \in \mathbb{R}$, i.e. we would like to give a notion for

$$\lim_{x \to c} f(x).$$

To do this, f does not need to be defined at c, but it needs to be defined "close to c".

Definition 5.2. Let $D \subset \mathbb{R}$ be a set.

- We define $\overline{D}$ the closure of the set D as follows
$$\overline{D} := \left\{ x \in \mathbb{R} : \quad \exists (x_n)_{n\in\mathbb{N}} \subset D \ \lim_{n\to\infty} x_n = x \right\}$$
That is $\overline{D}$ are all points (in $\mathbb{R}$) that can be approximated by sequences *from within D*.
- A set $D \subset \mathbb{R}$ is *closed*, if $D = \overline{D}$.
- A set $D \subset \mathbb{R}$ is *open*, if $\mathbb{R} \backslash D$ is closed[13]
- While not so relevant for our purposes, let us also define the *boundary* of a set D, usually denoted by ∂D,
$$\partial D = \overline{D} \cap \left(\mathbb{R}^n \setminus \bar{D} \right).$$
Equivalently ∂D is the set of all points x such that there exists a sequence $(y_n)_{n\in\mathbb{N}} \subset D$ and another sequence $(z_n)_{n\in\mathbb{N}} \subset \mathbb{R}^n \setminus D$ such that $x = \lim_{n\to\infty} y_n = \lim_{n\to\infty} z_n$. That is ∂D are the points that can be approximated from both within D, and from within the complement of D, $\mathbb{R}^n \setminus D$.
- A point $c \in \overline{D}$ is a *cluster point* of D, if there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset D \backslash \{c\}$ with $\lim_{n\to\infty} x_n = c$.
  That is, a point c is a cluster point of D if it can approximated by points within D *different from c itself*.

Exercise 5.3. *Show that*

a) the set $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$ has no cluster points.
b) every point in $\mathbb{R}$ is a cluster point of $\mathbb{Q}$.

Exercise 5.4. *The empty set $\emptyset$ is both open and closed. So is $\mathbb{R} = \mathbb{R} \setminus \emptyset$.*

Exercise 5.5. *$\overline{\mathbb{Q}} = \mathbb{R}$, $\mathbb{Q}$ is neither open nor closed*

Lemma 5.6. *We always have $D \subset \overline{D}$*

Proof. For any $x \in D$, the sequence $(x_n)_{n\in\mathbb{N}} := (x, x, x, x, \ldots)$ clearly converges to x, $\lim_{n\to\infty} x_n = x$, so $x \in \overline{D}$. □

Exercise 5.7. *Let $D \subset \mathbb{R}$ be a set. Show that $\bar{D}$ is closed, i.e. that $\bar{\bar{D}} = \bar{D}$.*

[13]Below we will see a equivalent but nicer/more intuitive definition of open: a set D is open if around any point $x \in D$ a whole neighborhood of that point belongs to D

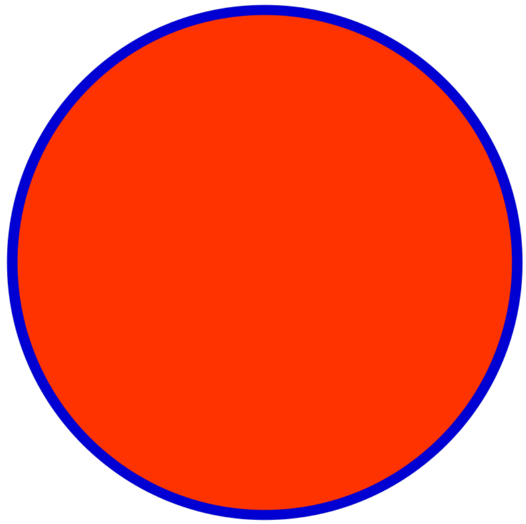

FIGURE 5.1. The blue circle represents the set of points (x, y) satisfying $x^2 + y^2 = 1$. The red disk represents the set of points (x, y) satisfying $x^2 + y^2 < 1$. The red set is an open set, the blue set is its boundary set, and the union of the red and blue sets is a closed set, the closure of the open set. (from: wikipedia)

Example 5.8. $\overline{(1,2)} = [1,2]$

Proof. Indeed, we already know $(1,2) \subset \overline{(1,2)}$. Now $1 \in \overline{(1,2)}$ because $x_n := 1 + \frac{1}{n} \in (1,2)$ for any $n \in \mathbb{N}$ and $\lim_{n\to\infty} x_n = 1$. Similar argument for 2. If $x \notin [1,2]$ then there must be a $\delta > 0$ such that $x < 1-\delta$ or $x > 2+\delta$. Now for any sequence $(x_n)_{n\in\mathbb{N}}$ with $\lim_{n\to\infty} x_n = x$ there exists $N \in \mathbb{N}$ such that

$$|x_n - x| < \frac{\delta}{2} \quad \forall n \geq N.$$

But then if $x < 1 - \delta$ we have

$$x_n < x_n - x + x < 1 - \delta + \frac{\delta}{2} = 1 - \frac{\delta}{2} < 1, \quad \forall n \geq N$$

or if $x > 2 + \delta$,

$$x_n > x_n - x + x > 2 + \delta - \frac{\delta}{2} = 2 + \frac{\delta}{2} > 2, \quad \forall n \geq N.$$

That is, in either case $x_n \notin (1,2)$ for all $n \geq N$. That means there is no sequence $(x_n)_{n\in\mathbb{N}} \in (1,2)$ such that $\lim_{n\to\infty} x_n = x$ if $x \notin [1,2]$ □

Exercise 5.9. $[1,2]$ *is closed, all points are cluster points.*

Lemma 5.10. *A set $D \subset \mathbb{R}$ is open, if and only if for any $x_0 \in D$ there exists $\varepsilon > 0$ such that $(x_0 - \varepsilon, x_0 + \varepsilon) \subset D$ (in words: any point $x_0 \in D$ has a small neighborhood also belonging to D).*

Proof. Assume that D is open and $x_0 \in D$, but for any $\varepsilon > 0$ there exists a point $x_\varepsilon \in (x_0 - \varepsilon, x_0 + \varepsilon)\backslash D$. Choosing $\varepsilon := \frac{1}{n}$ we then find a sequence $x_n \in (x_0 - \frac{1}{n}, x_0 + \frac{1}{n})\backslash D$. That is $|x_n - x_0| \xrightarrow{n\to\infty} 0$. That is $\lim_{n\to\infty} x_n = x_0$. On the other hand, $x_n \in \mathbb{R}\backslash D$ which is closed by assumption, so $x_0 \in \mathbb{R}\backslash D$. Contradiction to $x_0 \in D$.

For the other direction, assume that D is such that for any $x_0 \in D$ there exists $\varepsilon > 0$ such that $(x_0 - \varepsilon, x_0 + \varepsilon) \subset D$. Let $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}\backslash D$ be a converging sequence and set $x_0 := \lim_{n\to\infty} x_n$. We need to show $x_0 \in \mathbb{R}\backslash D$. To the contrary assume that $x_0 \in D$. By assumption there exists $\varepsilon > 0$ such that $(x_0 - \varepsilon, x_0 + \varepsilon) \subset D$, which means that $x_n \notin (x_0 - \varepsilon, x_0 + \varepsilon)$. But this means that $|x_n - x_0| \geq \varepsilon$ for all $n \in \mathbb{N}$, i.e. x_n does not converge to x_0. Contradiction. □

Lemma 5.11. *Let D be an open set, then any point $x_0 \in \overline{D}$ is a clusterpoint of D (and $\overline{D}$).*

Proof. Indeed, let $x_0 \in \overline{D}$. If $x_0 \notin D$ there is no problem: by definition there must be a sequence $(x_n)_{n\in\mathbb{N}} \subset D = D\backslash\{x_0\}$ with $\lim_{n\to\infty} x_n = x_0$.

If $x_0 \in D$, then we construct a sequence $(x_n)_{n\in\mathbb{N}} \subset D\backslash\{x_0\}$ as follows. Since D is open and $x_0 \in D$, there must be some $\varepsilon > 0$ such that $(x_0 - \varepsilon, x_0 + \varepsilon) \subset D$. Let N be such that $\frac{1}{N} < \varepsilon$. Then take x_n any element from $(x_0 - \frac{1}{N+n}, x_0 + \frac{1}{N+n})\backslash\{x_0\} \subset D\backslash\{x_0\}$. It is easy to show that $\lim_{n\to\infty} x_n = x_0$. □

Now we want to define

$$\lim_{x\to c} f(x)$$

There are issues to deal with: $x \to c$ is not a sequence, second $f(c)$ may not be defined.

Definition 5.12. Let $f : D \subset \mathbb{R} \to \mathbb{R}$ be a function and $c \in \overline{D}$ be a cluster point of D. We say that the *limit as $x \ni D$ approaches c of f* is a number $L \in \mathbb{R}$,

$$\lim_{x\to c} f(x) = L, \quad \text{or} \quad f(x) \xrightarrow{x\to c} L.$$

if [14]

$$\forall \varepsilon > 0 \quad \exists \delta = \delta(\varepsilon) > 0 : \quad \text{s.t. } |f(x) - L| < \varepsilon \quad \forall x \in D\backslash\{c\} : \quad |x - c| < \delta.$$

Cf. Figure 5.2.

Example 5.13. Let

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

Then the domain of f is $(-\infty, 0) \cup (0, \infty)$ and $\lim_{x\to 0} f(x)$ does not exist.

On the other hand, if we consider f as a function $f : (-\infty, 0) \to \mathbb{R}$ then $\lim_{x\to 0} f(x) = 1$.

[14] since c is a cluster point of D, the set $D\backslash\{c\} : \quad |x - c| < \delta$ is nonempty

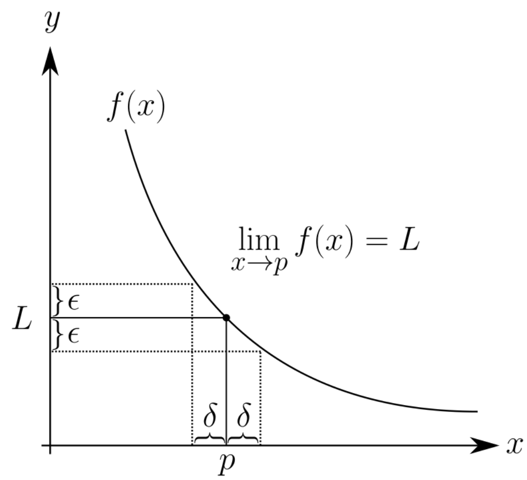

FIGURE 5.2. The limit of a function f for x to p is equal to L, $\lim_{x\to p} f(x) = L$ if and only if for any $\varepsilon > 0$ there exists $\delta > 0$ such that for all x with $0 < |x - p| < \delta$ we have $|f(x) - L| < \varepsilon$.
picture: Johannes Schneider , wikipedia.

First we show that if the limit exists, the limit is unique:

Lemma 5.14. *Let $f : D \subset \mathbb{R} \to \mathbb{R}$ be a function and $c \in \overline{D}$ be a cluster point. If $L_1, L_2 \in \mathbb{R}$ with $\lim_{x\to c} f(x) = L_1$ and $\lim_{x\to c} f(x) = L_2$ then $L_1 = L_2$.*

Proof. By the definition of a limit, for any $\varepsilon > 0$ there must be $\delta = \delta(\varepsilon) > 0$ such that

$$\max\{|f(x) - L_1|, |f(x) - L_2|\} < \frac{\varepsilon}{2} \quad \forall x \in D\backslash\{c\} : \quad |x - c| < \delta.$$

But then, if we pick any point $x \in D$, $x \neq c$, such that $|x - c| < \delta$ (this point must exist, since c is a cluster point)

$$|L_1 - L_2| \leq |f(x) - L_1| + |f(x) - L_2| < 2\frac{\varepsilon}{2} = \varepsilon$$

We can do this for any $\varepsilon > 0$, so $|L_1 - L_2| < \varepsilon$ for any $\varepsilon > 0$. By the Archimedian principle this means that $L_1 = L_2$. □

Recall again $x \to c$ doesnt make too much sense, since x is not a sequence. The meaning of $x \to c$ is: "take any possible seqeunce $(x_n)_{n\in\mathbb{N}}$ converging to c" (but no sequence element equal to c). More precisely, we have

Lemma 5.15 (sequential limits)**.** *Let $f : D \subset \mathbb{R} \to \mathbb{R}$ and $c \in \overline{D}$ be a cluster point of D. Then the following are equivalent for any $L \in \mathbb{R}$*

(1) $\lim_{x\to c} f(x) = L$
(2) for any sequence $(x_n)_{n\in\mathbb{N}} \subset D\backslash\{c\}$ with $\lim_{n\to\infty} x_n = c$ we have that the sequence $(f(x_n))_{n\in\mathbb{N}}$ is convergent to L, i.e. $\lim_{n\to\infty} f(x_n) = L$.

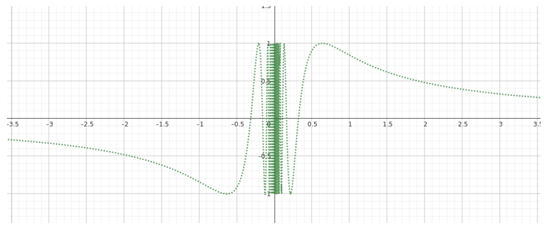

FIGURE 5.3. The graph of the function $f(x) = \sin(1/x)$

Proof. (1) ⇒ (2): Assume that $\lim_{x\to c} f(x) = L$, and let $(x_n)_{n\in\mathbb{N}} \subset D\backslash\{c\}$ with $\lim_{n\to\infty} x_n = c$. We need to show that $\lim_{n\to\infty} f(x_n) = L$. For this let $\varepsilon > 0$ be arbitrary. We need to find $N = N(\varepsilon)$ such that

$$|f(x_n) - L| < \varepsilon \quad \forall n \geq N.$$

Since by assumption $\lim_{x\to c} f(x) = L$ there must be $\delta > 0$ such that

$$|f(x) - L| < \varepsilon \quad \forall x \in D\backslash\{c\} : \ |x - c| < \delta.$$

Moreover, since $\lim_{n\to\infty} x_n = c$, for this δ there must be an $N = N(\delta)$ such that

$$|x_n - c| < \delta \quad \forall n \geq N.$$

So in particular, for any $n \geq N$ we have $|f(x_n) - L| < \varepsilon$, which is what we needed to show.

(2) ⇒ (1):

Assume that $\lim_{n\to\infty} f(x_n) = L$ holds for any sequence $(x_n)_{n\in\mathbb{N}} \subset D\backslash\{c\}$ with $\lim_{n\to\infty} x_n = c$. We need to show that $\lim_{x\to c} f(x) = L$, that is

$$\forall \varepsilon > 0\, \exists \delta > 0 : |f(x) - L| < \varepsilon \quad \forall x \in D\backslash\{c\}, |x - c| < \delta.$$

Assume this is not the case, then the logical negation is

$$\exists \varepsilon > 0\, \forall \delta > 0\, \exists x \in D\backslash\{c\} \text{ with } |x - c| < \delta \text{ but such that } |f(x) - L| > \varepsilon.$$

We can apply the above to $\delta := \frac{1}{n}$ for each $n \in \mathbb{N}$. Then for some $\varepsilon > 0$ fixed, we find for each $n \in \mathbb{N}$ some $x_n \in D\backslash\{c\}$ with $|x_n - c| < \frac{1}{n}$ but $|f(x_n) - L| > \varepsilon$.

The sequence $(x_n)_{n\in\mathbb{N}} \subset D\backslash\{c\}$ then converges, $\lim_{n\to\infty} x_n = c$. By assumption, this implies that $\lim_{n\to\infty} f(x_n) = L$, which contradicts that $|f(x_n) - L| > \varepsilon$ holds for any $n \in \mathbb{N}$. □

Example 5.16.
- $\lim_{x\to 0} \sin(1/x)$ does not exist, cf. Figure 5.3. Indeed take the sequence $x_n := \frac{1}{\pi n + \pi/2}$. Then $\sin(x_n) = (-1)^n$, $\lim_{n\to\infty} x_n = 0$, but $\lim_{n\to\infty} \sin(x_n)$ does not exist - Lemma 5.15 implies the limit of $\sin(1/x)$ cannot exist.
- $\lim_{x\to 0} x\sin(1/x) = 0$, cf. Figure 5.4: Indeed,

  $$|x\sin(1/x)| \leq |x|.$$

  So for any $\varepsilon > 0$ if we choose $\delta := \varepsilon$ we have

  $$|x\sin(1/x) - 0| \leq |x| < \varepsilon \quad \forall |x| < \delta.$$

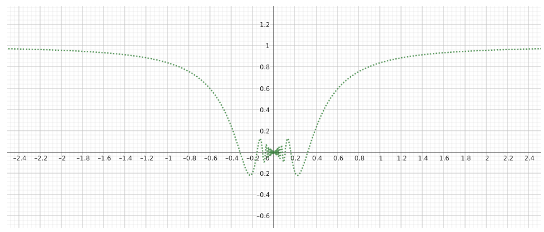

FIGURE 5.4. The graph of the function $f(x) = x\sin(1/x)$

Exercise 5.17. [Leb, Ex. 3.1.9]*: Let c_1 be a cluster point of $A \subset \mathbb{R}$ and c_2 be a cluster point of $B \subseteq \mathbb{R}$. Suppose that $f : A \to B$ and $g : B \to \mathbb{R}$ are functions such that $f(x) \to c_2$ as $x \to c_1$ and $g(y) \to L$ as $y \to c_2$. Let $h(x) := g(f(x))$ and show $h(x) \to L$ as $x \to c_1$.*

Since we know from Lemma 5.15 that the limit of a function f can be described as the limit of sequence $f(x_n)$ we can deduce the limit laws from the sequential limit laws.

Corollary 5.18. *Let $D \subset \mathbb{R}$ and $c \in \overline{D}$ a cluster point of D. Let $f, g, h : D \to \mathbb{R}$ be functions.*

(1) If
$$f(x) \le g(x) \quad \text{for all } x \in D,$$
then if $\lim_{x\to c} f(x)$ and $\lim_{x\to c} g(x)$ both exist we have
$$\lim_{x\to c} f(x) \le \lim_{x\to c} g(x).$$
(2) If for some $a, b \in \mathbb{R}$ we have
$$a \le f(x) \le b \quad \text{for all } x \in D,$$
then if $\lim_{x\to c} f(x)$ exists we have
$$a \le \lim_{x\to c} f(x) \le b$$
(3) If
$$f(x) \le g(x) \le h(x) \quad \text{for all } x \in D,$$
then if $\lim_{x\to c} f(x) = \lim_{x\to c} h(x)$ (i.e. they both exist and are equal) then $\lim_{x\to c} g(x)$ exists and we have
$$\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = \lim_{x\to c} h(x)$$

Proof. (1) In view of (5.15) for any sequence $(x_n)_{n\in\mathbb{N}} \subset D\backslash\{c\}$ with $\lim_{n\to\infty} x_n = c$ we have
$$\lim_{n\to\infty} f(x_n) = \lim_{x\to c} f(x), \quad \lim_{n\to\infty} g(x_n) = \lim_{x\to c} g(x).$$

On the other hand $f(x) \leq g(x)$ for all $x \in D$ implies $f(x_n) \leq g(x_n)$ for all $n \in \mathbb{N}$, and by the limit laws for sequences (monotonicity of the limit):

$$\lim_{x \to c} f(x) = \lim_{n \to \infty} f(x_n) \leq \lim_{n \to \infty} g(x_n) \leq \lim_{x \to c} g(x).$$

(2) Follows from (1): E.g. taking $g(x) := b$, observing that $\lim_{x \to c} g(x) = b$ we conclude from $f(x) \leq b$ for all $x \in D$ that

$$\lim_{x \to c} f(x) \leq b.$$

In a similar way we conclude from $f(x) \geq a$ for all $x \in D$ that

$$\lim_{x \to c} f(x) \geq b.$$

(3) This is a consequence of the squeeze lemma, Lemma 2.11. Let $(x_n)_{n \in \mathbb{N}} \subset D \backslash \{c\}$ be an arbitrary sequence with $\lim_{n \to \infty} x_n = c$.

If we set $a_n := f(x_n)$, $b_n := g(x_n)$, $c_n := h(x_n)$ and

$$\Gamma := \lim_{x \to c} f(x) = \lim_{x \to c} h(x),$$

then we have by assumption (and Lemma 5.15)

$$a_n \leq b_n \leq c_n \quad \text{for all } n \in \mathbb{N},$$

and

$$\Gamma = \lim_{n \to \infty} a_n = \lim_{n \to \infty} c_n.$$

By the squeeze lemma, Lemma 2.11,

$$\lim_{n \to \infty} b_n = \Gamma.$$

Thus, we have shown for any sequence $(x_n)_{n \in \mathbb{N}} \subset D \backslash \{c\}$ with $\lim_{n \to \infty} x_n = c$ that

$$\lim_{n \to \infty} g(x_n) = \Gamma.$$

By Lemma 5.15 we conclude that

$$\lim_{x \to c} g(x) = \Gamma = \lim_{x \to c} f(x) = \lim_{x \to c} h(x).$$

□

From Lemma 5.15 we also obtain that the usual limit laws hold for $\lim_{x \to c}$-operation:

Corollary 5.19. *Let $D \subset \mathbb{R}$ and $c \in \overline{D}$ a cluster point of D. Let $f, g : D \to \mathbb{R}$ be functions. Suppose that $\lim_{x \to c} f(x)$ and $\lim_{x \to c} g(x)$ both exist. Then*

(1) $\lim_{x \to c} (f(x) + g(x)) = (\lim_{x \to c} f(x)) + (\lim_{x \to c} g(x))$.
(2) $\lim_{x \to c} (f(x) - g(x)) = (\lim_{x \to c} f(x)) - (\lim_{x \to c} g(x))$.
(3) $\lim_{x \to c} (f(x)g(x)) = (\lim_{x \to c} f(x)) \, (\lim_{x \to c} g(x))$.
(4) If $g(x) \neq 0$ for all $x \in D$ and $\lim_{x \to c} g(x) \neq 0$ then

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \frac{\lim_{x \to c} f(x)}{\lim_{x \to c} g(x)}.$$

Exercise 5.20. *Prove Corollary 5.19.*

To compute limits $\lim_{n\to\infty} x_n$ we only care about all but finitely many sequence elements of $(x_n)_{n\in\mathbb{N}}$. A similar statement same is true for for $\lim_{x\to c} f(x)$: we only care about points x "close" to c, that is it suffices to consider f *restricted to a small (open) neighborhood of c*.

Definition 5.21. Let $f : D \to \mathbb{R}$ be a function and let $D_2 \subset D$. The function *f restricted to D_2*, $f\Big|_{D_2}$, is the function

$$f\Big|_{D_2} : D_2 \to \mathbb{R}$$

$$f\Big|_{D_2} : x \ni D_2 \mapsto f(x).$$

Lemma 5.22. *Let $D_2 \subset D \subset \mathbb{R}$, $c \in \overline{D} \cap \overline{D_2}$ be a cluster point of D and D_2. Let $f : D \to \mathbb{R}$ and let $f\Big|_{D_2} : D_2 \to \mathbb{R}$ be its restriction to D_2*

(1) If $\lim_{x\to c} f(x)$ exists then $\lim_{x\to c} f\Big|_{D_2}$ exist, and

$$\lim_{x\to c} f\Big|_{D_2} = \lim_{x\to c} f(x).$$

(2) If $\lim_{x\to c} f\Big|_{D_2}$ exists, in general $\lim_{x\to c} f$ may not exist.

(3) Assume that D_2 contains a relative open *neighborhood of c in D. That is, assume there exists $\varepsilon > 0$ such that $(c-\varepsilon, c+\varepsilon) \cap D \subset D_2$. Then $\lim_{x\to c} f\Big|_{D_2}$ exists if and only if also $\lim_{x\to c} f$ exists. Also if one of the limits exists, we have*

$$\lim_{x\to c} f\Big|_{D_2} = \lim_{x\to c} f(x).$$

Proof. (1) Let $(x_n)_{n\in\mathbb{N}} \subset D_2 \backslash \{c\}$ be any sequence with $\lim_{n\to\infty} x_n = c$. Since $D_2 \subset D$ we also have $(x_n)_{n\in\mathbb{N}} \subset D\backslash\{c\}$, and thus by assumption and Lemma 5.15,

$$\lim_{n\to\infty} f(x_n) = \lim_{x\to c} f(x).$$

Since for any $n \in \mathbb{N}$ we have $x_n \in D_2\backslash\{c\}$,

$$f(x_n) = f\Big|_{D_2}(x_n)$$

and consequently we have

$$\lim_{n\to\infty} f\Big|_{D_2}(x_n) = \lim_{n\to\infty} f(x_n) = \lim_{x\to c} f(x).$$

This holds for any sequence $(x_n)_{n\in\mathbb{N}} \subset D_2\backslash\{c\}$ with $\lim_{n\to\infty} x_n = c$, so again by Lemma 5.15

$$\lim_{x\to c} f\Big|_{D_2}(x) \lim_{n\to\infty} f\Big|_{D_2}(x_n) = \lim_{n\to\infty} f(x_n) = \lim_{x\to c} f(x).$$

(2) The typical example is the socalled Heaviside function,

$$f(x) := \begin{cases} 0 & x \le 0 \\ 1 & x > 0 \end{cases}$$

It is easy to show (exercise) that $\lim_{x\to 0} f(x)$ does not exist. However if we consider $f\Big|_{(-\infty,0)}$ then

$$f\Big|_{(-\infty,0)}(x) = 0,$$

so

$$\lim_{x\to 0} f\Big|_{(-\infty,0)} = 0.$$

(3) In (1) we have shown that if $\lim_{x\to c} f$ exists, then also $\lim_{x\to c} f\Big|_{D_2}$ exists and the two numbers are the same.

For the converse, assume that $\lim_{x\to c} f\Big|_{D_2}$ exists. Let $(x_n)_{n\in\mathbb{N}} \subset D\backslash\{c\}$ be a sequence with $\lim_{n\to\infty} x_n = c$. Since x_n converges to c, there must be a large index $N = N(\alpha) \in \mathbb{N}$ such that

$$|x_n - c| < \alpha \quad \forall n > N.$$

That is,

$$x_n \in (c-\alpha, c+\alpha) \cap D \subset D_2 \quad \forall n > N. \tag{5.1}$$

Set

$$(z_n)_{n\in\mathbb{N}} := (\underbrace{x_{N+1}, \ldots, x_{N+1}}_{N+1 \text{ times}}, x_{N+2}, x_{N+3}, \ldots)$$

Then from (5.1) we deduce that $(z_n)_{n\in\mathbb{N}} \subset (D \cap D_2)\backslash\{c\}$ and we have

$$\lim_n z_n = \lim_{n\to\infty} x_n = c.$$

Thus

$$f(z_n) = f\Big|_{D_2}(z_n) \xrightarrow{n\to\infty} \lim_{x\to c} f\Big|_{D_2}(z_n).$$

In other words,

$$\lim_{n\to\infty} f(z_n) = \lim_{x\to c} f\Big|_{D_2}(z_n).$$

Since $z_n = x_n$ (and thus $f(z_n) = f(x_n)$) for all but finitely many $n \in \mathbb{N}$ we have

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} f(z_n) = \lim_{x\to c} f\Big|_{D_2}(z_n).$$

The above holds for any sequence $(x_n)_{n\in\mathbb{N}} \subset D\backslash\{c\}$ with $\lim_{n\to\infty} x_n = c$. By Lemma 5.15 we conclude that

$$\lim_{x\to c} f(x) = \lim_{x\to c} f\Big|_{D_2}(z_n).$$

□

Exercise 5.23. *Use the precise ε-δ-definition of limit to prove*

$$\lim_{x\to 2} x^2 = 4.$$

In particular the *one-sided limits* are defined exactly the same way as in calculus:

Definition 5.24. Assume that $f : (a, b) \to \mathbb{R}$ then for $c \in [a, b)$ we set

$$\lim_{x\to c^+} f := \lim_{x\to c} f\Big|_{(c,b)}$$

and for $c \in (a, b]$ we set

$$\lim_{x\to c^-} f := \lim_{x\to c} f\Big|_{(a,c)}$$

Further (optional) exercises. Computing limits of functions is also very important, so here some practice exams

Exercise. *Use the precise ε-δ-definition of limit to prove the following statements.*

(1) $\lim_{x\to 10}(2x + 4) = 24$
(2) $\lim_{x\to -\frac{3}{2}}(1 - 4x) = 7$
(3) $\lim_{x\to 1}(x^2 + 3) = 4$
(4) $\lim_{x\to 3} \frac{2}{x+3} = \frac{1}{3}$
(5) $\lim_{x\to -6} \frac{x+4}{2-x} = -\frac{1}{4}$
(6) $\lim_{x\to 9}(\sqrt{x} + 2) = 5.$
(7) $\lim_{x\to 1} \frac{2+4x}{3} = 2$
(8) $\lim_{x\to -2} x^2 - 1 = 3$
(9) $\lim_{x\to 2} x^3 = 8$

6. Continuous functions

A function $f : D \to \mathbb{R}$ is *continuous* is small changes in the domain $x \in D$ imply small changes in the target $f(x)$.

Here is the precise definition of continuous functions that we are going to use for the rest of our (mathematical) life.

Definition 6.1 (ε-δ-definition)**.** Let $D \subset \mathbb{R}$ be a set and let $f : D \to \mathbb{R}$ be a function.

- f is *continuous* at a point $x_0 \in D$, if

$$\forall \varepsilon > 0\, \exists \delta = \delta(\varepsilon) > 0 : \quad |f(x) - f(x_0)| < \varepsilon \text{ holds whenever } x \in D \text{ and} |x - x_0| < \delta.$$

- f is *continuous in D* if f is continuous at any point $x_0 \in D$.

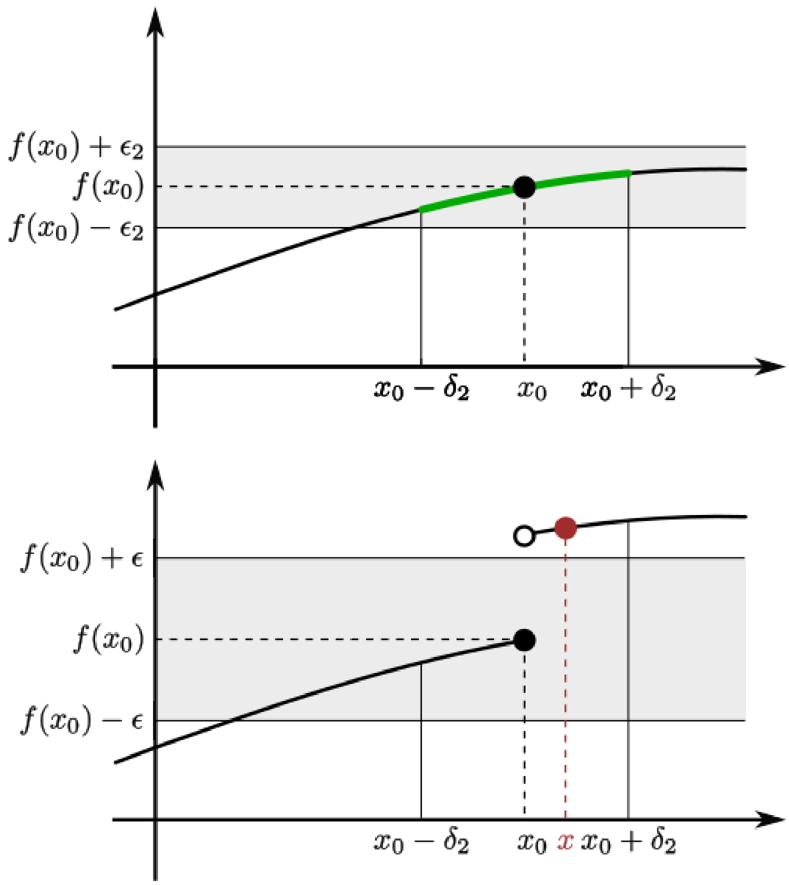

FIGURE 6.1. ε-δ-definition of continuity at x_0: For any $\varepsilon > 0$ we must be able to find a δ such that the function values $f(x)$ are ε-close to $f(x_0)$ for any x which is δ-close to x_0. In the first picture this works for any ε. In the second one this does not work at a jump discontinuity. Pictures: Stephan Kulla (User:Stephan Kulla), CC0, via Wikimedia Commons

Cf. Figure 6.1.

Exercise 6.2. [Leb, Ex. 3.2.1] *Use the definition of continuity from Definition 6.1 to prove that $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) := x^2$ is continuous.*

Functions are automatically continuous at "discrete" points, namely we have

Exercise 6.3. *Use the definition of continuity from Definition 6.1 to prove that if $f : D \to \mathbb{R}$ and $c \in D$ is not a cluster point of D, then f is continuous at c.*

Exercise 6.4. [Leb, Ex. 3.2.3] *Let $f : \mathbb{R} \to \mathbb{R}$ by defined by*

$$f(x) = \begin{cases} x & \text{if } x \text{ is rational} \\ x^2 & \text{if } x \text{ is irrational.} \end{cases}$$

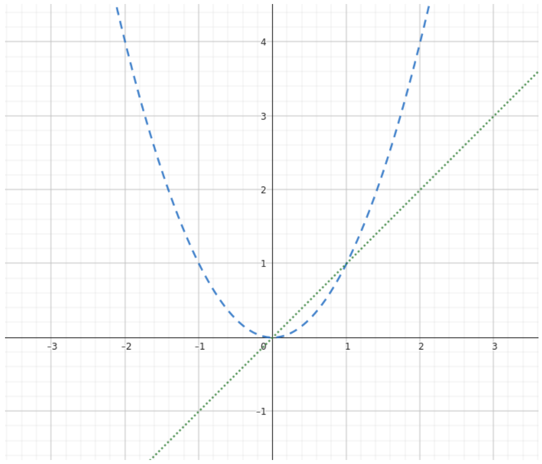

FIGURE 6.2. The function from Exercise 6.4

Cf. Figure 6.2. Using the definition of continuity from Definition 6.1, prove that f is continuous at 1 and discontinuous at 2.

Definition 6.1 is not the definition we have from Calculus 1 (which, we recall, was that f is continuous at x_0 if $\lim_{x\to x_0} f(x) = f(x_0)$). But it is very related.

Proposition 6.5 (Continuity via limits)**.** *Let $f : D \to \mathbb{R}$ be a function.*

(1) f is continuous at $x_0 \in D$ if and only if for any sequence $(x_n)_{n\in\mathbb{N}} \subset D$ with $\lim_{n\to\infty} x_n = x_0$ we have $\lim_{n\to\infty} f(x_n) = f(x_0)$ (the latter is called sequential continuity).
(2) For any $x_0 \in D$ which is not a cluster point of D, we have that f is continuous at x_0.
(3) Let $x_0 \in D$ and x_0 is a cluster point[15] of D. Then f is continuous at x_0 if and only if $\lim_{x\to x_0} f(x) = f(x_0)$.

Proof. (1) Step 1: continuity implies sequential continuity:

Let f be continuous at $x_0 \in D$. Take any sequence $(x_n)_{n\in\mathbb{N}} \in D$ with $\lim_{n\to\infty} x_n = x_0$. We need to show that $\lim_{n\to\infty} f(x_n) = f(x_0)$.

That is we need to show that for any $\varepsilon > 0$ there exists $N = N(\varepsilon) > 0$ such that

$$|f(x_n) - f(x_0)| < \varepsilon \quad \forall n > N.$$

From the definition of continuity, since f is continuous at x_0: there must be some $\delta > 0$ be such that

$$|f(x) - f(x_0)| < \varepsilon \quad \text{whenever } x \in D \text{ and } |x - x_0| < \delta.$$

[15]the point of this assumption is: the notion $\lim_{x\to x_0} f(x) = f(x_0)$ is not defined if x_0 is not a cluster point

On the other hand, since $\lim_{n\to\infty} x_n = x_0$ there must be some $N = N(\delta) > 0$ such that

$$|x_n - x_0| < \delta \quad \forall n > N.$$

Consequently,

$$|f(x_n) - f(x_0)| < \varepsilon \quad \forall n > N.$$

which implies that $\lim_{n\to\infty} f(x_n) = f(x_0)$.

Step 2: Sequential continuity implies continuity:

Assume that for any sequence $(x_n)_{n\in\mathbb{N}} \subset D$ with $\lim_{n\to\infty} x_n = x_0$ we have $\lim_{n\to\infty} f(x_n) = f(x_0)$. We need to show that f is continuous, i.e. that

$$\forall \varepsilon > 0\, \exists \delta > 0 : \quad |f(x) - f(x_0)| < \varepsilon \quad \forall x \in D : |x - x_0| < \delta. \tag{6.1}$$

Assume this is not the case, i.e. that (6.1) is false. Then (by logical negation of (6.1)):

$$\exists \varepsilon > 0\, \forall \delta > 0 : \quad |f(x) - f(x_0)| \geq \varepsilon \quad \text{for some } x = x_\delta \in D \text{ with } |x - x_0| < \delta.$$

Apply this statement (for this ε) to $\delta = \frac{1}{n}$ then for any $n \in \mathbb{N}$ we find a point $x_n \in D$ with $|x_n - x_0| < \frac{1}{n}$ but $|f(x_n) - f(x_0)| > \varepsilon$. That is $\lim_{n\to\infty} x_n = x_0$ but $\lim_{n\to\infty} f(x_n) \neq f(x_0)$. This contradicts the assumption that f is sequentially continuous at x_0. Thus (6.1) could not have been false so (6.1) must have been true all along.

(2) Let $x_0 \in D$ which is *not* a cluster point of D. That is assume there is no sequence $(x_n)_{n\in\mathbb{N}} \in D\backslash\{x_0\}$ with $\lim_{n\to\infty} x_n = x_0$. This means[16] there exists a $\delta > 0$ we have that $(x_0 - \delta, x_0 + \delta) \subset \{x_0\} \cup \mathbb{R}\backslash D$. In other words, there exists a $\delta > 0$ such that if $|x - x_0| < \delta$ *and* $x \in D$ then $x = x_0$.

That is, for any $\varepsilon > 0$

$$\forall x \in D\, |x - x_0| < \delta : \quad |f(x) - f(x_0)| = |f(x_0) - f(x_0)| = 0 < \varepsilon.$$

That is, f is continuous at x_0 (in a very pathological way).

(3) "$\Rightarrow$" follows from (1) and Lemma 5.15.

"$\Leftarrow$": Assume that $\lim_{x\to x_0} f(x) = f(x_0)$. That is,

$$\forall \varepsilon > 0 : \quad \exists \delta > 0 : \quad |f(x) - f(x_0)| < \delta \quad x \in D\backslash\{x_0\},\ |x - x_0| < \delta.$$

Clearly $|f(x_0) - f(x_0)| = 0$, so the above is equivalent to

$$\forall \varepsilon > 0 : \quad \exists \delta > 0 : \quad |f(x) - f(x_0)| < \delta \quad x \in D,\ |x - x_0| < \delta.$$

But this is the definition of continuity at x_0.

□

Example 6.6.

- $f(x) = 1/x$ is continuous in $(0, \infty)$, also $(-\infty, 0) \cup (0, \infty)$ but clearly not in $(-1, 1)$.
- Any map $f : \mathbb{N} \subset \mathbb{R} \to \mathbb{R}$ is continuous.
- Any continuous map $f : \mathbb{R} \to \mathbb{Z}$ is constant. (Exercise 6.15)

[16] good exercise!

- Polynomials are continuous (we can prove that now with the limit definition!)
- If $g : D \to \mathbb{R}$ is continuous at x_0 and $g(x_0) \neq 0$. Then there exists $\delta > 0$ such that $g(x) \neq 0$ for all $x \in D$, $|x - x_0| > \delta$.
  Indeed: Set $\Gamma := g(x_0) \neq 0$. For $\varepsilon := \frac{1}{2}|\Gamma|$ there must be a $\delta > 0$ such that
  $$|g(x) - g(x_0)| < \varepsilon \quad \forall |x - x_0| < \delta, \quad x \in D.$$
  and thus
  $$|g(x)| \geq |g(0)| - |g(x) - g(0)| \geq \Gamma - \frac{\Gamma}{2} = \frac{\Gamma}{2} \quad \forall |x - x_0| < \delta, \quad x \in D.$$
- $f, g : D \to \mathbb{R}$ continuous at x_0 and $g(x_0) \neq 0$ then $\frac{f}{g}$ continuous at x_0 (with D a small neighborhood of x_0), cf. Corollary 5.19.
- Let $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ be continuous functions with $g(B) \subset A$. Then $f \circ g$ is continuous. (cf. Exercise 5.17)

Exercise 6.7. *Assume $f, g : D \to \mathbb{R}$ continuous at x_0.*

Use each of the two definitions of continuity that we had for now, namely

(a) *the ε-δ-definition of continuity*
(b) *the sequential definition of continuity*

to show that

(1) *$f + g$ is continuous continuous at x_0*
(2) *fg is continuous at x_0*

Another equivalent definition of continuity (preferred in particular by topologists) is the property that the inverse of continuous functions maps open sets into open sets.

Exercise 6.8. *Recall the notion of open sets $A \subset \mathbb{R}$ (cf. Lemma 5.6)*
$$A \subset \mathbb{R} \text{ is open} \quad \Leftrightarrow \quad \forall x_0 \in A: \ \exists \epsilon > 0: \ (x_0 - \epsilon, x_0 + \epsilon) \subset A.$$
Show the following. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Then the following are equivalent

(1) $f : \mathbb{R} \to \mathbb{R}$ is continuous.
(2) the inverse f^{-1} maps open sets into open sets. That is: whenever $A \subset \mathbb{R}$ is an open set, then the $f^{-1}(A)$ defined as
$$f^{-1}(A) \equiv \{x \in \mathbb{R} : f(x) \in A\}$$
is an open set.

Exercise 6.9. *Let D be open*[17]*, $f : D \to \mathbb{R}$ be a map let $(x_n)_{n \in \mathbb{N}} \subset D$ be a sequence converging to $x_0 \in D$. Show the following two statements*

- *If $\lim_{n \to \infty} f(x_n) \neq f(x_0)$ then f is discontinuous at x_0.*

[17] so in view of Lemma 5.11 no worries about clusterpoints!

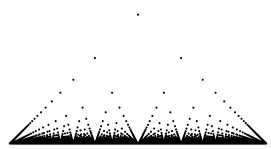

FIGURE 6.3. Graph of the "Popcorn function" . (wikipedia, public domain), cf. Example 6.10

- *Assume that* $\lim_{n\to\infty} f(x_n)$ *does not exist. Show there is no continuous replacement at* x_0*. That is, there is no continuous* $g : D \to \mathbb{R}$ *with* $g(x) = f(x)$ *for all* $x \in D \backslash \{x_0\}$*.*

Hint: Proposition 6.5.

Example 6.10.

- Let $D : \mathbb{R} \to \mathbb{R}$ be the *Dirichlet function*

$$D(x) := \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

  Then D is everywhere discontinuous.

  Indeed, for any $x_0 \in \mathbb{R}$ we can easily construct a sequence $x_n \in \mathbb{R}$ with $\lim_{n\to\infty} x_n = x_0$ such that

$$x_n \in \mathbb{Q}, \text{if and only if } n \text{ even.}$$

  Then

$$D(x_n) = \begin{cases} 1 & n \text{ even} \\ 0 & n \text{ odd} \end{cases}$$

  so $\lim_{n\to\infty} D(x_n)$ does not exist.
- $f(x) := xD(x)$ (where $D(x)$ is the Dirichlet function from above). Then f is continuous at $x = 0$ (squeeze lemma: $|f(x)| \leq |x| \xrightarrow{x\to 0} 0$), and discontinuous for $x \neq 0$ ($D(x) = \frac{1}{x} xD(x)$. If $xD(x)$ was continuous in $x_0 \neq 0$ then so was $D(x)$ by the limit laws).
- *Thomae's function* (also "Popcorn function"). $f : (0,1) \to \mathbb{R}$ defined as

$$f(x) := \begin{cases} \frac{1}{q} & \text{if } x = \frac{p}{q},\ p, q \in \mathbb{N} \text{ with no common divisors} \\ 0 & \text{if } x \text{ is irrational.} \end{cases}$$

  Then f is discontinuous at all *rational* numbers $x_0 \in (0,1) \cap \mathbb{Q}$[18] and continuous at all *irrational* $x_0 \in (0,1) \cap \backslash\mathbb{Q}$ (this is a bit more work). (for a precise proof see [Leb, Example 3.2.12], picture see Figure 6.3).
- So it is possible for the irrationals to be the set of continuity points of a function. However, a fun fact is that it is impossible to construct a function that is continuous only on the rational numbers. See G_δ sets on wikipedia

[18] this is easy to see, for any rational x we have $f(x) \neq 0$, but there exists an irrational sequence $\mathbb{R}\backslash\mathbb{Q} \ni x_n \xrightarrow{n\to\infty} x$, so $f(x_n) = 0$ and so $\lim_{n\to\infty} f(x_n) = 0 \neq f(x)$

Exercise 6.11. [Leb, Ex 3.2.4] *Let* $f : \mathbb{R} \to \mathbb{R}$ *be defined by*

$$f(x) = \begin{cases} \sin(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Is f *continuous (and if not: where is it continuous, and where not)? Prove your assertion. Cf. Figure 5.3.*

Exercise 6.12. [Leb, Ex. 3.2.5] *Let* $f : \mathbb{R} \to \mathbb{R}$ *be defined by*

$$f(x) = \begin{cases} x\sin(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Is f *continuous? Prove your assertion.*

Cf. Figure 5.4.

Exercise 6.13. [Leb, Ex. 3.2.9] *Give an example of functions* $f : \mathbb{R} \to \mathbb{R}$ *and* $g : \mathbb{R} \to \mathbb{R}$ *such that the function* h *defined by* $h(x) := f(x) + g(x)$ *is continuous, but* f *and* g *are not continuous. Can you find* f *and* g *that are nowhere continuous, but* h *is a continuous function?*

Exercise 6.14. [Leb, 3.2.11] *Let* $f : \mathbb{R} \to \mathbb{R}$ *be continuous. Suppose that* $f(c) > 0$. *Show that there exists an* $\alpha > 0$ *such that for all* $x \in (c - \alpha, c + \alpha)$ *we have* $f(x) > 0$.

Exercise 6.15. *Show that any continuous map* $f : \mathbb{R} \to \mathbb{Z}$ *is constant. Do not use the Intermediate Value theorem, Theorem 8.2, but the definition(s) of continuous functions from above.*

Exercise 6.16. [Leb, ex. 3.2.10] *Let* $f : \mathbb{R} \to \mathbb{R}$ *and* $g : \mathbb{R} \to \mathbb{R}$ *be continuous functions. Suppose that for all rational numbers* $r \in \mathbb{Q}$, $f(r) = g(r)$. *Show that* $f(x) = g(x)$ *for all* $x \in \mathbb{R}$.

Exercise 6.17. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a function. Assume that for some* $\alpha \in (0, 1]$ *and some* $\Lambda > 0$ *the function* f *satisfies*

$$|f(x) - f(y)| \leq \Lambda |x - y|^\alpha \quad \forall x, y \in \mathbb{R}$$

Show that f *is continuous.*

Remark 6.18.
(1) If $|f(x) - f(y)| \leq \Lambda |x - y| \quad \forall x, y \in \mathbb{R}$ for some $\Lambda > 0$ we say that f is (uniformly) *Lipschitz continuous*, and Λ is called the *Lipschitz constant* of f.
(2) If f is Lipschitz constant with $\Lambda \leq 1$ then f is called a *contraction* (and if $\Lambda < 1$ then f is a *strict contraction*).
(3) If $|f(x) - f(y)| \leq \Lambda |x - y|^\alpha \quad \forall x, y \in \mathbb{R}$ for some $\Lambda > 0$ and $\alpha > 0$ we say that f is (uniformly) *Hölder continuous*, and Λ is called the *Hölder constant* of f.
(4) Once we have derivatives it is easy to show that if

$$|f(x) - f(y)| \leq \Lambda |x - y|^\alpha \quad \forall x, y \in \mathbb{R}$$

holds for some $\alpha > 1$ and $\Lambda > 0$ then f is constant. See Exercise 12.5.

(5) more generally (and not relevant in this course): a *modulus of continuity* is a function $\omega : [0,\infty] \to [0,\infty]$ which is increasing which continuously vanishes at 0, i.e.

$$\lim_{t\to 0} \omega(t) = \omega(0) = 0$$

A function $f : D \to \mathbb{R}$ *has the modulus of continuity* ω at a point x if

$$|f(x) - f(y)| \leq \omega(|x-y|) \quad \forall y \in D.$$

So Hölder continuous functions have the modulus of continutiy $\omega(t) = \Lambda t^\alpha$ for some $\Lambda > 0$.

Here is a cool result about *contractive* maps and Cauchy sequences (this is a special case of the *Banach Fixed Point theorem*

Theorem 6.19. *Assume $f : \mathbb{R} \to \mathbb{R}$ is a strict contraction, i.e. f is continuous and moreover there exists some $\lambda \in (0,1)$ such that*

$$|f(x) - f(y)| \leq \lambda |x-y| \quad \forall x, y \in \mathbb{R}$$

Then f has a fixed point, namely there exists exactly one $x \in \mathbb{R}$ with

$$f(x) = x.$$

We indeed need $\lambda < 1$ in Theorem 6.19. Take $f(x) = x+1$ then $|f(x) - f(y)| \leq 1|x-y|$, but f has no fixed point.

Proof.
- *Uniqueness*: assume we have

$$f(x) = x, \quad \text{and} \quad f(y) = y.$$

Then

$$|x-y| = |f(x) - f(y)| \leq \lambda|x-y|.$$

Since $\lambda < 1$ we find that $|x-y| = 0$.
- *Existence*: The idea is to use an iteration argument to produce a Cauchy sequence x_k (which thus converges).

Let $x_0 \in \mathbb{R}$ be any arbitary point. Set

$$x_n := f(x_{n-1}).$$

We then have

$$|x_{n+1} - x_{n+2}| = |f(x_n) - f(x_{n+1})| \leq \lambda |x_n - x_{n+1}|.$$

Since $\lambda < 1$ from Exercise 4.10 we know that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. By Theorem 4.4 we conclude (since we work in $\mathbb{R}$, which is a complete space) that $(x_n)_{n\in N}$ is actually converging.

Set $x := \lim_{n\to\infty} x_n$.

Since f is continuous we have $\lim_{n\to\infty} f(x_n) = f(x)$. Thus

$$x \xleftarrow{n\to\infty} x_{n+1} = f(x_n) \xrightarrow{n\to\infty} f(x)$$

Thus $x = f(x)$, and we have found our fixed point.

□

It is fun to observe that we use very few things in the above proof, namely we only need that Cauchy sequences converge.

For example the following is quite immediate

Exercise (Banach Fixed Point theorem in $\mathbb{R}^n$)**.** *This is an optional exercise*

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be strict contraction: For some $\lambda \in (0, 1)$ we have

$$|f(x) - f(y)|_{\mathbb{R}^n} \leq \lambda |x - y|_{\mathbb{R}^n}.$$

Here we use the usual norm for $\mathbb{R}^n$-vectors:

$$|(p_1, \ldots, p_n)|_{\mathbb{R}^n} := \left(\sum_{i=1}^{n} |p_i|^2 \right)^{\frac{1}{2}}.$$

Then there exists exactly one $x \in \mathbb{R}^n$ with $T(x) = x$.

Hint: Either you define the notion of Cauchy sequences in $\mathbb{R}^n$ or you just argue componentwise...

One important application of the $\mathbb{R}^n$ (this is a bit easier in $\mathbb{R}^1$, so we do it in $\mathbb{R}^n$) is the following (which is a simple version of the *inverse function theorem*, which says that any C^1-function $f : \mathbb{R}^n \to \mathbb{R}^n$ is locally invertable around a point $x_0 \in \mathbb{R}^n$ if the matrix $Df(x_0)$ is invertible)

Corollary 6.20 (Small distorions of invertible maps are invertible)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be Lipschitz, i.e. assume there exist $\Lambda > 0$ such that*

$$|f(x) - f(y)| \leq \Lambda |x - y| \quad \forall x \in \mathbb{R}^n.$$

Let $A \in \mathbb{R}^{n \times n}$ be invertible matrix.

We consider the map

$$T_\varepsilon := A \cdot + \varepsilon f(\cdot) : \mathbb{R}^n \to \mathbb{R}^n$$

defined as

$$T_\varepsilon(x) := Ax + \varepsilon f(x).$$

There exists an $\varepsilon_0 = \varepsilon_0(\Lambda, A) > 0$ such that $T_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ is bijective from $\mathbb{R}^n \to \mathbb{R}^n$ whenever $\varepsilon \in \mathbb{R}$, $|\varepsilon| < \varepsilon_0$.

Proof. This is relatively easy linear algebra if $f = Bx$ for a matrix $B \in \mathbb{R}^{n \times n}$, i.e. if f is *linear*. Since then $\det(A - \varepsilon B)$ is nonzero if and only if $|\varepsilon| \ll 1$.

But f might be very nonlinear, indeed it is not even differentiable!

We want to show that that T_ε is bijective (for suitably small ε). That is we want to show

$$\forall p \in \mathbb{R}^n : \quad \exists! x \in \mathbb{R}^n : T_\varepsilon(x) = p. \tag{6.2}$$

So fix some $p \in \mathbb{R}^n$. By the definition of T_ε we want to find x such that

$$Ax + \varepsilon f(x) = p.$$

Denoting $A^{-1} \in \mathbb{R}^{n\times n}$ the inverse matrix of A (exists by assumption), the above is equivalent to

$$x + \varepsilon A^{-1} f(x) = A^{-1} p.$$

or, more usefully, the *fixed point* equation

$$x = A^{-1}p - \varepsilon A^{-1} f(x).$$

So if we call

$$S_\varepsilon(x) := A^{-1}p - \varepsilon A^{-1} f(x)$$

then $S_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$. Moreover we have

$$S_\varepsilon(x) - S_\varepsilon(y) = \varepsilon A^{-1} \left(f(y) - f(x)\right)$$

Using the matrix operator norm we then have

$$|S_\varepsilon(x) - S_\varepsilon(y)| \leq \varepsilon |A^{-1}||f(y) - f(x)| \leq \varepsilon \Lambda |A^{-1}||x - y|$$

So if we set $\lambda := \varepsilon\Lambda|A^{-1}|$ we see that for $\varepsilon_0 := \frac{1}{|A^{-1}|\Lambda|}$ we have that $S_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ is a strict contraction!

The above Banach Fixed Point theorem tells us that indeed, there exists exactly one $x \in \mathbb{R}^n$ such that $S_\varepsilon(x) = x$. That is we have established (6.2). □

Ineed it is easy to extend the Banach Fixed Point theorem to any metric space which is *complete*:

Exercise (Banach Fixed Point theorem). *This is an optional exercise*

Let (X, d) be a metric space, cf. Definition 1.7. Assume the metric space is complete, i.e. assume that any Cauchy sequence converges.

Let $T : X \to X$ be a continuous map (in the sense that whenever $(x_n)_{n\in\mathbb{N}} \subset X$ converges to some x, that is if $\lim_{n\to\infty} d(x_n, x) = 0$ then $\lim_{n\to\infty} f(x_n) = x$). Assume T additionally is a (strict) contraction:

$$d(T(x), T(y)) \leq \lambda d(x, y) \quad \forall x, y \in X.$$

Then there exists exactly one $x \in X$ with $T(x) = x$.

Let us two standard applications of Banach Fixed Point theorem (which are typically treated in Advanced Calculus II):

- Picard-Lindelöf theorem (also Cauchy-Lipschitz theorem): exisence and uniqueness for (ordinary) differential equations, Section 20.

- Implicit and inverse function theorem (IFT): a function $f : \mathbb{R}^n \to \mathbb{R}^m$, when it is (in what sense) invertible?

For more, see https://en.wikipedia.org/wiki/Banach_fixed-point_theorem

Further (optional) exercises. Computing limits of functions is also very important, so here some practice exams

Exercise. *Determine if the function is continuous. Prove your claim!*

(1)

$$f(x) = \begin{cases} x+1 & x \leq 1 \\ \frac{1}{x} & 1 < x < 3 \\ \sqrt{x-4} & x \geq 3 \end{cases}$$

(2)

$$g(x) = \begin{cases} x+2 & x \leq 1 \\ e^x & 0 \leq x \leq 3 \\ 2-x & x > 1 \end{cases}$$

(you can assume that e^x is continuous in $\mathbb{R}$, we will later obtain this from the Weierstrass M-test, Example 17.3.)

(3)

$$h(x) = \begin{cases} \sqrt{x} & x \leq 1 \\ 1 & x > 1 \end{cases}$$

Exercise. *(1) Let $f : \mathbb{R} \to \mathbb{R}$ be a function with $f(0) = 0$ and f is continuous around 0. Assume moreover that $f(x + y) = f(x) + f(y)$ for all $x, y \in \mathbb{R}$. Show that f is continuous in all of $\mathbb{R}$.*

(2) Let $f, g : [a, b] \to \mathbb{R}$ be continuous functions and set

$$m(x) := \max\{f(x), g(x)\} \quad x \in [a, b].$$

Show that $m(x)$ is continuous.

7. Min-Max Theorem

From Calculus we remember the beautiful result that *continuous function on a bounded closed interval attain their maximum and minimum*. Now we are going to prove it.

Definition 7.1. Assume $f : D \to \mathbb{R}$ is a function.

- $\sup_D f := \sup f(D)$. If $\sup_D f < \infty$ then f is *bounded from above*.
- $\inf_D f := \inf f(D)$.If $\inf_D f > -\infty$ then f is *bounded from below*.

- f is called a *bounded function* if $\sup_D f < \infty$ and $\inf_D f > 0$. Equivalently, a function is bounded if and only if there exists $M > 0$ such that
$$|f(x)| \leq M \quad \forall x \in D$$
- If there exists $x \in D$ such that $f(x) = \sup_D f$ then we say that x is a *maximum point* of f on D, and $f(x)$ is the maximum value. We might also say "the maximum of f is attained/achieved on D (in x)". We then write $f(x) = \max_D f$, and $x = \text{argmax}_D f$.
- If there exists $x \in D$ such that $f(x) = \inf_D f$ then we say that x is a *minimum point* of f on D, and $f(x)$ is the minimum value. We might also say "the minimum of f is attained/achieved on D (in x)". We then write $f(x) = \min_D f$, and $x = \text{argmin}_D f$.

Theorem 7.2. *Let $f : [a,b] \to \mathbb{R}$ be a continuous function $-\infty < a < b < \infty$. Then f achieves its maximum and minimum on $[a,b]$, that is there exists $x_{max}, x_{min} \in [a,b]$ with*
$$f(x_{max}) \geq f(y) \geq f(x_{min}) \quad \forall y \in [a,b].$$

Example 7.3.
- If f is discontinuous on $[a,b]$ the statement in Theorem 7.2 can be false. Take, e.g.
$$f(x) = \begin{cases} \frac{1}{x} & x \in [-1,1]\backslash\{0\} \\ 0 & x = 0 \end{cases}$$
- If f is continuous on (a,b) the statement of Theorem 7.2 can be false: Again take f from above, but on $D = (0,1]$. It's continuous, but the "maximum" is $+\infty$.

Proof of Theorem 7.2. Let $S := \sup_{[a,b]} f = \sup f([a,b])$. Observe that $S \in (-\infty, \infty]$, i.e. as of now we cannot rule out $S = \infty$.

We consider two cases:

If $S < \infty$: By the definition of the supremum, Lemma 1.5, there must be a sequence $(x_n)_{n\in\mathbb{N}} \subset [a,b]$ with
$$S - \frac{1}{n} \leq f(x_n) \leq S,$$
i.e.
$$\lim_{n\to\infty} f(x_n) = S.$$
If $S = \infty$: Again by Lemma 1.5 there must be a sequence $(x_n)_{n\in\mathbb{N}} \subset [a,b]$ with
$$n \leq f(x_n),$$
i.e.
$$\lim_{n\to\infty} f(x_n) = \infty.$$
Observe that $[a,b]$ is a bounded interval, so the sequence $(x_n)_{n\in\mathbb{N}}$ is bounded. By Bolzano-Weierstrass, Theorem 3.8, there exists a convergent subsequence $(x_{n_i})_{i\in\mathbb{N}} \subset [a,b]$ and $x \in \mathbb{R}$ with $\lim_{i\to\infty} x_{n_i} = x$. Since $[a,b]$ is a closed set, $x \in [a,b]$.

In both cases: by the sequential characterization of continuity, Proposition 6.5, we have

$$f(x) = \lim_{i\to\infty} f(x_{n_i}) = S = \sup_{[a,b]} f \geq f(y) \quad \forall y \in [a,b].$$

In particular we infer that $S < \infty$ (because $f(x)$ is a real number, and thus $f(x) < \infty$). That is if we set $x_{max} := x$ then we have found our maximum.

We argue similarly for the existence of x_{min}, setting $I := \inf_{[a,b]} g = \inf f([a,b])$, observing that $I \in [-\infty, \infty)$. □

As a consequence of Theorem 7.2 we get the following corollary.

Corollary 7.4. *Let $f : [a,b] \to \mathbb{R}$ be a continuous function $-\infty < a < b < \infty$. Then f is bounded.*

Exercise 7.5. *Prove Corollary 7.4 and show its sharp. More precisley show the following:*

(1) Let f be continuous on $[a,b]$. Then f is bounded.
(2) Give an example of a continuous function $f : (a,b] \to \mathbb{R}$ such that f is not bounded.

Example 7.6.
- If f is discontinuous on $[a,b]$ the statement of Corollary 7.4 is false. Take, e.g.

$$f(x) = \begin{cases} \frac{1}{x} & x \in [-1,1]\backslash\{0\} \\ 0 & x = 0 \end{cases}$$

- Just because a function $f : \mathbb{R} \to \mathbb{R}$ attains its maximum and infimum on any closed interval $[a,b]$, does not mean it is continuous: $f(x) = \sin(1/x)$.
- The function $f(x) = x$ is continuous, but on *unbounded* intervals $\mathbb{R}$ or $[0,\infty)$ it does not attain necessarily maximum and minimum.

If one inspects the proof of Theorem 7.2 then we see that we did not use full continuity for the existence of the maximum or minimum. This is important in applications in Analysis where functions (or functionals) are indeed only *lower semicontinuous*. This happens a lot in Partial Differential Equations (cf. Viscosity solutions, Direct Method of Calculus etc.)

Definition 7.7 (liminf and limsup)**.** Let $D \subset \mathbb{R}$ and $c \in \overline{D}$ a *cluster point*. Assume that $f : D \to \mathbb{R}$ is a function.

We then define

- the lim sup as

$$\limsup_{D \ni y \to c} f(y) = \lim_{\delta \to 0} \sup_{D\cap(c-\delta,c+\delta)\backslash\{c\}} f \equiv \inf_{\delta>0} \sup_{D\cap(c-\delta,c+\delta)\backslash\{c\}} f \in \mathbb{R} \cup \{+\infty\}.$$

Here we observe that $g(\delta) := \sup_{D\cap(c-\delta,c+\delta)\backslash\{c\}} f$ is a monotone function (taking values possibly $+\infty$, since c is a clusterpoint $g(\delta) > -\infty$): For $\delta < \tilde{\delta}$ we have $g(\delta) < g(\tilde{\delta})$. Thus the limit exists and equals the infimum.

- the sequential definition of lim sup is a bit more complicated than for the limit, and we will not prove it. However we note (without proof):

$$\limsup_{D \ni y \to c} f(y) = L \in \mathbb{R} \cup \{+\infty\}$$

is equivalent to the fact that

(1) for all $(y_n)_{n\in\mathbb{N}} \subset D \setminus \{c\}$ with $\lim_{n\to\infty} y_n = c$ we have

$$\limsup_{n\to\infty} f(y_n) \leq L$$

and

(2) there exists at least one sequence $(y_n)_{n\in\mathbb{N}} \subset D \setminus \{c\}$ such that

$$\lim_{n\to\infty} f(y_n) = L.$$

(This sequence is called the *recovery sequence*.)

Similarly we can define the lim inf

Exercise 7.8. *A function $f : D \subset \mathbb{R} \to \mathbb{R}$ is called (sequentially) lower semicontinuous at a cluster point $x \in D$ if we have*

$$f(x) \leq \liminf_{D \ni y \to x} f(y),$$

in the following sense: for any sequence $(y_n)_{n\in\mathbb{N}} \subset D$ with $\lim_{n\to\infty} y_n = x$ we have

$$f(x) \leq \liminf_{n\to\infty} f(y_n).$$

In a similar spirit, a function is called (sequentially) upper semicontinuous if

$$f(x) \geq \limsup_{D \ni y \to x} f(y).$$

Cf. Figure 7.1.

(1) *Give an example of lower semicontinuous functions which is not continuous*
(2) *Give an example of upper semicontinuous functions which are not continuous*
(3) *Show that f is continuous at $x \in D$ if and only if f is lower and upper semicontinuous at x.*
(4) *Show that if $f : [a, b] \to \mathbb{R}$ is lower semicontinuos in every $x \in [a, b]$, then f attains its minimum value in $[a, b]$*
(5) *Show that if $f : [a, b] \to \mathbb{R}$ is upper semicontinuos in every $x \in [a, b]$, then f attains its maximum value in $[a, b]$.*

Definition 7.9 (Limits at $\pm\infty$)**.** For a function $f : \mathbb{R} \to \mathbb{R}$ we say

- For some $L \in \mathbb{R}$ we say $\lim_{x\to\infty} f(x) = L$ if

$$\forall \varepsilon > 0 \quad \exists N > 0 \quad |f(x) - L| \leq \varepsilon \quad \forall x > N.$$

- For some $L \in \mathbb{R}$ we say $\lim_{x\to-\infty} f(x) = L$ if

$$\forall \varepsilon > 0 \quad \exists N > 0 \quad |f(x) - L| \leq \varepsilon \quad \forall x < -N.$$

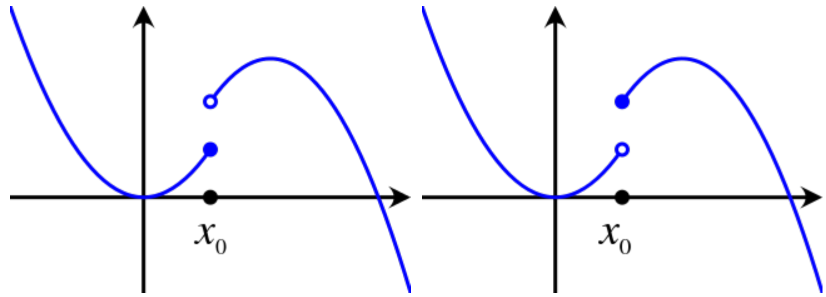

FIGURE 7.1. the graph of a *lower* semicontinuous function (left) and *upper* semicontinuous function (right). Image: Wikipedia:Mktyscn (public domain)

- We say $\lim_{x\to\infty} f(x) = +\infty$ if
$$\forall M > 0 \quad \exists N > 0 \quad f(x) > M \quad \forall x > N.$$
- We say $\lim_{x\to\infty} f(x) = -\infty$ if
$$\forall M > 0 \quad \exists N > 0 \quad f(x) < -M \quad \forall x > N.$$
- We can also define the $\limsup_{x\to\infty}$:
$$\limsup_{x\to\infty} f(x) := \lim_{c\to\infty} \sup_{x>c} f(x) \equiv \inf_c \sup_{x>c} f(x) \in \mathbb{R} \cup \{+\infty\}$$
and
$$\limsup_{x\to-\infty} f(x) := \lim_{c\to-\infty} \sup_{x<c} f(x) \equiv \inf_c \sup_{x<c} f(x) \in \mathbb{R} \cup \{+\infty\}$$
(observe that the limit above exists for the same it exists for the usual lim sup and lim inf – by monotonicity!)
- Similarly we define $\liminf_{x\to\infty}$
$$\liminf_{x\to\infty} f(x) := \lim_{c\to\infty} \inf_{x>c} f(x) \equiv \sup_c \inf_{x>c} f(x) \in \mathbb{R} \cup \{+\infty\}$$
and
$$\liminf_{x\to-\infty} f(x) := \lim_{c\to-\infty} \inf_{x<c} f(x) \equiv \sup_c \inf_{x<c} f(x) \in \mathbb{R} \cup \{+\infty\}$$
(observe that the limit above exists for the same it exists for the usual lim sup and lim inf – by monotonicity!)

Exercise 7.10. *Assume that $f : \mathbb{R} \to \mathbb{R}$ is continuous and*
$$\limsup_{x\to-\infty} f(x) \leq 0, \quad \limsup_{x\to\infty} f(x) \leq 0$$
and
$$f(0) = 0.$$
Show

(1) f attains its maximum in $\mathbb{R}$

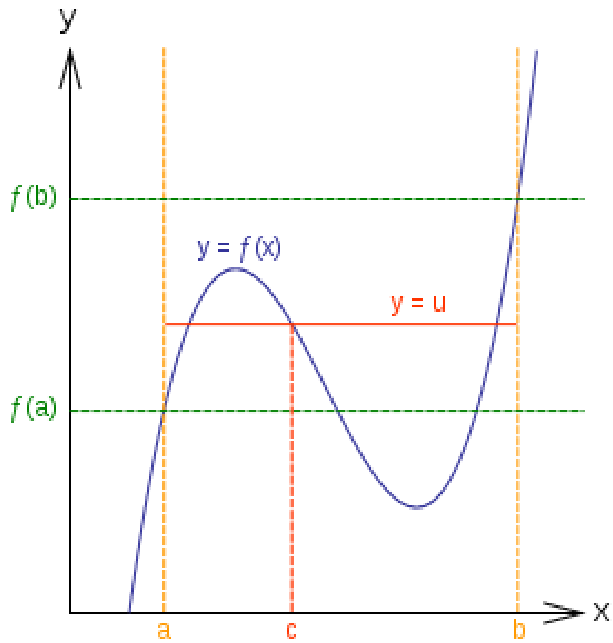

FIGURE 8.1. If we draw without lifting the pen any line connecting $(a, f(a))$ and $(b, f(b))$ that represents the graph of a function, then the y-values of this line pass through any value between $f(a)$ and $f(b)$. So for any $y \in (f(a), f(b))$ there exists $c \in (a, b)$ with $f(c) = y$. image source: wikipedia.

(2) *Give an example of a function with the above properties where f does not attains it minimum.*

8. Intermediate Value Theorem

Again, this is a statement we know (and love?) from Calculus 1. If a continuous function satisfies $f(a) = A$ and $f(b) = B$ then f has to attain any value between A and B in the interval (a, b) – cf. Figure 8.1.

We proof first a slightly simplified version of this statement (from which it will be easy to deduce the full statement).

Lemma 8.1. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function.*

Suppose that $f(a) < 0$ and $f(b) > 0$. Then there exists a number $c \in (a, b)$ with $f(c) = 0$.

Proof. The idea is to play catch with the the point c via a sequence [19], cf. Figure 8.2.

[19] The following might remind you of a binary search algorithm (since it essentially is that. The main difference is that we have a non-discrete set where we are looking, hence we need to talk about convergence to eventually find the point x

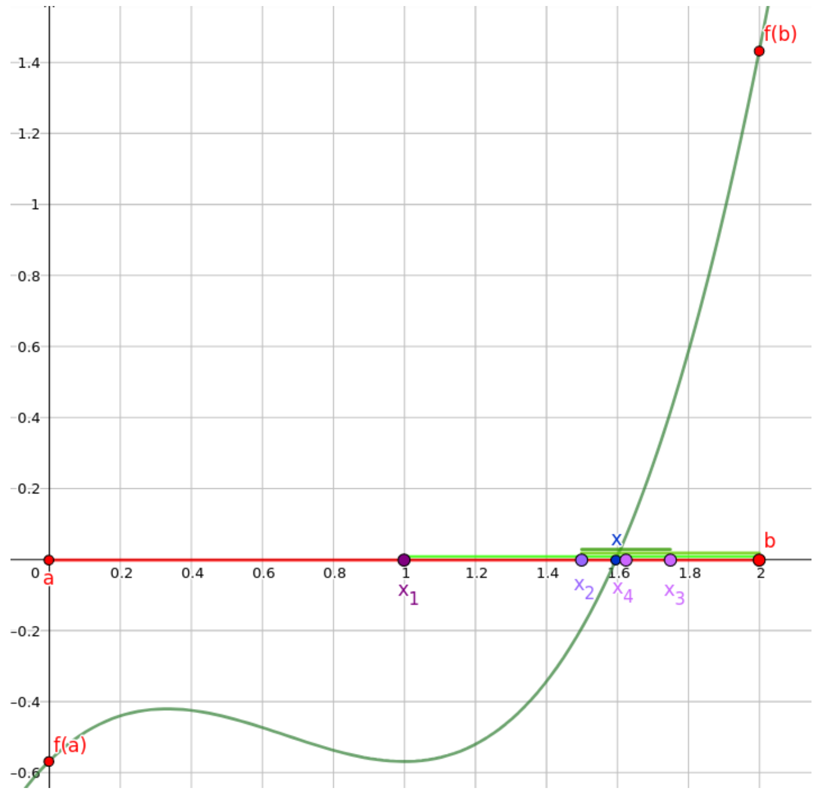

FIGURE 8.2. we find the point x as the limit of a sequence x_n which is constructed out of the midpoints of shrinking intervals (a_i, b_i) chosen such that the $f(a_i) < 0$ and $f(b_i) > 0$

We have $f(a) < f(b)$.

So let us assume $f(a) < f(b)$. We define the sequences $(a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty} \in [a, b]$ with the following properties:

- $a_n, b_n \in [a, b]$
- $a_n \leq a_{n+1}$ and $b_{n+1} \leq b_n$ (for $n \geq 1$)
- $f(a_n) \leq 0 \leq f(b_n)$
- $|a_n - b_n| \leq 2^{-n}|a - b|$.

We do so by induction:

- Set $a_0 = a$ and $b_0 = b$. Then $f(a_0) \leq f(b_0)$ and $|a_0 - b_0| = 2^{-0}|a - b|$.
- Assume as induction hypothesis that we have constructed $a_n, b_n \in [a, b]$ with the desired properties. We now construct a_{n+1} and b_{n+1}:
  Compute $f(\frac{a_n+b_n}{2})$. There are two possibilities:
  - If $f(\frac{a_n+b_n}{2}) \leq 0$ then we set $a_{n+1} := \frac{a_n+b_n}{2}$, $b_{n+1} = b_n$.
  - Otherwise $f(\frac{a_n+b_n}{2}) > 0$. In this case we set $a_{n+1} := a_n$ and $b_{n+1} = \frac{a_n+b_n}{2}$.

In both cases we check that

- $a_{n+1}, b_{n+1} \in [a, b]$ (since we take midpoints)
- $f(a_{n+1}) \leq 0 \leq f(b_{n+1})$ (by construction)
- $|a_{n+1} - b_{n+1}| \leq 2^{-n-1}|a - b|$ since $|a_n - \frac{a_n+b_n}{2}| = |\frac{a_n+b_n}{2} - b_n| = \frac{1}{2}|a_n - b_n| \leq 2^{-n}\frac{1}{2}|a - b|$ by induction hypothesis.

So, we have found the sequence $a_n, b_n \in [a, b]$ with the claimed properties.

Observe that $(a_n)_{n=1}^{\infty}$ is bounded and monotone, and thus convergent. The same holds for $(b_n)_{n=1}^{\infty}$.

Set

$$\bar{a} := \lim_{n\to\infty} a_n$$

$$\bar{b} := \lim_{n\to\infty} b_n$$

- then $\bar{a}, \bar{b} \in [a, b]$
- $f(\bar{a}) \leq 0 \leq f(\bar{b})$ – indeed by continuity $0 \geq \lim_{n\to\infty} f(a_n) = f(\bar{a})$ and similarly for $\bar{b}$.
- Moreover
$$|\bar{a} - \bar{b}| = \lim_{n\to\infty} |a_n - b_n| \leq \lim_{n\to\infty} 2^{-n}|\bar{a} - \bar{b}| = 0.$$

Set $c : \bar{a} = \bar{b}$. Then

$$0 \leq f(c) = f(\bar{b}) = f(\bar{a}) \leq 0$$

That is $f(c) = 0$, and we can conclude. □

As a corollary, we obtain the intermediate value theorem, originally due to Bolzano.

Theorem 8.2 (Intermediate Value Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be continuous. For any*

$$\min\{f(a), f(b)\} \leq M \leq \max\{f(a), f(b)\}$$

there exists $x \in [a, b]$ with $f(x) = M$.

If

$$\min\{f(a), f(b)\} < M < \max\{f(a), f(b)\}$$

(in particular $f(a) \neq f(b)$) then we can find such an x in (a, b)

Proof. If $f(a) = f(b)$ there is nothing to show, simply take $x = a$. More generally, if $M = f(a)$ or $M = f(b)$ there is noting to show.

So we may assume w.l.o.g.

$$f(a) < M < f(b)$$

Set $g(x) := f(x) - M$.

Then $g(a) < 0$ and $g(b) > 0$, so we can apply Lemma 8.1 and find some $x \in [a, b]$ with $g(x) = 0$, that is $f(x) = M$. □

From Calculus 1 we recall exercises like the following:

Exercise 8.3. *Show that there exists at least one solution to* $x - \cos(x) = 0$.

Exercise 8.4. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a continuous function such that* $\lim_{x\to\infty} f(x) = M_+$ *and* $\lim_{x\to-\infty} = M_-$ *for some* $M-, M_+ \in \mathbb{R} \cup \{\infty, -\infty\}$. *Then for any* $L \in (M_-, M_+)$ *there exists* $x \in \mathbb{R}$ *with* $f(x) = L$.

Exercise 8.5. *Prove Exercise 8.5 using the Intermediate Value theorem, Theorem 8.2: show that any continuous map* $f : \mathbb{R} \to \mathbb{Z}$ *is constant.*

Exercise 8.6. *Show that any continuous map* $f : \mathbb{R} \to \mathbb{Q}$ *is constant.*

Observe, there are *discontinuous* functions that still satisfy the conclusion of the Intermediate Value theorem, Theorem 8.2:

Exercise 8.7. [Leb, Ex. 3.3.4] *Let*

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

Show that f *has the intermediate value property. That is, for any* $a < b$, *if there exists a* y *such that* $f(a) < y < f(b)$ *or* $f(b) < y < f(a)$, *then there exists* $c \in (a, b)$ *such that* $f(c) = y$.

Exercise 8.8. [Leb, ex. 3.3.7] *Suppose that* $f : [a, b] \to \mathbb{R}$ *is a continuous function. Prove that the image* $f([a, b])$ *is a closed and bounded interval or a single number.*

Hint: Combine the min-max theorem, Theorem 7.2, with the intermediate value theorem, Theorem 8.2.

9. Uniform continuity

We learned the notion of continuity above, f is continuous at x if

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : |f(x) - f(y)| < \varepsilon \quad \forall |x - y| < \delta.$$

If we consider $f(x) := 1/x$ or $f(x) = \sin(1/x)$ on $(0, 1)$ then we learned that these are continuous function on $(0, 1)$ (as composition of continuous function). However, they are not *uniformly* continuous functions. Lets check that again for $1/x$. Let $\varepsilon \in (0, 1)$ and $x \in (0, 1)$ be given. We can choose e.g. $\delta := \frac{1}{2}\varepsilon x^2$ then

$$\left|\frac{1}{x} - \frac{1}{y}\right| = \frac{|x - y|}{xy} \leq \frac{\delta}{xy} \leq \varepsilon\frac{1}{2}\frac{x}{y} =\leq \varepsilon\frac{1}{2}\frac{x - y}{y} + \frac{1}{2}\varepsilon \leq \frac{1}{4}\varepsilon + \frac{1}{2}\varepsilon < \varepsilon.$$

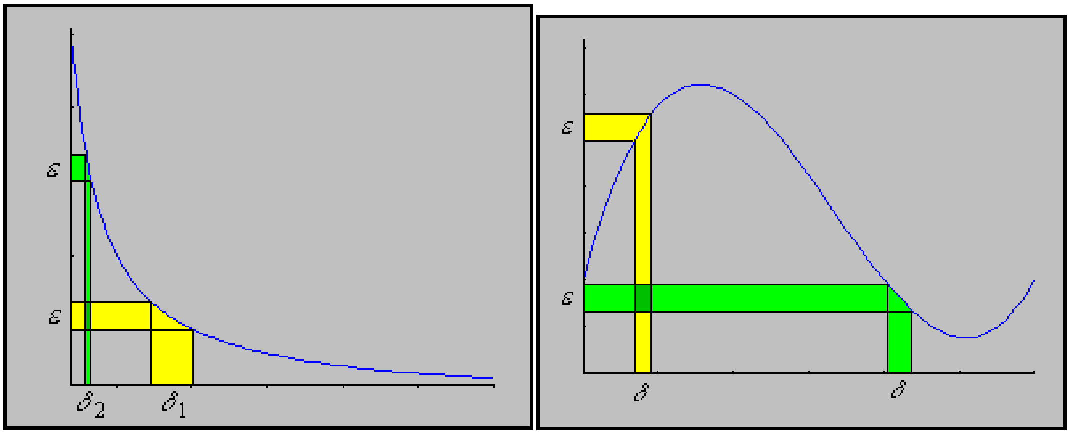

FIGURE 9.1. Uniform continuity: The function on the left is not uniformly continuous, since for some fixed $\varepsilon > 0$ δ has to be chosen differently as x changes. The function on the right is uniformly continuous, since for fixed $\varepsilon > 0$ we can choose a fixed δ independent of the point x. Picture source: Mathcs.org

So we choose δ in dependence on x, that means as x goes to zero, δ needs to be smaller and smaller. This is not only our incompetence, indeed, the function $1/x$ becomes crazy around $x \approx 0$, in that it blows up to infinity.

Uniform continuity is a property that rules out this sort of behaviour: it assumes that δ can be chosen independent of x, only dependent on ε.

Definition 9.1 (Uniform continuity)**.** A function $f : D \to \mathbb{R}$ is

(1) *(pointwise) continuous* in D if
$$\forall x \in D : \forall \varepsilon > 0 \quad \exists \delta > 0 : |f(x) - f(y)| < \varepsilon \quad \forall y \in D,\, |x - y| < \delta.$$
(2) *uniformly continuous* in D, if
$$\forall \varepsilon > 0 \quad \exists \delta > 0 : |f(x) - f(y)| < \varepsilon \quad \forall x, y \in D,\, |x - y| < \delta.$$

Cf. Figure 9.1

Example 9.2.
- As we have seen above, $\frac{1}{x}$, $\sin(1/x)$ are continuous, but *not* uniformly continuous in $(0, 1)$,
- Assume that $f : D \to \mathbb{R}$ is *Lipschitz continuous*, that is there exists $K > 0$ such that
$$|f(x) - f(y)| \leq K|x - y| \quad \forall x, y \in D.$$
then f is uniformly continuous. Example for Lipschitz continuous function is for example $f(x) = |x|$.

- More generally, if $f : D \to \mathbb{R}$ is *Hölder continuous*, that is there exists $K > 0$ and $\alpha > 0$ such that
$$|f(x) - f(y)| \leq K|x - y|^\alpha \quad \forall x, y \in D.$$
then f is uniformly continuous. For example $\sqrt{|x|}$ is $\frac{1}{2}$-Hölder continuous, but not Lipschitz continuous.

Exercise 9.3. *Show*

(1) If $f : D \to \mathbb{R}$ is uniformly continuous, then f is continuous.
(2) The converse is false (give a counterexample)

The important observation is that functions that are pointwise continuous in *a closed, bounded interval*, then they are uniformly continuous. Example 9.2 shows this is not true for open intervals

Theorem 9.4. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function.*[20] *Then f is uniformly continuous.*

Proof of Theorem 9.4. Assume to the contrary that $f : [a, b] \to \mathbb{R}$ is a continuous function but f is not uniformly continuous. Then
$$\exists \varepsilon > 0 \quad \forall \delta > 0 : \exists x, y \in [a, b], \ |x - y| < \delta : \quad |f(x) - f(y)| \geq \varepsilon$$
Fix this $\varepsilon > 0$. Taking $\delta = \frac{1}{n}$ we find sequences $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}} \subset [a, b]$ such that
$$|x_n - y_n| < \frac{1}{n} \tag{9.1}$$
and
$$|f(x_n) - f(y_n)| \geq \varepsilon. \tag{9.2}$$
Since $[a, b]$ is bounded, $(x_n)_{n\in\mathbb{N}}$ is bounded. By Bolzano Weierstrass, Theorem 3.8 there must be a subsequence such that $(x_{n_i})_{i\in\mathbb{N}}$ is convergent to some $x \in [a, b]$ (because $[a, b]$ is closed), $\lim_{i\to\infty} x_{n_i} = x$.

Now we have
$$|y_{n_i} - x| \overset{(9.1)}{\leq} |x_{n_i} - x| + \frac{1}{n_i} \xrightarrow{i\to\infty} 0,$$
so we also have $\lim_{i\to\infty} y_{n_i} = x$.

Since f is continuous at x, using the sequential characterization of continuity from Proposition 6.5, there must be some $N_2 = N_2(\varepsilon) \in \mathbb{N}$ such that
$$|f(x) - f(y_{n_i})| < \frac{\varepsilon}{4}, \ |f(x) - f(x_{n_i})| < \frac{\varepsilon}{4} \quad \forall i \geq N_2.$$

[20] observe: closed interval, and in particular we assume $-\infty < a < b < \infty$

But then

$$|f(x_{n_i}) - f(y_{n_i})| \leq |f(x) - f(x_{n_i})| + |f(x) - f(y_{n_i})| < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2} < \varepsilon.$$

This contradicts (9.2), so the initial assumption that f is not uniformly continuous must have been false. □

Exercise 9.5. *Repeating the proof of Theorem 9.4, show that any continuous map $f : D \to \mathbb{R}$ is actually uniformly continuous, if D is a compact set.*

A set $A \subset \mathbb{R}$ is (sequentially) compact if any sequence $(x_n)_{n\in\mathbb{N}} \subset A$ has a converging subsequence $(x_{n_i})_{i\in\mathbb{N}}$ with $\lim_{i\to\infty} x_{n_i} = x \in A$.

Why do we care about uniform ellipticity? It rules out "degeneracy" at the boundary.

Theorem 9.6. *Let $f : (a, b) \to \mathbb{R}$ be uniformly continuous. Then there exists a continuous extension $\tilde{f}$ on $[a, b]$. That is: there is $\tilde{f} : [a, b] \to \mathbb{R}$ continuous (by Theorem 9.4: uniformly continuous) such that $\tilde{f}(x) = f(x)$ for all $x \in (a, b)$.*

Also, $\tilde{f}$ is unique. That is any $\tilde{g} : [a, b] \to \mathbb{R}$ which is continuous and satisfies $\tilde{g}(x) = f(x) = \tilde{f}(x)$ for all $x \in (a, b)$ satisfies $\tilde{f} = \tilde{g}$ in all of $[a, b]$.

Example 9.7. The statement of the theorem is false in general if f is not uniformly continuous, but merely continuous: Let $f(x) = \frac{1}{x}$ for $D = (0, 1)$. There is no continuous extension to $[0, 1]$ because then $\tilde{f}(0) = \infty$. Similar argument shows that $\sin(1/x)$ on $(0, 1)$ cannot continuously be extended to $[0, 1]$.

Exercise 9.8. *Show the last statement of Theorem 9.6:*

Assume $\tilde{f}, \tilde{g} : [a, b] \to \mathbb{R}$ are both continuous and satisfy $\tilde{g}(x) = \tilde{f}(x)$ for all $x \in (a, b)$. Then $\tilde{f}(a) = \tilde{g}(a)$ and $\tilde{f}(b) = \tilde{g}(b)$.

The main ingredient for the proof of Theorem 9.6 is the following Lemma

Lemma 9.9. *Let $f : D \to \mathbb{R}$ be uniformly continuous and let $(x_n)_{n\in\mathbb{N}} \subset D$ be a Cauchy sequence. Then $(f(x_n))_{n\in\mathbb{N}}$ is a Cauchy sequence.*

Proof. Let $\varepsilon > 0$ we need to show that there exists $N \in \mathbb{N}$ such that

$$|f(x_n) - f(x_m)| < \varepsilon \quad \forall n, m > N. \tag{9.3}$$

Since f is uniformly continuous, there must be some $\delta = \delta(\varepsilon) > 0$ such that

$$|f(x) - f(y)| < \varepsilon \quad \forall x, y \in D : \quad |x - y| < \delta. \tag{9.4}$$

On the other hand, since $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, there exists an $N = N(\delta)$ such that

$$|x_n - x_m| < \delta \quad \forall n, m > N. \tag{9.5}$$

Note that (9.5) together with (9.4) implies (9.3). □

Proof of Theorem 9.6. So let $f : (a, b) \to \mathbb{R}$ be uniformly continuous. Set for $x \in [a, b]$ set

$$\tilde{f}(x) := \lim_{(a,b)\ni y \to \bar{x}} f(y).$$

We first need to show that this makes sense.

Fix now $\bar{x} \in [a, b]$ is a cluster point, Lemma 5.6, so the above notion of limit is defined (but we have to ensure that the limit exists as a real number).

Let now $(x_n)_{n\in\mathbb{N}} \subset (a, b)$ be any sequence converging to a. Then $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, Theorem 4.4. By uniform continuity, $(f(x_n))_{n\in\mathbb{N}}$ is a Cauchy sequence, Lemma 9.9. Cauchy sequences converge in $\mathbb{R}$, Theorem 4.4, so there exists $\Gamma_{\bar{x}} \in \mathbb{R}$ with

$$\lim_{n\to\infty} f(x_n) = \Gamma_{\bar{x}} \in \mathbb{R}. \tag{9.6}$$

We claim[21] that

$$\lim_{(a,b)\ni y\to\bar{x}} f(y) = \Gamma_{\bar{x}}.$$

In order to prove this we have to show (by definition of the limit)

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad |f(y) - \Gamma_{\bar{x}}| < \varepsilon \quad \forall y \in (a, b),\ y \neq \bar{x},\ |y - \bar{x}| < \delta. \tag{9.7}$$

Let us gather what we know

- From uniform continuity we have
$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad |f(y) - f(x)| < \frac{\varepsilon}{2} \quad \forall x, y \in (a, b), |y - x| < 2\delta.$$
- Since $\lim_{n\to\infty} x_n = \bar{x}$ we have
$$\forall \delta > 0 \quad \exists N_1 > 0 : \quad |x_n - \bar{x}| < \delta \quad \forall n \geq N_1.$$
- From the definition of $\Gamma_{\bar{x}}$ in (9.6) we know
$$\forall \varepsilon > 0 \quad \exists N_2 = N_2(\varepsilon) : |f(x_n) - \Gamma_{\bar{x}}| < \frac{\varepsilon}{2} \quad \forall n \geq N_2$$

So for $\varepsilon > 0$ we take δ from the first bullet point, $N := \max\{N_1, N_2\}$ where N_1, N_2 are from second and third bullet point, respectively.

If $y \in (a, b)$ such that $|y - \bar{x}| < \delta$ then we have

$$|y - x_n| < 2\delta \quad \forall n \geq N$$

so from the first bullet point above we find

$$|f(y) - f(x_n)| < \frac{\varepsilon}{2} \quad \forall n \geq N.$$

Thus, with the third bullet point we find

$$|f(y) - \Gamma_{\bar{x}}| \leq |f(y) - f(x_n)| + |f(x_n) - \Gamma_{\bar{x}}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

[21] this is not obvious, from what we have so far

That is we have established (9.7). Consequently, we may define

$$\tilde{f}(\bar{x}) := \lim_{(a,b)\ni y\to\bar{x}} f(y) = \Gamma_{\bar{x}}.$$

For now $\tilde{f} : [a,b] \to \mathbb{R}$ is only a function. We need to establish it is continuous and coincides with f on (a,b).

The latter is easy. Since f is continuous,

$$f(\bar{x}) \overset{\text{continuity}}{=} \lim_{y\to\bar{x}} f(y) \overset{\text{def}}{=} \tilde{f}(\bar{x}) \quad \forall \bar{x} \in (a,b).$$

But then we also have

$$\tilde{f}(a) \overset{\text{def}}{=} \lim_{y\to a} f(y) \overset{y\neq a}{=} \lim_{y\to a} \tilde{f}(y),$$

that is $\tilde{f}$ is continuous at a. By the same argument $\tilde{f}$ is continuous at b. We can conclude. □

Exercise 9.10. *Let $f : \mathbb{Q} \to \mathbb{R}$ be a uniformly continuous function.*

Show that there exist a unique extension $\tilde{f} : \mathbb{R} \to \mathbb{R}$ which is continuous. Namely,

(1) for any $x \in \mathbb{R}$ the function $\tilde{f}(x) := \lim_{y\ni\mathbb{Q}\to x} f(y)$ is well defined
(2) show that $\tilde{f}$ is uniformly continuous and that $\tilde{f}(y) = f(y)$ for any $y \in Q$
(3) show that any other continuous function $g : \mathbb{R} \to \mathbb{R}$ with $g(y) = f(y)$ for all $y \in Q$ is equal to $\tilde{f}$: $\tilde{f}(x) = g(x)$ for all $x \in \mathbb{R}$.

Hint: for (1) Cauchy sequences

Exercise 9.11. [Leb, ex. 3.4.3] *Show that $f : (c,\infty) \to \mathbb{R}$ for some $c > 0$ and defined by $f(x) := 1/x$ is Lipschitz continuous (for the definition, see Example 9.2)*

Exercise 9.12. [Leb, ex. 3.4.4] *Show that $f : (0,\infty) \to \mathbb{R}$ defined by $f(x) := 1/x$ is not Lipschitz continuous (for the definition, see Example 9.2).*

Exercise 9.13. *A function $f : D \subset \mathbb{R} \to \mathbb{R}$ is called (sequentially) lower semicontinuous at a point $x \in D$ if we have*

$$f(x) \leq \liminf_{D\ni y\to x} f(y),$$

in the sense that for any sequence $(y_n)_{n\in\mathbb{N}} \subset D$ with $\lim_{n\to\infty} y_n = x$ we have

$$f(x) \leq \liminf_{n\to\infty} f(y_n).$$

In a similar spirit, a function is called (sequentially) upper semicontinuous if

$$f(x) \geq \limsup_{D\ni y\to x} f(y).$$

(a) Give an example of a lower semicontinuous function which is not continuous.

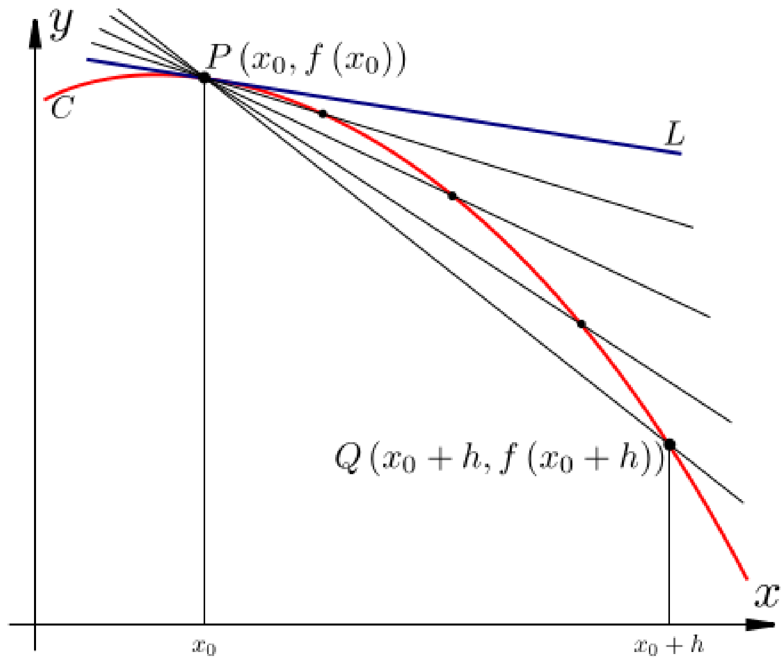

FIGURE 10.1. The derivative of the function f (in red) is the slope of the tangent line (blue), which is the limit of the slopes of the secant lines.

(b) *Give an example of an upper semicontinuous function which is not continuous.*
(c) *Show that f is continuous at $x \in D$ if and only if f is lower and upper semicontinuous at x.*
(d) *Show that if $f : [a, b] \to \mathbb{R}$ is lower semicontinuous in every $x \in [a, b]$, then f attains its minimum value in $[a, b]$.*
(e) *Show that if $f : [a, b] \to \mathbb{R}$ is upper semicontinuous in every $x \in [a, b]$, then f attains its maximum value in $[a, b]$.*

10. Derivatives

We remember from Calculus the definition of the derivative, as the limit of the slopes of secant lines (cf. Figure 10.1). Recall that the slope of the secant line of between $(x, f(x))$ and $(x + h, f(x + h))$ is

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

Definition 10.1. Let $f : D \to \mathbb{R}$ be a function and $x \in D$ a cluster point of D. We say that f is *differentiable at x*, if

$$L = \lim_{y \to x} \frac{f(y) - f(x)}{y - x}$$

exists[22]. In that case we write $f'(x) = L$.

In general we mostly are interested in the situation where $D = I$ is an interval (open or closed). In that case we have some equivalent conditions:

Lemma 10.2. *Let $f : I \to \mathbb{R}$ be a function defined on any open interval $I = (a, b)$. Then the following are equivalent for $x \in I$ and $L \in \mathbb{R}$.*

(1) *f is differentiable at x and $L := f'(x)$*
(2) $\lim_{h \to 0} \frac{f(x+h)-f(x)}{h} = L$
(3) *$L \in \mathbb{R}$ is such that such that*
$$\lim_{y \to x} \frac{|f(y) - f(x) - L(y - x)|}{|y - x|} = 0.$$

The formulation of (3) has the advantage that it is easily generalizable to (higher dimensional) vector spaces (cf. Frechet derivative). E.g. if $f : \mathbb{R}^n \to \mathbb{R}^m$ then this definition makes complete sense if L is a $\mathbb{R}^{m \times n}$-matrix and $L(y - x)$ is the matrix product $\mathbb{R}^{m \times n}\ \mathbb{R}^n \subset \mathbb{R}^m$. Observe that $\frac{f(x+h)-f(x)}{h}$ does not make any sense if h is a vector!

All formulations are also equivalent in closed intervals $[a, b]$; for $x = a$, condition (2) changes to
$$\lim_{h \to 0^+} \frac{f(a + h) - f(a)}{h} = L$$
for $x = b$, condition (2) changes to
$$\lim_{h \to 0^-} \frac{f(b + h) - f(b)}{h} = L$$

Proof. Observe that any $x \in I$ is a cluster point of I.

(1) $\Leftrightarrow$ (2):

Setting $g(h) := \frac{f(x+h)-f(x)}{h}$ we observe that $h = 0$ is a cluster point of the domain of g which is $D_g = \{h \in \mathbb{R} : x + h \in I\}$. Also observe that if $y \in I$, then $h := y - x \in D_g$, and we have
$$g(y - x) = \frac{f(y) - f(x)}{y - x}.$$
Moreover, for $h \in D_g$ we have that $y = x + h \in I$, and
$$g(h) = \frac{f(x + h) - f(x)}{x + h - x}.$$

[22]Observe: Since x is a cluster point of D, then
$$g(y) := \frac{f(y) - f(x)}{y - x}$$
is defined in $D\backslash\{x\}$ and x is still a cluster point of $D\backslash\{x\}$.

By the limit laws we then find

$$\lim_{h\to 0} g(h) = L \quad \Leftrightarrow \quad \lim_{y\to x} g(y-x) = L \quad \Leftrightarrow \quad \lim_{y\to x} \frac{f(y)-f(x)}{y-x} = L$$

This implies (2).

<u>(1) ⇔ (3)</u>: By the limit laws, we have

$$\begin{aligned} &\lim_{y\to x} \frac{f(y)-f(x)}{y-x} = L \\ \Leftrightarrow &\lim_{y\to x} \left| \frac{f(y)-f(x)}{y-x} - L \right| = 0 \\ \Leftrightarrow &\lim_{y\to x} \frac{|f(y)-f(x)-L(y-x)|}{|y-x|} = 0. \end{aligned}$$

□

Example 10.3.

- Set $f(x) := \ln x$ then $f'(x) = \frac{1}{x}$ for any $x > 0$. Indeed,

$$\begin{aligned} \frac{f(x+h)-f(x)}{h} =& \frac{\ln(x+h)-\ln(x)}{h} = \frac{1}{h}\ln\left(\frac{x+h}{x}\right) \\ =& \ln\left(\frac{x+h}{x}\right)^{\frac{1}{h}} \\ =& \frac{1}{x}\ \ln\left(\left(1+\frac{h}{x}\right)^{\frac{x}{h}}\right) \end{aligned}$$

Now we observe

$$\begin{aligned} &\lim_{h\to 0} \ln\left(\left(1+\frac{h}{x}\right)^{\frac{x}{h}}\right) \\ =& \ln \lim_{h\to 0}\left(\left(1+\frac{h}{x}\right)^{\frac{x}{h}}\right) \\ =& \ln \lim_{z\to\infty}\left(\left(1+\frac{1}{z}\right)^{z}\right) \\ =& \ln e = 1 \end{aligned}$$

Exercise 10.4. *Use the limit definition above to show that for*

$$f(x) = x^2$$

we have $f'(x) = 2x$.

From Calculus we know that in order to be differentiable a function needs to be at least continuous. Now we can prove it:

Lemma 10.5 (Differentiability implies continuity)**.** *Assume that $f:(a,b)\to\mathbb{R}$ is differentiable at $x\in(a,b)$. Then f is continuous at x.*

Proof. Assume that f is differentiable at x, then for some $L\in\mathbb{R}$,

$$\lim_{(a,b)\ni y\to x}\frac{|f(y)-f(x)-L(y-x)|}{|y-x|}=0.$$

That is, for any $\varepsilon>0$ there exists $\delta>0$ such that

$$\frac{|f(y)-f(x)-L(y-x)|}{|y-x|}<\varepsilon\quad\forall y\in(a,b)\backslash\{x\}:\,|x-y|<\delta.$$

In order show that f is continuous, let $\varepsilon>0$ be given. Let δ be from the differentiability condition and set $\delta_1:=\min\{\delta,1\}$. Then

$$|f(y)-f(x)|<(\varepsilon+|L|)\,|x-y|\quad\forall y\in(a,b)\backslash\{x\}:\,|x-y|<\delta_1,$$

Set $\gamma:=\min\{\delta_1,\frac{\varepsilon}{|L|}\}$. Since $\gamma\le\delta_1$ we still have

$$|f(y)-f(x)|<(\varepsilon+|L|)\,|x-y|\quad\forall y\in(a,b)\backslash\{x\}:\,|x-y|<\gamma,$$

Observe that $|L|\gamma\le\varepsilon$ and $\varepsilon\gamma\le\varepsilon\delta_1\le\varepsilon$, then we have

$$|f(y)-f(x)|<2\varepsilon\quad\forall|x-y|<\gamma.$$

that is, we have shown

$$\forall\varepsilon>0,\quad\exists\gamma=\gamma(\varepsilon)>0:\quad|f(y)-f(x)|<\varepsilon\quad\forall y:\,|x-y|<\gamma.$$

This is exactly the definition of f being continuous at x. □

Even more that Lemma 10.5 is true: differentiable functions are locally Lipschitz continuous, see Exercise 10.6. A famous theorem, Rademacher's theorem, provides the opposite direction. Lipschitz functions are differentiable at almost all points.

Exercise 10.6. *Assume that $f:(a,b)\to\mathbb{R}$ is differentiable at $x\in(a,b)$. Then f is locally Lipschitz continuous around x, i.e. there exists $L>0$ and $\delta>0$ such that*

$$|f(x)-f(y)|\le L|x-y|\quad\text{for all } y\in(a,b),\ |x-y|<\delta.$$

Proposition 10.7. *Let $f,g:I\to\mathbb{R}$ be differentiable at $c\in I$, $I=(a,b)$. Then*

(1) λf is differentiable at c and $(\lambda f)'(c)=\lambda f'(c)$.
(2) $f+g$ is differentiable at c and $(f+g)'(c)=f'(c)+g'(c)$.
(3) fg is differentiable at c and $(fg)'(c)=f'(c)g(c)+f(c)g'(c)$.
(4) $\frac{f}{g}$ is differentiable if $g(c)\neq0$ at c and $\left(\frac{f}{g}\right)'(x)=\frac{f'(c)g(c)-f(c)g'(c)}{(g(c))^2}$.
(5) If $f:J\to\mathbb{R}$ and $g:I\to\mathbb{R}$ with $g(I)\subset J$ then if g is differentiable at $c\in I$ and f is differentiable at $g(c)$ then $f(g(\cdot)):I\to\mathbb{R}$ is differentiable at $x=c$ and we have $(f(g(\cdot)))'(c)=f'(g(c))\,g'(c)$.

Proof. (1) exercise

(2) exercise
(3) fg is differentiable at c and $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$:
We have

$$f(c+h)g(c+h) - f(c)g(c) = (f(c+h) - f(c))g(x+h) + f(c)(g(c+h) - g(c)).$$

Since f, g is differentiable at c we have

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

$$\lim_{h\to 0} \frac{g(c+h) - g(c)}{h} = g'(c)$$

Moreover, by Lemma 10.5, g is continuous at c, so

$$\lim_{h\to 0} g(c+h) = g(c).$$

From the limit laws we then find

$$\begin{aligned}(fg)'(c) =& \lim_{h\to 0} \frac{f(c+h)g(c+h) - f(c)g(c)}{h} \\ =& \lim_{h\to 0} \frac{(f(c+h) - f(c))g(c+h) + f(c)(g(c+h) - g(c))}{h} \\ =& \lim_{h\to 0} \frac{f(c+h) - f(c)}{h} \lim_{h\to 0} g(c+h) + f(c) \lim_{h\to 0} \frac{g(c+h) - g(c)}{h} \\ =& f'(c)\, g(c) + f(c)g'(c)\end{aligned}$$

Which is what was claimed.
(4) exercise
(5) exercise

□

Exercise 10.8. *Prove Proposition 10.7(1), (2), (4), (5).*

10.1. further exercises.

Exercise. *Use the limit definition of derivative to show*

- $f(x) = \frac{x+1}{x+4}$ *then* $f'(x) = \frac{3}{(t+4)^2}$
- $f(x) = 5$ *then* $f'(x) = 0$
- $f(x) = \frac{1}{x}$ *then* $f'(x) = -\frac{1}{x^2}$
- $f(x) = \sqrt{x}$ *then* $f'(x) = \frac{1}{2\sqrt{x}}$.

FIGURE 11.1. Pierre de Fermat, 1608-1665. French, lawyer, amateur mathematician.

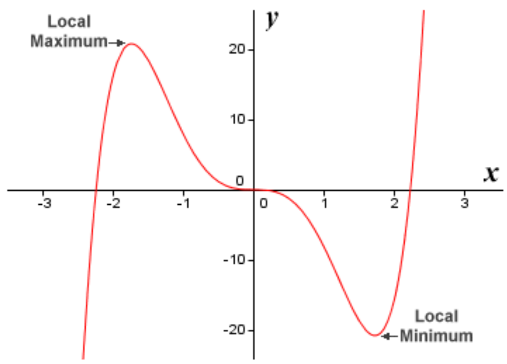

FIGURE 11.2. Fermats theorem: the derivative of a differentiable function is zero at its local maximum and minimum points. However, the derivative can be zero at points which are not the maximum and minimum, these are called stationary (or critical) points.

11. FERMAT'S THEOREM

This is another theorem we know from Calculus: If $f : (a, b) \to \mathbb{R}$ is differentiable at $c \in (a, b)$ and if f has a local maximum or minimum, then $f'(c) = 0$.

To make this statement precise, let us begin with

Definition 11.1 (Local maximum/minimum)**.** Let $f : D \to \mathbb{R}$ be a function.

- We say that f has a *local maximum* at $c \in D$ if there exists $\delta > 0$ such that
$$f(x) \geq f(c) \quad \forall x \in D : |x - c| < \delta.$$
- We say that f has a *local minimum* at $c \in D$ if there exists $\delta > 0$ such that
$$f(x) \leq f(c) \quad \forall x \in D : |x - c| < \delta.$$
- if f has either a local maximum or a local minimum at $c \in D$ then we say f has a *local extremum* at $c \in D$.

See Figure 11.2.

Theorem 11.2 (Fermat)**.** *Let $f : (a,b) \to \mathbb{R}$, f differentiable at $c \in (a,b)$ which is a local extremum. Then $f'(c) = 0$.*

Proof. Assume that c is a local minimum (the maximum argument follows the same strategy).

Observe that since (a,b) is an open set and $c \in (a,b)$ we have that $0 < \delta_0 := \min\{c-a, b-c\}$, and $c + h \in (a,b)$ for any $|h| < \delta_0$.

Moreover, since c is a local minimum of f, there exists $\delta_1 > 0$ such that

$$f(y) - f(c) \leq 0 \quad \forall y \in (a,b),\ |c - y| < \delta_1.$$

set $\delta := \frac{1}{2}\min\{\delta_0, \delta_1\}$. For $h > 0$, $|h| < \delta$ we then have

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

Since we know that f is differentiable at c, we conclude that

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

On the other hand, for $h < 0$ and $|h| < \delta$ we have

$$\frac{f(c+h) - f(c)}{h} \geq 0.$$

and from the differentiability of f at c we find

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

Thus $f'(c) = 0$. □

Example 11.3.
- The typical example to show this is false if f is not differentiable is $f(x) = |x|$ which has a local minimum at $x = 0$ but there is no sort of derivative equal to zero at $x = 0$[23].
- If $f(x) = x^3$ then $f'(0) = 0$, but the point $c = 0$ is neither a local minimum or a local maximum. We call points c where $f'(c) = 0$ *critical points* (because they might be min or max, or might not be either).
- Take $f(x) = x$ on the interval $[0,1]$. f attains its minimum at $x = 0$ and maximum at $x = 1$, however $f'(0) = f'(1) = 1 \neq 0$. This is a not a contradiction to Theorem 11.2 which is stated on *open intervals*. So we see: It is important that the interval is open, otherwise Theorem 11.2 is false. However see Exercise 11.4.
- From the picture, Figure 11.2 we see that $f'(x) = 0$ corresponds to horizontal tangent planes (we learned this from calculus).

[23]well, there is: the notion of *subdifferential* can be used here

Exercise 11.4. *Let $f : [a, b] \to \mathbb{R}$ be continuous, and assume that f has a local maximum at $c \in [a, b]$. Assume that f is differentiable at c. Show that*

(1) *If $c \in (a, b)$ then $f'(c) = 0$*
(2) *If $c = a$ then $f'(c) \leq 0$*
(3) *If $c = b$ then $f'(c) \geq 0$*
(4) *What can we say about the derivatives if f has a local minimum at c?*

Exercise 11.5. *Let $\mathcal{E} : \mathbb{R}^n \to \mathbb{R}$ be a function (this is often called the energy) and assume that for some $\bar{x} \in \mathbb{R}^n$ we have*

$$\mathcal{E}(\bar{x}) \leq \mathcal{E}(x) \quad \forall x \in \mathbb{R}^n.$$

Assume that for any $v \in \mathbb{R}^n$ the directional derivative exists, i.e.

$$\frac{d}{dt}\Big|_{t=0} \mathcal{E}(\bar{x} + tv) := \lim_{t \to 0} \frac{\mathcal{E}(\bar{x} + tv) - \mathcal{E}(\bar{x})}{t} \in \mathbb{R}.$$

Show that then

$$\frac{d}{dt}\Big|_{t=0} \mathcal{E}(\bar{x} + tv) = 0. \tag{11.1}$$

Equation (11.1) is essentially what is called the *Euler-Lagrange equation*

The above arguments works also if $\mathbb{R}^n$ is replaced with any other linear space, and (11.1) is a useful equation to derive properties of potential minimizers.

Exercise 11.6 (Euler-Lagrange equations)**.** *(be generous about the definitions of integrals below, we haven't defined it yet)*

Consider differentiable functions $f : [a, b] \to \mathbb{R}$ whose derivative is integrable (whatever that means), and consider the energy

$$\mathcal{E}(f) = \frac{1}{2} \int_{(a,b)} |f'|^2 dx.$$

Assume that $\bar{f} : [a, b] \to \mathbb{R}$ is a minimizer

$$\mathcal{E}(\bar{f}) \leq \mathcal{E}(f) \quad \forall f : [a, b] \to \mathbb{R} \quad \text{such that } f(a) = \bar{f}(a) \text{ and } f(b) = \bar{f}(b).$$

Then we have (assuming the second derivative makes sense and is continuous)

$$\bar{f}''(x) = 0 \quad \text{for all } x \in (a, b)$$

Hints:

- *Show that for any $\varphi : [a, b] \to \mathbb{R}$, $\varphi(a) = \varphi(b) = 0$ we have*

$$\frac{d}{dt}\Big|_{t=0} \mathcal{E}(\bar{f} + t\varphi) = 0.$$

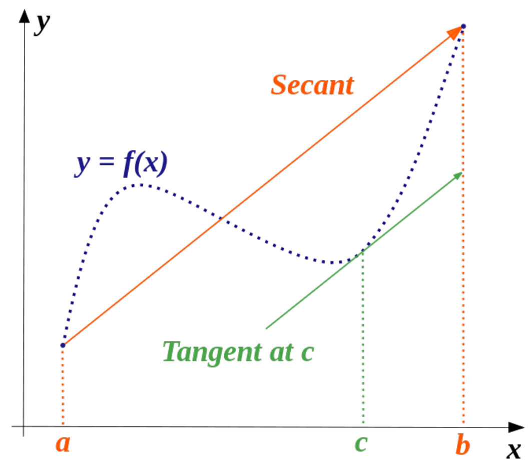

FIGURE 12.1. For any function that is continuous on $[a, b]$ and differentiable on (a, b) there exists some c in the interval (a, b) such that the secant joining the endpoints of the interval $[a, b]$ is parallel to the tangent at c. Source: Wikipedia, robert alexander ortiz, GFDL

- *Show that this implies (assuming derivative and integral converge)*

$$\int_{(a,b)} \bar{f}' \varphi' dx = 0$$

- *By an integration by parts show that this implies*

$$\int_{(a,b)} \bar{f}'' \varphi dx = 0$$

- *The above implies (try to prove it for fun) that $\bar{f}'' = 0$ – we will discuss this later Exercise 14.34.*

12. Mean Value Theorem

Fermat's theorem is incredibly important in Analysis (in particular, many "equilibria of systems" are often stationary points which may or may not be minimizers e.g. of sort of physical energies, cf. Exercise 11.5). There is a theoretical consequence which is the Mean Value Theorem, cf. Figure 12.1 – from calculus we know that the mean value theorem for suitable f says: for any a, b there exists ξ such that

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

As it happens often, mathematically the mean value theorem is a consequence of a much simplified situation. In this case its Rolle's theorem.

FIGURE 12.2. Michel Rolle, 1652-1719. French, "early critic of infinitesimal calculus, arguing that it was inaccurate, based upon unsound reasoning, and was a collection of ingenious fallacies, but later changed his opinion" (wikipedia).

Theorem 12.1 (Rolle)**.** *Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). If $f(a) = f(b) = 0$ then there exists $c \in (a,b)$ with $f'(c) = 0$.*

Proof. Since f is continuous on $[a,b]$ we know by the max-min theorem, Theorem 7.2, that there exists $c_{max}, c_{min} \in [a,b]$ where f attains its maximum and minimum respectively. Now we proceed by a case study.

- If c_{max} in (a,b) then by Fermat, Theorem 11.2, $f'(c_{max}) = 0$ and $c := c_{max}$ is (one of) the point we are looking for.
- If c_{min} in (a,b) then by Fermat, Theorem 11.2, $f'(c_{min}) = 0$ and $c := c_{min}$ is (one of) the point we are looking for.
- If neither c_{min} nor c_{max} are in (a,b) then $c_{min}, c_{max} \in \{a,b\}$. But then because of $f(a) = f(b) = 0$ we have
$$0 = f(c_{min}) \le f(x) \le f(c_{max}) = 0 \quad \forall x \in [a,b].$$
That is $f(x) = 0$ for all $x \in [a,b]$. In particular $f'(x) = 0$ for all $[a,b]$ (constant functions have zero derivative!). So $c := \frac{a+b}{2} \in (a,b)$ is the point we are looking for.

□

The mean value theorem is a consequence of Theorem 12.1.

Theorem 12.2. *Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). There exists $c \in (a,b)$ with*
$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Proof. The idea is to linearly change f into a g with $g(a) = 0$ and $g(b) = 0$. Namely,
$$g(x) = f(x) + Ax + B,$$

where we choose A and B such that $g(a) = 0$ and $g(b) = 0$, i.e. we need to solve for A and B the linear system

$$\begin{cases} 0 = f(a) + Aa + B \\ 0 = f(b) + Ab + B \end{cases}$$

Subtracting the second line from the first we find that $f(a) - f(b) = -A(a - b)$, i.e.

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

Having A we find that

$$B = -f(a) + a\frac{f(b) - f(a)}{b - a}.$$

With this choice of A and B the function g satisfies the assumptions of Theorem 12.1, so there exists $c \in (a, b)$ with

$$0 = g'(c) = f'(c) + A.$$

That is,

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

□

The mean value theorem has many applications. The most important part is that it allows us to relate properties of a function f by the properties of f'.

Proposition 12.3. *Let $f : (a, b) \to \mathbb{R}$ be a differentiable function.*

(1) If $f'(x) \geq 0$ for all $x \in I$ then f is monotonically increasing
(2) If $f'(x) \leq 0$ for all $x \in I$ then f is monotonically decreasing
(3) If $f'(x) > 0$ for all $x \in I$ then f is strictly monotonically increasing
(4) If $f'(x) < 0$ for all $x \in I$ then f is strictly monotonically decreasing
(5) If $f'(x) = 0$ for all $x \in I$ then f is constant.

Proof. (1) If $f'(x) \geq 0$ for all $x \in I$ then f is monotonically increasing

Let $a < x < y < b$. Then $f : [x, y] \to \mathbb{R}$ is continuous and differentiable, so by the mean value theorem, Theorem 12.2, there exists $c \in (x, y)$ with

$$\frac{f(y) - f(x)}{y - x} = f'(c) \geq 0. \tag{12.1}$$

Multiplying with $(y - x) > 0$ we find

$$f(y) \geq f(x) \quad \forall a < x < y < b.$$

This is the definition of monotonically increasing.

(2) If $f'(x) \leq 0$ for all $x \in I$ then f is monotonically decreasing:

Consider $g(x) := -f(x)$, then $g'(x) \geq 0$ so g is monotonically increasing by the above argument. That is f is montonically decreasing.

(3) If $f'(x) > 0$ for all $x \in I$ then f is strictly monotonically increasing
We argue as above, instead of (12.1) we obtain

$$\frac{f(y) - f(x)}{y - x} = f'(c) > 0,$$

that is

$$f(y) > f(x) \quad \forall x, y : \, a < x < y < b.$$

This is the definition of f being monotonically increasing.

(4) If $f'(x) < 0$ for all $x \in I$ then f is strictly monotonically decreasing
Same as above. Set $g(x) := -f(x)$.

(5) If $f'(x) = 0$ for all $x \in I$ then f is constant.
Since $f'(x) \geq 0$ we have from the above that f is montonically increasing, i.e.

$$f(y) \geq f(x) \quad \forall x \geq y.$$

Since on the other hand $f'(x) \leq 0$ we have

$$f(y) \leq f(x) \quad \forall x \geq y.$$

Together we obtain $f(y) = f(x)$ for all $x \geq y$, $x, y \in (a, b)$ which is the claim.

□

Exercise 12.4. *Show that*

$$\arctan\left(\frac{1+x}{1-x}\right) = \arctan(x) + \frac{\pi}{4} \quad \forall x \in (-\infty, 1)$$

Hint:, set

$$f(x) := \arctan\left(\frac{1+x}{1-x}\right) - \arctan(x),$$

show that f is constant.

Exercise 12.5. *Assume that $f : (a, b) \to \mathbb{R}$ is Hölder continuous with $\alpha > 1$, cf. Remark 6.18. I.e. assume there exists $\Lambda > 0$ such that*

$$|f(x) - f(y)| \leq \Lambda |x - y|^{\alpha}.$$

Show that f is constant.

Hint: Recall that $\alpha > 1$ – what is $f'(x)$?

From Calculus we know the following *first derivative test*.

Proposition 12.6 (First derivative test)**.** *Let $f \in (a, b)$ be continuous and differentiable. Assume that for some $c \in (a, b)$ we have $f'(x) \leq 0$ if $x \leq c$ and $f'(x) \geq 0$ if $x \geq c$. Then f has a minimum at at $x = c$, that is*

$$f(c) \leq f(x) \quad \forall x \in (a, b).$$

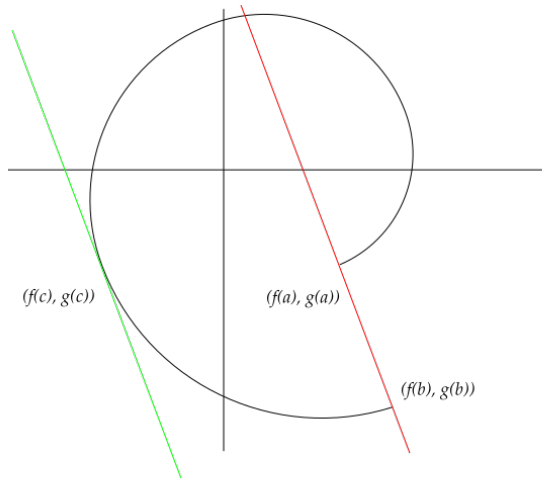

FIGURE 12.3. Cauchy's mean value theorem, Theorem 12.7, geometrically means the following: there is some tangent to the graph of the curve $\gamma : [a,b] \to \mathbb{R}^2$ given as $\gamma(t) := (f(t), g(t))$ which is parallel to the line defined by the points $(f(a), g(a))$ and $(f(b), g(b))$ – unless the curve becomes stationary at some point $c \in (a,b)$, i.e. $f'(c) = g'(c) = 0$. Source: wikipedia.

Proof. Let $x \in (c,b)$. Then by the mean value theorem, Theorem 12.2, for some $y \in (c,x)$,

$$\frac{f(c) - f(x)}{c - x} = f'(y) \geq 0.$$

Observe that $c < x$ so this implies

$$f(c) - f(x) \leq 0 \quad \Leftrightarrow \quad f(c) \leq f(x) \quad \forall x \in (c,b).$$

Let now $x \in (a,c)$, then there exists $y \in (a,c)$ such that

$$\frac{f(c) - f(x)}{c - x} = f'(y) \leq 0.$$

Now $c - x \geq 0$ so this implies

$$f(c) - f(x) \leq 0 \quad \Leftrightarrow \quad f(c) \leq f(x) \quad \forall x \in (a,c).$$

Since clearly $f(c) \leq f(x)$ for $x = c$ we have shown that

$$f(c) \leq f(x) \quad \forall x \in (a,b).$$

□

We also have the following "generalization" of the Mean Value Theorem (for $g(x) = x$ it is just the statement of the Mean Value Theorem).

Theorem 12.7 (Cauchy's mean value theorem)**.** *Let $f, g : [a, b] \to \mathbb{R}$ continuous in $[a, b]$ and differentiable in (a, b). Assume that $g(b) \neq g(a)$. Then there exists $c \in (a, b)$ such that*

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

If $g'(x) \neq 0$ for all $x \in (a, b)$ then this is equivalent to

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

Cf. Figure 12.3.

Proof. Like the mean value theorem, this statement follows from Rolle's theorem Theorem 12.1.

Let

$$h(x) := f(x) - \frac{f(b) - f(a)}{g(b) - g(a)} g(x).$$

Observe that $h(a) = h(b)$, indeed

$$\begin{aligned} &h(a) = h(b) \\ \Leftrightarrow &f(a) - \frac{f(b) - f(a)}{g(b) - g(a)} g(a) = f(b) - \frac{f(b) - f(a)}{g(b) - g(a)} g(b) \\ \Leftrightarrow &f(a) - f(b) - \frac{f(b) - f(a)}{g(b) - g(a)} (g(a) - g(b)) = 0 \end{aligned}$$

By the mean value theorem, Theorem 12.2, (or rather Rolle's theorem, Theorem 12.1) there must be $c \in (a, b)$ such that

$$h'(c) = 0.$$

That is

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

This is the claimed equation. □

Cauchy's version of the mean value theorem also implies the $\frac{0}{0}$-type for L'Hopital.

Proposition 12.8 (L'Hopital)**.** *Let $f, g : (a, b) \to \mathbb{R}$ differentiable and $g(x), g'(x) \neq 0$ for all (a, b). Assume that $\lim_{x \to a^+} f(x) = \lim_{x \to a^+} g(x) = 0$. If*

$$L := \lim_{x \to a+} \frac{f'(x)}{g'(x)} \text{ exists and } L \in \mathbb{R}$$

then

$$L := \lim_{x \to a+} \frac{f(x)}{g(x)}.$$

FIGURE 12.4. Guillaume de l'Hôpital. 1661 - 1704. Fench Mathematician

Proof. Since $\lim_{x\to a^+} f(x) = \lim_{x\to a^+} g(x) = 0$ we can assume that $f, g : [a, b) \to \mathbb{R}$ is continuous with $f(0) = g(0) = 0$ (we extend f and g into a). By Cauchy's mean value theorem, Theorem 12.7, for any $x \in (a, b)$ there exists $c \in (a, x)$ such that

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

Observe that as $x \to a$ we necessarily have $c \to a$ (since $c \in (a, x)$, using the squeeze theorem). That is,

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = \lim_{c\to a^+} \frac{f'(c)}{g'(c)} = L,$$

since the right-hand side exists. □

As we learned in Calculus, we derive the other types of L'Hopital's rule from the $\frac{0}{0}$-case, e.g.

Theorem 12.9 (L'Hopital $\frac{\infty}{\infty}$)**.** *Show the following: for any* $0 < a < b \leq \infty$*:*

Let $f, g : (a, b) \to \mathbb{R}$ *differentiable and* $g(x), g'(x) \neq 0$ *for all* (a, b)*. Assume that* $\lim_{x\to b^-} f(x) = \lim_{x\to b^-} g(x) = \infty$*. If*

$$L := \lim_{x\to b^-} \frac{f'(x)}{g'(x)} \text{ exists and } L \in \mathbb{R}$$

then

$$L := \lim_{x\to b^-} \frac{f(x)}{g(x)}.$$

Proof. This is actually more complicated than the $\frac{0}{0}$-case, so here is the solution inspired by [ht]:

Fix $\varepsilon > 0$. Since

$$L := \lim_{x\to b^-} \frac{f'(x)}{g'(x)}$$

there exists $\delta > 0$ such that

$$\left|\frac{f'(z)}{g'(z)} - L\right| < \varepsilon \quad \forall z \in (b - \delta, b).$$

Fix now, $x, y \in (b-\delta, b)$. Then we have by Cauchy's mean value theorem, Theorem 12.7, for some $z = z(x,y) \in (b-\delta, b)$,

$$\left|\frac{f(x)-f(y)}{g(x)-g(y)} - L\right| = \left|\frac{f'(z)}{g'(z)} - L\right| < \varepsilon$$

Let us write this again,

$$\left|\frac{f(x)-f(y)}{g(x)-g(y)} - L\right| < \varepsilon \quad \forall x, y \in (b-\delta, b).$$

Now we have

$$\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{\frac{f(x)}{g(x)} - \frac{f(y)}{g(x)}}{1 - \frac{g(y)}{g(x)}}$$

And we have

$$\frac{f(x)}{g(x)} = \frac{\frac{f(x)}{g(x)} - \frac{f(y)}{g(x)}}{1 - \frac{g(y)}{g(x)}} \left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)}$$

and thus

$$\begin{aligned}\frac{f(x)}{g(x)} - L &= \left(\frac{\frac{f(x)}{g(x)} - \frac{f(y)}{g(x)}}{1 - \frac{g(y)}{g(x)}} - L\right)\left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)} - L + L\left(1 - \frac{g(y)}{g(x)}\right)\\ &= \left(\frac{\frac{f(x)}{g(x)} - \frac{f(y)}{g(x)}}{1 - \frac{g(y)}{g(x)}} - L\right)\left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)} - L\left(\frac{g(y)}{g(x)}\right)\end{aligned}$$

That is, for any $x, y \in (b-\delta, b)$, (keep in mind that $g(x) \xrightarrow{x\to b^-} \infty$)

$$\left|\frac{f(x)}{g(x)} - L\right| \leq \underbrace{\left|\frac{\frac{f(x)}{g(x)} - \frac{f(y)}{g(x)}}{1 - \frac{g(y)}{g(x)}} - L\right|}_{\leq \varepsilon} \underbrace{\left|1 - \frac{g(y)}{g(x)}\right|}_{\xrightarrow{x\to b^-} 1} + \underbrace{\left|\frac{f(y)}{g(x)}\right|}_{\xrightarrow{x\to b^-} 0} + L \underbrace{\left|\frac{g(y)}{g(x)}\right|}_{\xrightarrow{x\to b^-} 0}$$

Keeping $y \in (b-\delta, b)$ fixed we now take the $\limsup_{x\to b^-}$ in the above inequality. Then

$$\limsup_{x\to b^-} \left|\frac{f(x)}{g(x)} - L\right| \leq \varepsilon 1 + 0 + L0 = \varepsilon.$$

That is we have

$$\forall \varepsilon > 0 : \limsup_{x\to b^-} \left|\frac{f(x)}{g(x)} - L\right| \leq \varepsilon$$

Clearly this implies

$$\limsup_{x\to b^-} \left|\frac{f(x)}{g(x)} - L\right| = 0$$

and thus

$$\lim_{x\to b^-} \frac{f(x)}{g(x)} = L.$$

□

Exercise 12.10 (Otto Stolz counterexample)**.** *Let*

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

$$g(x) := e^{\sin\left(\frac{1}{x}\right)}\left(\frac{1}{x} + \cos\left(\frac{1}{x}\right)\sin\left(\frac{1}{x}\right)\right)$$

(1) Show that

$$\lim_{x\to 0^+} \frac{f'(x)}{g'(x)} = 0$$

(2) however show that

$$\lim_{x\to 0^+} \frac{f(x)}{g(x)} \quad \textit{does not exist}$$

(3) Why is this no contradiction to Proposition 12.8?

Theorem 12.11. *Let $f : [a,b] \to \mathbb{R}$ be continuous and differentiable in $[a,b]$. Assume that L is a value strictly between $f'(a)$ and $f'(b)$,*

$$\min\{f'(a), f'(b)\} < L < \max\{f'(a), f'(b)\}.$$

Then there exists $c \in (a,b)$ with $f'(c) = L$.

This looks a lot like the *intermediate value theorem* for f', Theorem 8.2. But observe that f' may not be continuous so Theorem 8.2 is not applicable! Observe that this statement is false e.g. for $f(x) = |x|$.

Proof. Fix L strictly between $f'(a)$ and $f'(b)$. Assume w.l.o.g.

$$f'(a) < L < f'(b). \tag{12.2}$$

We set

$$g(x) := Lx - f(x).$$

Why do we do this? Because then

$$L = f'(x) \quad \Leftrightarrow \quad g'(x) = 0.$$

So we need to find critical points of g, and for this we can use min-max theorem and Fermat's theorem:

The function $g : [a,b] \to \mathbb{R}$ is continuous and differentiable in $[a,b]$. By the min-max theorem, Theorem 7.2, there exists $x_{max} \in [a,b]$ with where g attains its global maximum. If we can ensure that $x_{max} \in (a,b)$ then Fermat's theorem, Theorem 11.2, implies that $g'(x_{max}) = 0$, that is $L = f'(x_{max})$.

So it only remains to show that $x_{max} \notin \{a,b\}$. Indeed, observe that

$$0 \overset{(12.2)}{<} L - f'(a) = g'(a) = \lim_{x\to a^+} \frac{g(x)-g(a)}{x-a}.$$

In particular there must be some $x_1 \in (a,b)$ such that

$$0 < \frac{g(x_1) - g(a)}{x_1 - a} \quad \text{or, equivalently,} \quad g(x_1) > g(a).$$

Similarly,

$$0 \overset{(12.2)}{>} L - f'(b) = g'(b) = \lim_{x \to b^-} \frac{g(x) - g(b)}{x - b}.$$

which implies that there must be some $x_2 \in (a,b)$ such that (using that $x_2 - b < 0$)

$$0 > \frac{g(x_2) - g(b)}{x_2 - b} \quad \Leftrightarrow \quad g(x_2) > g(b).$$

That is a and b are no maxima for g, that is $x_{max} \notin \{a,b\}$. □

13. Continuous and differentiable function spaces

We have defined what it means for $f : I \to \mathbb{R}$ to be differentiable. If $f' : I \to \mathbb{R}$ is again differentiable, we say that f is twice differentiable, etc. By $f^{(n)}$ we denote the n-th derivative, $f^{(0)} = f$, $f^{(1)} = f'$, $f^{(2)} = f''$ etc.

Definition 13.1. Let $k \in \mathbb{N} \cup \{0\}$. We say that $f \in C^k(D)$, in words, f is *k times continuously differentiable*, if f is k times differentiable in D, and $f, f', \ldots, f^{(k)}$ are continuous in D.

If f is k-times differentiable for any $k \in \mathbb{N}$ we say that f is infinitely many times differentiable, and write $f \in C^\infty(I)$.

Example 13.2.

- $f \in C^0(D)$ means that f is continuous. If D is a closed interval $D = [a,b]$ then this implies that f is uniformly continuous,Theorem 9.4. If $D = (a,b)$ is open, then $f \in C^0(I)$ might not be uniformly continuous.
- if $f \in C^k(D)$ then $f \in C^{k-1}(D)$.
- if $f \in C^k([a,b])$ then f, f^1,. . .,f^k are uniformly continuous and bounded, assuming as always that $-\infty < a < b < \infty$.
- A non-polynomial function which is infinitely many times differentiable is

$$f(x) = \begin{cases} 0 & |x| \geq 1 \\ e^{\frac{1}{x^2-1}} & |x| < 1. \end{cases}$$

  This is called a *bump function* (also used as *mollifier* function). Observe that $f(0) = 1$. See Figure 13.1. These kind of functions are used often in analysis to localize differential equations or mollify (smoothen) functions.

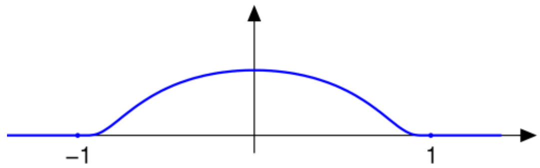

FIGURE 13.1. A bump function: a C^∞-smooth function f with $f(0) = 1$ and $f(x) = 0$ for $|x| > 1$.

FIGURE 14.1. Bernhard Riemann, 1826 - 1866. He is considered by many to be one of the greatest mathematicians of all time.

FIGURE 14.2. Jean Gaston Darboux, 1842 - 1917. French, mathematician

14. The Riemann integral

We are going to introduce (a version of) the *Riemann integral*, which is the area below a curve. Before we come to the fundamental theorem of calculus, this has nothing to do with antiderivatives. The area below a curve is approximated by area boxes, cf. Figure 14.3; the lower and upper *Darboux sum*.

Definition 14.1 (Partition)**.** A *partition of size n* of the interval $[a, b]$ is a set of numbers $\{x_0, x_1, \ldots, x_n\}$ such that

$$a = x_0 < x_1 < x_2 \ldots < x_{n-1} < x_n = b.$$

We write

$$\Delta x_i := x_i - x_{i-1}, \quad i \geq 1.$$

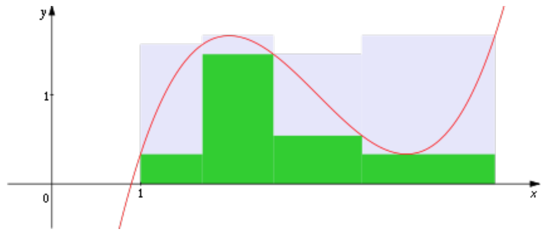

FIGURE 14.3. We approximate the sum below the (red) curve by boxes. The area of the green boxes is the lower Darboux sum, the area of the grey boxes plus the green boxes is the upper Darboux sum, cf Definition 14.2

We plan approximate the area below a curve from above and below, cf. Figure 14.3. For this, we define

Definition 14.2 (Darboux sums)**.** Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Let P be a partition of $[a, b]$. Define

$$m_i := \inf_{[x_{i-1},x_i]} f, \quad M_i := \sup_{[x_{i-1},x_i]} f.$$

Then the *lower Darboux sum* is defined

$$L(P, f) := \sum_{i=1}^{n} m_i \, \Delta x_i.$$

and the *upper Darboux sum* is defined by

$$U(P, f) := \sum_{i=1}^{n} M_i \, \Delta x_i$$

See Figure 14.3.

From pictures, e.g. Figure 14.3, it seems obvious that any lower Darboux sum delivers a smaller area than the actual area below the curve. and any upper darboux sum delivers a larger area than the area below the curve. The idea of the Riemann integral is that if we just take a fine enough partition, then upper and lower Darboux sum should approximate the actual area below the curve. We will later see this is true for continuous functions.

Definition 14.3. Let $f : [a, b] \to \mathbb{R}$ be bounded.

We define the *upper Riemann integral*,

$$\overline{\int_{[a,b]}} f(x) \, dx := \inf_{\text{P partition}} U(P, f).$$

and the *lower Riemann integral*

$$\underline{\int_{[a,b]}} f(x) \, dx := \sup_{\text{P partition}} L(P, f).$$

We say that $f : [a, b] \to \mathbb{R}$ is *Riemann integrable* (notation: $f \in \mathcal{R}([a, b])$ if

$$\overline{\int_{[a,b]}} f(x)\,dx = \underline{\int_{[a,b]}} f(x)\,dx \in (-\infty, \infty).$$

In this case we write

$$\int_{[a,b]} f(x)\,dx = \overline{\int_{[a,b]}} f(x)\,dx = \underline{\int_{[a,b]}} f(x)\,dx$$

To make sense of the above notation, first we observe some basic properties of L and U and the upper and lower Riemann integral.

First we observe that the supremums/infimum condition for the upper/lower Riemann integral implies that the partition chosen can be assumed to be arbitrarily fine (meaning that the Δx_i can be assumed to be arbitrarily small).

Lemma 14.4 (Refinement property)**.** *Let $f : [a, b] \to \mathbb{R}$ be a function and P be a partition of $[a, b]$. Let Q be a* refinement *of P, i.e. Q is another partition of $[a, b]$ with $Q \supset P$. Then*

$$L(P, f) \leq L(Q, f)$$

and

$$U(P, f) \geq U(Q, f)$$

Proof. By an induction argument we can assume that Q only has one more element than P, i.e. $P = \{x_0, \ldots, x_N\}$ and $Q = P \cup \{y\}$ with $x_n < y < x_{n+1}$ for some $n = 0, \ldots, N-1$. Recall that we have $\inf_A f \leq \inf_B f$ for $A \supset B$. Thus,

$$\begin{aligned}
&\left(\inf_{[x_n,x_{n+1}]} f\right)(x_{n+1} - x_n)\\
=&\left(\inf_{[x_n,x_{n+1}]} f\right)(x_{n+1} - y) + \left(\inf_{[x_n,x_{n+1}]} f\right)(y - x_n)\\
\leq&\left(\inf_{[y,x_{n+1}]} f\right)(x_{n+1} - y) + \left(\inf_{[x_n,y]} f\right)(y - x_n)
\end{aligned}$$

Now by the definition of $L(P, f)$ and $L(Q, f)$ we see that $L(P, f) \leq L(Q, f)$.

A similar argument shows that $U(P, f) \leq U(Q, f)$. □

Lemma 14.5. *Let $f : [a, b] \to \mathbb{R}$ bounded.*

(1) If $f(x) = c$ (i.e. f is contant) then

$$\int_{[a,b]} f(x)\,dx = c(b - a).$$

(2) If $f(x) \leq g(x)$ for all $x \in [a,b]$ then

$$\underline{\int}_{[a,b]} f(x)\,dx \leq \underline{\int}_{[a,b]} g(x)\,dx$$

and

$$\overline{\int}_{[a,b]} f(x)\,dx \leq \overline{\int}_{[a,b]} g(x)\,dx$$

In particular if f and g are Riemann integrable then

$$\int_{[a,b]} f(x)\,dx \leq \int_{[a,b]} g(x)\,dx$$

(3) We have

$$(b-a)\inf_{[a,b]} f \leq \underline{\int}_{[a,b]} f(x)\,dx \leq \overline{\int}_{[a,b]} f(x)\,dx \leq (b-a)\sup_{[a,b]} f$$

In particular if f is Riemann integrable

$$(b-a)\inf_{[a,b]} f \leq \int_{[a,b]} f(x)\,dx \leq (b-a)\sup_{[a,b]} f$$

Proof. (1) If $f(x) = c$ then for any partition $P = \{x_0, \dots, x_n\}$ we have $m_i = M_i = c$, so that

$$L(P,f) = U(P,f) = \sum_{i=1}^{n} c\,\Delta x_i = c\,(b-a).$$

In particular,

$$\underline{\int}_{[a,b]} c\,dx = c(b-a) = \overline{\int}_{[a,b]} c\,dx.$$

Thus, by definition,

$$\int_{[a,b]} c\,dx = c(b-a).$$

(2) Let $f(x) \leq g(x)$ for all $x \in [a,b]$ then for any partition $P = \{x_0, \dots, x_n\}$ we have $m_i(f) \leq m_i(g)$ and $M_i(f) \leq M_i(g)$. This readily implies

$$L(P,f) \leq L(P,g), \quad \text{and} \quad U(P,f) \leq U(P,g).$$

Since $\underline{\int}_{[a,b]} dx$ is the supremum over all partitions P, we conclude

$$L(P,f) \leq \underline{\int}_{[a,b]} g(x)\,dx \quad \forall \text{partitions } P.$$

Taking the supremum over all P of this inequality, we have

$$\underline{\int}_{[a,b]} f(x)\,dx \leq \underline{\int}_{[a,b]} g(x)\,dx.$$

We argue similarly to obtain

$$\overline{\int}_{[a,b]} f(x)\,dx \leq \overline{\int}_{[a,b]} g(x)\,dx.$$

(3) follows from (1) and (2).

□

Exercise 14.6. [Leb, 5.1.3] *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Suppose that there exists a sequence of partitions $\{P_k\}$ of $[a, b]$ such that*

$$\lim_{k\to\infty} (U(P_k, f) - L(P_k, f)) = 0.$$

Show that f is Riemann integrable and that

$$\int_a^b f = \lim_{k\to\infty} U(P_k, f) = \lim_{k\to\infty} L(P_k, f).$$

The above criterion is sometimes called the Cauchy criterion

Exercise 14.7. [Leb, ex. 5.1.6] *Let $c \in (a, b)$ and let $d \in \mathbb{R}$. Define $f : [a, b] \to \mathbb{R}$ as*

$$f(x) := \begin{cases} d & \text{if } x = c, \\ 0 & \text{if } x \neq c. \end{cases}$$

Prove that f is Riemann integrable and compute $\int_{[a,b]} f$ using the definition of the integral and, if you want, Exercise 14.6.

Exercise 14.8. [Leb, Ex. 5.2.1] *Let f be Riemann integrable on $[a, b]$. Prove that $-f$ is Riemann integrable on $[a, b]$ and that*

$$\int_{[a,b]} -f(x)\,dx = -\int_{[a,b]} f(x)\,dx.$$

Example 14.9. Lemma 14.5 implies that for any bounded function $f : [a, b] \to \mathbb{R}$ the upper and lower Riemann integral exists. However, the Riemann integral may not exist for discontinuous functions.

Take the Dirichlet function, $D : [0, 1] \to \mathbb{R}$

$$D(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R}\backslash\mathbb{Q}. \end{cases}$$

Observe that for any partition $P = \{x_1, \ldots, x_n\}$ we have $M_i = 1$ and $m_i = 0$, by density of $\mathbb{Q}$ and $\mathbb{R}\backslash\mathbb{Q}$ in $[0, 1]$. So we have

$$\underline{\int_{[a,b]}} f(x)\,dx = 0 < \overline{\int_{[a,b]}} f(x)\,dx = 1.$$

With our current definition, f needs to be bounded for the Riemann integral to exist. In particular $\frac{1}{\sqrt{x}}$ is not integrable in $[0, 1]$ – even though we know from calculus that $\int_0^1 \frac{1}{\sqrt{x}} dx = 2\sqrt{x}\Big|_0^1 = 1$. To make sense of this one would need to introduce the notion of *improper integral*. We will not do that now, but assume that f is always bounded, justified by the following

Lemma 14.10. *Let $f : [a,b] \to \mathbb{R}$ be Riemann-integrable. Then f is bounded*[24].

More precisely, if $f : [a,b] \to \mathbb{R}$ is unbounded then at least one of the following is true:

$$\inf_{P \text{ partition of } [a,b]} U(P,f) = +\infty,$$

or

$$\sup_{P \text{ partition of } [a,b]} L(P,f) = -\infty,$$

Proof. Assume that f is unbounded. Then there exist $(z_n)_{n\in\mathbb{N}} \subset [a,b]$ such that $f(z_n) \xrightarrow{n\to\infty} \pm\infty$. For simplicity assume that $f(z_n) \xrightarrow{n\to\infty} +\infty$ (the $-\infty$ case goes similar).

By Bolzano-Weierstrass, Theorem 3.8, we may assume that (up to taking a subsequence) $z_n \xrightarrow{n\to\infty} z \in [a,b]$.

Let $P = \{x_0, \dots, x_N\}$ be any partition of $[a,b]$. Since $z \in [a,b]$ there must be some $m \in \{1, \dots, N\}$ such that $z_n \in [x_{m-1}, x_m]$ for infinitely many $n \in \mathbb{N}$. Taking a subsequence we can assume w.l.og. that $z_n \in [x_{m-1}, x_m]$. But this implies that

$$\max_{[x_{m-1},x_m]} f = \infty.$$

So $U(P,f) = \infty$. This holds for any partition P, so $\inf_P U(P,f) = \infty$ (and in particular f is not Riemann integrable).

□

So unbounded functions are never integrable.

For a positive result: continuous functions on closed bounded sets $[a,b]$ are always integrable.

Proposition 14.11. *Let $f : [a,b] \to \mathbb{R}$ be continuous. Then f is Riemann-integrable.*

Proof. Since $f : [a,b] \to \mathbb{R}$ is continuous it is bounded, by Theorem 7.2 (or Corollary 7.4). So we already have from Lemma 14.5

$$-\infty < \underline{\int_{[a,b]}} f(x)\,dx \le \overline{\int_{[a,b]}} f(x)\,dx < \infty$$

We are now going to show that for any $\varepsilon > 0$

$$\overline{\int_{[a,b]}} f(x)\,dx \le \underline{\int_{[a,b]}} f(x)\,dx + \varepsilon(b-a). \tag{14.1}$$

If we can do so, we can let ε go to zero, and conclude that $\overline{\int_{[a,b]}} f(x)\,dx = \underline{\int_{[a,b]}} f(x)\,dx$.

[24]this is kind of non-suprising, since in the definition of Riemann-integrability we assume boundedness

Let $\varepsilon > 0$. Since $f : [a,b] \to \mathbb{R}$ is continuous, by Theorem 9.4 we have that f is uniformly continuous. That is, there exists $\delta > 0$ such that

$$|f(x) - f(y)| < \varepsilon \quad \forall x, y \in [a,b];\ |x-y| < \delta. \tag{14.2}$$

Let $P = \{x_1, \ldots, x_N\}$ be a partition of $[a,b]$ such that $\Delta x_i < \delta$. For any i, by the min-max theorem, *Theorem* 7.2, there exists y_i and z_i in $[x_{i-1}, x_i]$ such that

$$M_i = \sup_{[x_{i-1},x_i]} f = f(y_i)$$

and

$$m_i = \inf_{[x_{i-1},x_i]} f = f(z_i).$$

Thus, by (14.2)

$$|M_i - m_i| = |f(y_i) - f(z_i)| < \varepsilon$$

In particular

$$M_i \leq m_i + \varepsilon.$$

But then

$$U(P,f) = \sum_i M_i \Delta x_i \leq \sum_i m_i \Delta x_i + \sum_i \varepsilon \Delta x_i = L(P,f) + \varepsilon(b-a).$$

Since $\overline{\int_{[a,b]}} f(x)\,dx$ is an infinmum over all partitions, and $\underline{\int_{[a,b]}} f(x)\,dx$ is a supremum over all partitions we have

$$\overline{\int_{[a,b]}} f(x)\,dx \leq U(P,f) \leq L(P,f) + \varepsilon(b-a) \leq \underline{\int_{[a,b]}} f(x)\,dx + \varepsilon(b-a).$$

This establishes (14.1) and concludes the proof. ☐

Lemma 14.12 (Splitting domains)**.** *Let $f : [a,b] \to \mathbb{R}$ be bounded.*

(1) for any $c \in (a,b)$ we have

$$\underline{\int_{[a,b]}} f = \underline{\int_{[a,c]}} f + \underline{\int_{[c,b]}} f$$

(2) for any $c \in (a,b)$ we have

$$\overline{\int_{[a,b]}} f = \overline{\int_{[a,c]}} f + \overline{\int_{[c,b]}} f$$

(3) $f : [a,b] \to \mathbb{R}$ is Riemann integrable if and only if for any $c \in (a,b)$ we have that $f : [a,c] \to \mathbb{R}$ and $f : [c,b] \to \mathbb{R}$ are Riemann integrable. Moreover in that case we have

$$\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f.$$

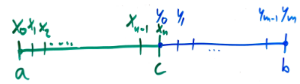

FIGURE 14.4. two partitions can be joined

Proof. (1) First we show

$$\underline{\int}_{[a,b]} f \geq \underline{\int}_{[a,c]} f + \underline{\int}_{[c,b]} f \tag{14.3}$$

Let $P_1 = \{x_0, \dots, x_n\}$ be a partition of $[a, c]$ and $P_2 = \{y_0, \dots, y_m\}$ a partition of $[c, b]$. Observe that this implies that $x_n = y_0 = c$. Let

$$P := \{x_0, \dots, x_{n-1}, c, y_1, \dots, y_m\}$$

Cf. Figure 14.4. P is a partition of $[a, b]$. Since $\underline{\int}_{[a,b]}$ is a supremum we have

$$\underline{\int}_{[a,b]} f \geq L(P, f) = L(P_1, f) + L(P_2, f).$$

This holds for any partition P_1 of $[a, c]$ and any partition P_2 of $[c, b]$, so we have

$$\underline{\int}_{[a,b]} f \geq \sup_{P_1} L(P_1, f) + \sup_{P_2} L(P_2, f) = \underline{\int}_{[a,c]} f + \underline{\int}_{[c,b]} f.$$

That is (14.3) is established.

For the reverse let $\varepsilon > 0$ and pick a a partition $P = \{x_0, \dots, x_N\}$ of $[a, b]$ such that

$$\underline{\int}_{[a,b]} f \leq L(P, f) + \varepsilon.$$

Let $Q := P \cup \{c\}$. Then Q is a refinement of P, and by Lemma 14.4 $L(Q, f) \geq L(P, f)$.

$$\underline{\int}_{[a,b]} f \leq L(Q, f) + \varepsilon.$$

However, now we may split $Q = Q_1 \cup Q_2$ with Q_1 a partition of $[a, c]$ and Q_2 a partition of $[c, d]$. By the definition of $L(Q, f)$ and $\underline{\int}$ we have

$$L(Q, f) = L(Q_1, f) + L(Q_2, f) \leq \underline{\int}_{[a,c]} f + \underline{\int}_{[c,b]} f.$$

So we arrive at

$$\underline{\int}_{[a,b]} f \leq \underline{\int}_{[a,c]} f + \underline{\int}_{[c,b]} f + \varepsilon.$$

This holds for any $\varepsilon > 0$, letting $\varepsilon \to 0$ we obtain

$$\underline{\int}_{[a,b]} f \leq \underline{\int}_{[a,c]} f + \underline{\int}_{[c,b]} f.$$

Since by (14.3) we have the converse inequality we conclude that

$$\underline{\int}_{[a,b]} f = \underline{\int}_{[a,c]} f + \underline{\int}_{[c,b]} f.$$

(2) Analogous to the above (exercise)

(3) If $f : [a,c] \to \mathbb{R}$ is Riemann integrable and $f : [c,b] \to \mathbb{R}$ is Riemann integrable we have by the above

$$\begin{aligned}\underline{\int}_{[a,b]} f =& \underline{\int}_{[a,c]} f + \underline{\int}_{[c,b]} f \\ &= \int_{[a,c]} f + \int_{[c,b]} f \\ &= \overline{\int}_{[a,c]} f + \overline{\int}_{[c,b]} f \\ = \overline{\int}_{[a,b]} f\end{aligned}$$

So $f : [a,b] \to \mathbb{R}$ is Riemann integrable and we can split the integral as claimed.

For the converse assume that we know $f : [a,b] \to \mathbb{R}$ is Riemann integrable. Then with the above argument

$$\begin{aligned}&\underline{\int}_{[a,c]} f + \underline{\int}_{[c,b]} f \\ =&\underline{\int}_{[a,b]} f \\ =&\overline{\int}_{[a,b]} f \\ =&\overline{\int}_{[a,c]} f + \overline{\int}_{[c,b]} f\end{aligned}$$

This implies

$$\underline{\int}_{[a,c]} f - \overline{\int}_{[a,c]} f = \overline{\int}_{[c,b]} f - \underline{\int}_{[c,b]} f$$

but from Lemma 14.5 we know then

$$0 \overset{\textit{Lemma } 14.5}{\geq} \underline{\int}_{[a,c]} f - \overline{\int}_{[a,c]} f = \overline{\int}_{[c,b]} f - \underline{\int}_{[c,b]} f \overset{\textit{Lemma } 14.5}{\geq} 0.$$

This implies

$$\underline{\int}_{[a,c]} f - \overline{\int}_{[a,c]} f = \overline{\int}_{[c,b]} f - \underline{\int}_{[c,b]} f = 0,$$

which implies that $f : [a,c] \to \mathbb{R}$ and $f : [c,b] \to \mathbb{R}$ are both Riemann integrable.

□

Lemma 14.13 (Linearity of the integral)**.** *Let $f, g : [a,b] \to \mathbb{R}$ be Riemann integrable and $\lambda, \mu \in \mathbb{R}$. Then $\lambda f + \mu g$ is Riemann integrable and we have*

$$\int_{[a,b]} (\lambda f + \mu g) = \lambda \int_{[a,b]} f + \mu \int_{[a,b]} g$$

Exercise 14.14. *Prove Lemma 14.13.*

Lemma 14.15. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function*[25] *such that for any c, d with $a < c < d < b$ we have that $f : [c,d] \to \mathbb{R}$ is Riemann integrable. Then $f : [a,b] \to \mathbb{R}$ is Riemann integrable and we have for any sequence $a < a_n < b_n < b$ with $\lim_{n\to\infty} a_n = a$ and $\lim_{n\to\infty} b_n = b$ that*

$$\int_{[a,b]} f = \lim_{n\to\infty} \int_{[a_n,b_n]} f.$$

Proof. We have by Lemma 14.12 (where in the second equality we use that $\underline{\int_{[a_n,b_n]}} f = \overline{\int_{[a_n,b_n]}} f$ by assumption, since f is integrable on $[a_n, b_n]$)

$$\begin{aligned}
\underline{\int_{[a,b]}} f &= \underline{\int_{[a,a_n]}} f + \underline{\int_{[b_n,b]}} f + \underline{\int_{[a_n,b_n]}} f \\
&= \underline{\int_{[a,a_n]}} f + \underline{\int_{[b_n,b]}} f + \overline{\int_{[a_n,b_n]}} f \\
&= \underline{\int_{[a,a_n]}} f + \underline{\int_{[b_n,b]}} f - \overline{\int_{[a,a_n]}} f - \overline{\int_{[b_n,b]}} f + \overline{\int_{[a,b]}} f
\end{aligned} \tag{14.4}$$

Thus,

$$\left| \underline{\int_{[a,b]}} f - \overline{\int_{[a,b]}} f \right| \leq 2\left(|a - a_n| + |b - b_n|\right) \sup_{[a,b]} |f|.$$

Since $a_n \to a$ and $b_n \to b$, and since $\sup_{[a,b]} |f| < \infty$, we let $n \to \infty$ to find that

$$\left| \underline{\int_{[a,b]}} f - \overline{\int_{[a,b]}} f \right| = 0.$$

That is f is Riemann integrable.

From (14.4) we also obtain that

$$\left| \int_{[a,b]} f - \int_{[a_n,b_n]} f \right| < \left(|a_n - a| + |b_n - b|\right) \sup_{[a,b]} |f|$$

Letting $n \to \infty$ we find that

$$\lim_{n\to\infty} \int_{[a_n,b_n]} f = \int_{[a,b]} f.$$

□

[25] so this lemma doesn't work for $\int_0^1 \frac{1}{x}$

A very annoying property of the integral (Riemann, but also the later Lebesgue integral) is the following. Assume that $f \leq g \leq h$, and f and h are Riemann integrable:

Exercise 14.16. *Find a function $g : [0,1] \to \mathbb{R}$ such that $0 \leq g(x) \leq 1$ for every $x \in [0,1]$, but so that g is not integrable.*

Hint: take the Dirichlet function from Example 14.9.

Example 14.17. (1) So we have seen that any continuous function $f : [a,b] \to \mathbb{R}$ is Riemann integrable. However there are discontinuous functions that are Riemann integrable. For example let

$$f(x) := \begin{cases} 1 & x \in [0,1] \\ 2 & x \in (1,2] \end{cases}$$

Then f is Riemann integrable. Indeed

f is clearly bounded, and since it is constant $\int_{[0,1]} f dx$ exists, i.e. f is integrable on $[0,1]$. f is also integrable on $[1,2]$, using Lemma 14.15: f is constant on $[1+\frac{1}{n}, 2]$, thus integrable, and thus also integrable on $[1,2]$. By Lemma 14.12 f is thus integrable on $[0,2]$.

(2) Crazy functions might be integrable. Let

$$f(x) := \begin{cases} \sin(1/x) & x \in (0,1] \\ 0 & x = 0. \end{cases}$$

Then f is Riemann integrable.

Indeed, f is continuous on $[0+\frac{1}{n}, 1]$ for any $n \in \mathbb{N}$. So f is integrable on $[0+\frac{1}{n}, 1]$ by Proposition 14.11. Since $f : [0,1] \to \mathbb{R}$ is bounded (not continuous) we find that f is integrable, by Lemma 14.15.

Exercise 14.18. *Set*

$$f(x) := \begin{cases} \arctan(1/x) & x \neq 0 \\ 25 & x = 0 \end{cases}$$

Show that f is Riemann-integrable on $[-1,1]$.

(Do not use Proposition 14.19)

We conclude: *continuity is not necessary for integrability*.

Indeed, from Lemma 14.15 we easily obtain

Proposition 14.19. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function, such that for a finite set $\Sigma = \{c_1, \ldots, c_N\}$ we have f is continuous in $[a,b] \backslash \Sigma$. Then f is Riemann integrable.*

Proof. We may assume that $a \leq c_1 < c_2 < \ldots < c_N \leq b$.

We divide $[a,b]$ into the intervals $[a,c_1]$, $[c_1,c_2]$, … ,$[c_N,b]$.

We first observe that f is integrable on each of these subintervalls. Indeed, f is integrable on

$$[a+\frac{1}{n}, c_1-\frac{1}{n}], \quad [c_1+\frac{1}{n}, c_2-\frac{1}{n}], \ldots, [c_N+\frac{1}{n}, b-\frac{1}{n}],$$

because it is continuous and we have Proposition 14.11. Since f is moreover bounded by assumption, Lemma 14.15 implies that f is integrable on each of the intervals $[a, c_1]$, $[c_1, c_2]$, ... ,$[c_N, b]$.

By Lemma 14.12 f is continuous on $[a, b]$. □

Exercise 14.20. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function, such that for a countable set $\Sigma = \{c_1, \ldots\}$ we have f is continuous in $[a, b]\backslash\Sigma$. Without using Riemann-Lebesgue theorem, Theorem 14.25 below, show that f is Riemann integrable.*

Exercise 14.21. *The Dirichlet function*

$$D(x) := \begin{cases} 1 & x \in [0,1] \cap \mathbb{Q} \\ 0 & x \in [0,1] \setminus \mathbb{Q} \end{cases}$$

is not Riemann-integrable, see Example 14.9.

Why is this no contradiction to Exercise 14.20?

The strongest result for integrability is the Riemann-Lebesgue Theorem, Theorem 14.25 below. It states that bounded functions are Riemann integrable, if and only if they are continuous *almost everywhere*. This means that $f : [a, b]\backslash\Sigma$ is continuous outside of a set Σ which has *zero measure*. There is a larger theory behind this, the theory of measures.

Definition 14.22. A set $\Sigma \subset \mathbb{R}$ has (Lebesgue-)*measure zero* if for all $\varepsilon > 0$ there is a countable collection of open intervals $\{I_1, I_2, \ldots\}$ such that

$$\Sigma \subset \bigcup_{i=1}^{\infty} I_i$$

and $\sum_{i=1}^{\infty} \mu(I_i) < \varepsilon$. Here $\mu : A \subset \mathbb{R} \to \mathbb{R}_+$ is the Lebesgue measure which for intervals is simply

$$\mu((a, b)) = \mu([a, b)) = \mu((a, b]) = \mu([a, b]) = b - a.$$

That is in other words: if the intervals $I_i = [a_i, b_i]$ then we need to have

$$\sum_{i=1}^{\infty} |a_i - b_i| < \varepsilon.$$

Example 14.23.

- It is easy to show that if $\Sigma = \{x_1, \ldots, x_n, \ldots\}$ (i.e. a countable set), then Σ has Lebesgue-measure zero. Indeed, for any given $\varepsilon > 0$ let

  $$r_i := 2^{-i-2}\varepsilon,$$

  and set

  $$I_i := (x_i - r_i, x_i + r_i).$$

FIGURE 14.5. The Construction of the Cantor set: Take a the line $C_1 := [0,1]$ (first line in the image). Now split this interval into three equal parts and remove the middle part (second line in the image), call this C_2. Take the intervals of C_2, split them into thirds, remove the middle part and obtain C_3. etc. Doing this infinitely many times, i.e. considering $C := \bigcap_i C_i$ is the cantor set. It is uncountable, has *Hausdorff dimension* $\ln 2/\ln 3 < 1$, and it has Lebesgue measure zero (exercise!).

Then $\Sigma \subset \bigcup_{i=1}^N I_i$ and we have

$$\sum_{i=1}^{\infty} \mu(I_i) = \sum_{i=1}^{\infty} 2^{-i-1}\varepsilon = \frac{1}{2}\varepsilon < \varepsilon.$$

- What is curious is that there are uncountable sets Σ that still have Lebesgue-measure zero. E.g. the *Cantor set*, see Figure 14.5

Exercise 14.24. *Let C be the* Cantor set *which is defined as follows:*

Take $C_0 = [0,1]$.

Then C_n is defined iteratively by taking out the open middle third from any segment, i.e. C_1 deletes the middle third line segment of C_0, that gives:

$$C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1].$$

C_2 now deletes the middle third of the two line segments from C_1, that is

$$C_2 := [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1].$$

And so on, See Figure 14.5.

We could also write this[26]

$$C_n := \frac{C_{n-1}}{3} \cup \left(\frac{2}{3} + \frac{C_{n-1}}{3}\right)$$

Then

$$C := \bigcap_{n=0}^{\infty} C_n$$

show that

(1) *the Cantor set C has Lebesgue measure zero, in the sense of Definition 14.22*

[26]Here: for a set $A \subset \mathbb{R}$, a point $x \in \mathbb{R}$ and $\lambda \in \mathbb{R}$ we write

$$x + \lambda A = \{z \in \mathbb{R} : z \text{ can be written as } z = x + \lambda a \text{ for some } a \in A\}.$$

FIGURE 14.6. Henri Lebesgue, 1875-1841. French, known for the Lebesgue Integral.

FIGURE 14.7. Felix Hausdorff, 1868-1942. German, one of the founders of modern topology.

(2) (difficult) the Cantor set is uncountable (see wikipedia)

Theorem 14.25 (Riemann Lebesgue Theorem)**.** *A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann-integrable if and only if there exists some $\Sigma \subset \mathbb{R}$ with Lebesgue measure zero such that $f : [a, b] \to \mathbb{R}$ is continuous at every point $x \in [a, b] \backslash \Sigma$.*

Exercise 14.26. *Let C the Cantor set.*

Define $f : [0, 1] \to \mathbb{R}$ as

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

Show that f is continuous in $[0,1]\backslash C$ (hint: C is closed), and thus f is Riemann integrable.

Exercise 14.27. *We know that countable sets Σ have zero Lebesgue measure, cf. Example 14.23.*

- *Consider the Dirichlet-function from Example 6.10. We know that it is not integrable, Example 14.9. Why does this not contradict Theorem 14.25?*
- *Consider the Thomae-function (popcorn function) from Example 6.10. Show that it is integrable in $[0,1]$.*

Corollary 14.28. *Let $f:[a,b]\to\mathbb{R}$ be Riemann-integrable function. Let $g:D\to\mathbb{R}$ be a continuous map on D such that $f([a,b])\subset D$. Then if $g\circ f:[a,b]\to\mathbb{R}$ is bounded, then $g\circ f:[a,b]\to\mathbb{R}$ is Riemann integrable.*

In particular, if $f:[a,b]\to\mathbb{R}$ is Riemann-integrable then $|f|:[a,b]\to\mathbb{R}$ is Riemann integrable, and we have

$$|\int_{[a,b]} f| \leq \int_{[a,b]} |f|.$$

Proof. Observe that f is continuous at $c\in[a,b]$ then $g\circ f$ is continuous at $c\in[a,b]$. That is, if Σ is the set of discontinuities of f, the set of discontinuities of $g\circ f$ is equal or smaller. In particular since f is Riemann integrable and $g\circ f$ is bounded then by Theorem 14.25 $g\circ f$ is still Riemann integrable.

Since $g(x):=|x|$ is uniformly continuous and takes bounded sets into bounded sets, we have that $|f|$ is Riemann integrable if f is[27]. Since moreover $f(x)\leq|f(x)|$ and $-f(x)\leq|f(x)|$ for all $x\in[a,b]$ we have

$$\int_{[a,b]} f(x)\,dx \leq \int_{[a,b]} |f(x)|\,dx$$

and

$$-\int_{[a,b]} f(x)\,dx \leq \int_{[a,b]} |f(x)|\,dx$$

This implies

$$\left|\int_{[a,b]} f(x)\,dx\right| \leq \int_{[a,b]} |f(x)|\,dx$$

□

A curious property is that we can change functions in very tiny sets, without changing the integral.

Theorem 14.29. *Let $f,g:[a,b]\to\mathbb{R}$ be Riemann-integrable such that $f(x)=g(x)$ for all $x\in[a,b]\backslash\Sigma$ where for some $\Sigma\subset\mathbb{R}$ with Lebesgue measure zero[28]. Then*

$$\int_{[a,b]} f = \int_{[a,b]} g.$$

[27] observe the converse is false, take the Dirichlet function with $+1$ and -1!

[28] We call this: $f=g$ *almost everywhere* in $[a,b]$

Remark 14.30. Theorem 14.29 is false, if g is not assumed to be Riemann integrable. E.g. $f(x) = 0$, and g the Dirichlet function

$$g(x) := \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

Then $\Sigma := \{x \in \mathbb{R}, f(x) \neq g(x)\} = \mathbb{Q}$ is a zero set, however f is integrable and g is not.

Exercise 14.31. *Show Theorem 14.29 for Σ a finite set.*

Exercise 14.32. [Leb, ex. 5.2.4] *Prove the mean value theorem for integrals. That is, prove that if $f : [a, b] \to \mathbb{R}$ is continuous, then there exists a $c \in [a, b]$ such that $\int_{[a.b]} f = f(c)(b - a)$.*

Hint: Use the min-max theorem and the intermediate value theorem.

Exercise 14.33. [Leb, ex. 5.2.6] *Suppose $f : [a, b] \to \mathbb{R}$ is a continuous function and $\int_{[a,b]} f = 0$. Prove that there exists a $c \in [a, b]$ such that $f(c) = 0$.*

Exercise 14.34. *Let $f : [a, b] \to \mathbb{R}$ be continuous and is such that*

$$\int_{[a,b]} f\eta = 0 \quad \text{for any continuous function } \eta : [a, b] \to \mathbb{R}.$$

Show that $f \equiv 0$.

Hint: Prove (picture proof is fine) that for any $c_1 < c_2 < d_1 < d_2$ there exists a bump function, cf. Example 13.2, with the following properties:

- *η is Lipschitz continuous on $\mathbb{R}$*
- *$\eta \geq 0$ in $\mathbb{R}$*
- *$\eta \equiv 0$ in $\mathbb{R} \setminus [c_1, d_2]$*
- *$\eta \equiv 1$ in $[c_2, d_2]$*

Then argue by contradiction, assume that there exists $x_0 \in [a, b]$ with (say) $f(x_0) > 0$ and choose η wisely.

The above property is called the fundamental theorem of calculus of variations, not to be confused with the fundamental theorem of calculus in the next section, Section 15.

Exercise. *Let $(r_n)_{n\in\mathbb{N}}$ be an enumeration of $\mathbb{Q}$. I.e. assume that*

$$\mathbb{Q} = \{r_1, r_2, r_3, \dots\}.$$

Set for $x \in \mathbb{R}$

$$f(x) := \sum_{\{n\in\mathbb{N}: r_n \leq x\}} 2^{-n}.$$

(1) Show that for each $x \in \mathbb{R}$, $f(x) < \infty$, i.e. f is well-defined.
(2) Show that f is discontinuous for any rational $x \in \mathbb{R} \setminus \mathbb{Q}$.
(3) Show that f is continuous for any irrational $x \in \mathbb{R} \setminus \mathbb{Q}$.

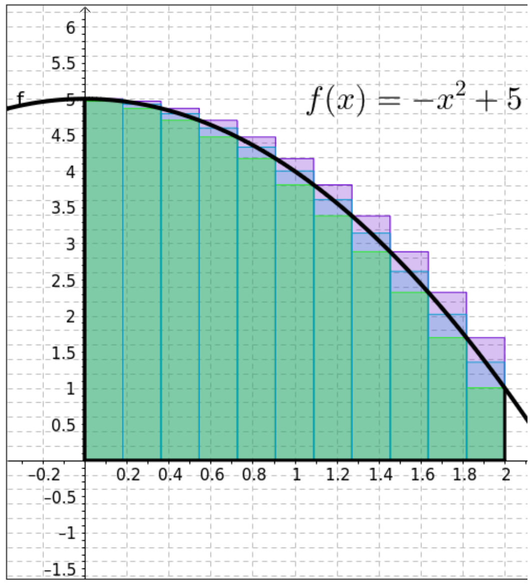

FIGURE 14.8. The area below a curve approximated via a Riemann sum, cf. Definition 14.35, with $y_i = x_{i-1}$ (green), y_i midpoint between x_{i-1} and x_i (blue), and $y_i = x_i$ (purple). (image and applet: https://www.geogebra.org/m/yGNEAgcw)

(4) Show that $f\Big|_{[0,1]}$ *is an integrable function.*

14.1. **Riemann sums.**

Definition 14.35. let $f : [a, b] \to \mathbb{R}$ be any function. Let $n \in \mathbb{N}$ and set $P_n = \{x_0, \dots, x_n\}$ be the equidistributed partition of $[a, b]$, namely

$$x_i := a + i\frac{b-a}{n}.$$

Take furthermore a collection $(y_i)_{i=1}^n \subset [a, b]$ such that $y_i \in [x_{i-1}, x_i]$. An n-th *Riemann sum* of f over $[a, b]$ with points $(y_i)_{i=1}^n$ is defined as

$$S_n := \sum_{i=1}^{n} f(y_i)\Delta x_i.$$

Proposition 14.36 (Riemann Sums converge)**.** *Let* $f : [a, b] \to \mathbb{R}$ *be continuous, and set* S_n *the* n*-the Riemann sum as above. Then* $\lim_{n\to\infty} S_n = \int_{[a,b]} f dx$*.*

Proof. Since f is continuous on $[a, b]$, f is integrable.

Fix $\varepsilon > 0$. And let $M := \sup_{[a,b]} |f| < \infty$.

Let

$$P_n := \{x_i \quad i \in \{0, \ldots, n\}\},$$

where we recall $x_i := a + i\frac{b-a}{n}$. This is clearly a partition of $[a, b]$.

Fix $\varepsilon > 0$. Since f is continuous on $[a, b]$ it is uniformly continuous. Thus there exists for any $\delta > 0$ some $\varepsilon > 0$ such that

$$|f(x) - f(y)| < \varepsilon \quad \forall |x - y| < \delta.$$

So whenever $n > \frac{1}{\delta}$, since $y_i \in [x_{i-1}, x_i]$ and $|x_{i-1} - x_i| = \frac{1}{n} < \delta$ we have

$$|f(y_i) - f(z)| < \varepsilon \quad \forall z \in [x_{i-1}, x_i].$$

In particular

$$\left| f(y_i) - \inf_{[x_{i-1}, x_i]} f \right| < \varepsilon \quad \text{and} \quad \left| f(y_i) - \sup_{[x_{i-1}, x_i]} f \right| < \varepsilon.$$

Since we can write

$$S_n = \sum_{i=1}^n f(y_i)\Delta x_i = \underbrace{\sum_{i=1}^n \inf_{[x_{i-1}, x_i]} f \ \Delta x_i}_{L(P_n, f)} + \sum_{i=1}^n \underbrace{\left(f(y_i) - \inf_{[x_{i-1}, x_i]} f(z)\right)}_{|\cdot| \leq \varepsilon} \Delta x_i$$

and

$$S_n = \sum_{i=1}^n f(y_i)\Delta x_i = \underbrace{\sum_{i=1}^n \sup_{[x_{i-1}, x_i]} f \ \Delta x_i}_{U(P_n, f)} + \sum_{i=1}^n \underbrace{\left(f(y_i) - \sup_{[x_{i-1}, x_i]} f(z)\right)}_{|\cdot| \leq \varepsilon} \Delta x_i$$

we then find

$$U(P_n, f) - \varepsilon|b - a| \leq S_n \leq L(P_n, f) + \varepsilon|b - a|$$

By the definition of upper and lower integral we thus have

$$\underline{\int_{[a,b]}} f - \varepsilon|b - a| \leq S_n \leq \overline{\int_{[a,b]}} f + \varepsilon|b - a|$$

Since f is integrable, Proposition 14.11, we have $\underline{\int_{[a,b]}} f = \overline{\int_{[a,b]}} f = \int_{[a,b]} f$, so the above inequality is actually

$$\int_{[a,b]} f - \varepsilon|b - a| \leq S_n \leq \int_{[a,b]} f + \varepsilon|b - a|,$$

or

$$\left| S_n - \int_{[a,b]} f \right| \leq \varepsilon|b - a|, \quad \forall n > \frac{1}{\delta}.$$

That is we have shown (for $N := \frac{1}{\delta}$)

$$\forall \varepsilon > 0 \exists N : \left| S_n - \int_{[a,b]} f \right| \leq \varepsilon|b - a| \quad \forall n \geq N.$$

This is the same as

$$\lim_{n \to \infty} S_n = \int_{[a,b]} f.$$

□

Observe that the Riemann sum converging does not mean integrability! (that's why we assumed continuity in the statment of Proposition 14.36). Indeed:

Example 14.37. Let for $x \in [0,1]$,

$$f(x) := \begin{cases} 0 & \frac{1}{x} \notin \mathbb{N} \\ k & x = \frac{1}{k}, \text{for some } k \in \mathbb{N}. \end{cases}$$

Clearly f is not bounded (and as we have seen in Lemma 14.10). But if we choose y_i such that $\frac{1}{y_i} \notin \mathbb{N}$, then the Riemann sum is zero (and thus convergent).

Example 14.38. Using Riemann sums, prove that

$$\lim_{n\to\infty} \left(\frac{1}{n} + \frac{1}{n+1} + \ldots + \frac{1}{2n} \right) = \ln 2.$$

Proof. Denote the sequence whose limit we must find by a_n and let $f(x) = (1+x)^{-1}$. Then $a_n = \frac{1}{n} + b_n$, where

$$b_n = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{n+n} = \frac{1}{n} \left(\frac{1}{1+\frac{1}{n}} + \frac{1}{1+\frac{2}{n}} + \cdots + \frac{1}{1+\frac{n}{n}} \right) = \frac{1}{n} \sum_{i=1}^{n} f\left(\frac{i}{n}\right).$$

We observe that this is of the form of a Riemann sum:

$$b_n = \sum_{i=1}^{n} f(x_i)\, \Delta x_i$$

where $x_i = \frac{i}{n}$ and $\Delta x_i = \frac{1}{n}$.

Therefore b_n is a Riemann sum approximating the integral $\int_{[0,1]} (1+x)^{-1}\, dx$ – by Proposition 14.36. So

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{1}{n} + \lim_{n\to\infty} b_n = 0 + \int_{[0,1]} \frac{dx}{1+x} = \ln(1+x)\Big|_0^1 = \ln 2.$$

□

Exercise 14.39. *Find*

$$\lim_{n\to\infty} \frac{\sqrt{1}+\sqrt{2}+\sqrt{3}+\sqrt{4}+\ldots+\sqrt{n}}{n\sqrt{n}}$$

Hint: Use Riemann sums

15. Fundamental Theorem of Calculus

We recall from Calculus, the *Fundamental Theorem of Calculus*:

$$\int_a^b f'(x)dx = f(b) - f(a).$$

We have so far talked only about the integral $\int_{[a,b]} g$, but not about $\int_a^b g$. Its almost the same, but if $b < a$ it gets a minus sign:

Definition 15.1 (The Calculus integral)**.** Let $f : [a,b] \to \mathbb{R}$ be integrable. Then

$$\int_a^b f(x)dx := \int_{[a,b]} f(x)\, dx$$

$$\int_b^a f(x)dx := -\int_{[a,b]} f(x)\, dx.$$

For simplicity we also set

$$\int_a^a f(x)dx = 0.$$

Here is the *fundamental theorem of calculus*.

Theorem 15.2. *Let $f : [a,b] \to \mathbb{R}$ be a continuous and differentiable function such that $f'(x)$ is Riemann integrable on $[a,b]$. Then*

$$f(y) - f(x) = \int_x^y f'(z)\, dz.$$

Proof. Assume w.l.o.g. $y > x$.

Let $P := \{z_0, \ldots, z_n\}$ be a partition of $[x,y]$. In the interval $[z_{i-1}, z_i]$ we can use the mean value theorem, Theorem 12.2, and find $c_i \in (z_{i-1}, z_i)$ such that

$$f(z_i) - f(z_{i-1}) = f'(c_i)\Delta z_i.$$

That is,

$$\inf_{[z_{i-1},z_i]} f'\ \Delta z_i \le f(z_i) - f(z_{i-1}) \le \sup_{[z_{i-1},z_i]} f'\ \Delta z_i$$

Summing over $i = 1, \ldots, n$ (observe the telescoping for the middle part) we find

$$L(f', P) \le f(y) - f(x) \le U(f', P)$$

This holds for any partition P of $[a,b]$ so we find

$$\underline{\int_{[x,y]}} f'(z)dz \le f(y) - f(x) \le \overline{\int_{[x,y]}} f'(z)dz.$$

Since f' is integrable in $[a,b]$ it is also integrable in $[x,y]$. This implies

$$f(y) - f(x) = \int_{[x,y]} f'(z)dz = \underline{\int_{[x,y]}} f'(z)dz = \overline{\int_{[x,y]}} f'(z)dz.$$

□

Another version is the following statement

Theorem 15.3 (The fundamental theorem of Calculus)**.** *Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable. Then $F : [a,b] \to \mathbb{R}$ defined as*

$$F(x) := \int_{[a,x]} f(z)\, dz \equiv \int_a^x f(z)\, dz$$

is a Lipschitz continuous function in $[a,b]$ with

$$|F(x) - F(y)| \le \sup_{[a,b]} |f|\, |x-y|.$$

If f is continuous[29] *at $c \in [a,b]$ then F is differentiable at c and $F'(c) = f(c)$.*

Proof. As for Lipschitz continuity observe that by Lemma 14.12 (also cf. Corollary 14.28)

$$|F(x) - F(y)| = |\int_{[x,y]} f| \le \int_{[x,y]} |f| \le \sup_{[a,b]} |f|\, |x-y|.$$

Next let $c < y$. Then

$$\begin{aligned} &F(y) - F(c) - f(c)(y-c) \\ =&\int_{[c,y]} f(z)dz - f(c)(y-c) \\ =&\int_{[c,y]} (f(z) - f(c))dz. \end{aligned}$$

Now assume that f is continuous at c. Then for any $\varepsilon > 0$ there exist $\delta > 0$ such that $|f(c) - f(z)| < \frac{\varepsilon}{2}$ for all z such that $|c-z| < \delta$. So if $y > c$ is such that $|c-y| < \delta$ we have

$$\left|\int_{[c,y]} (f(z) - f(c))dz\right| \le |c-y| \sup_{z:|z-c|<\delta} |f(z) - f(c)| \le \frac{\varepsilon}{2}|c-y|.$$

That is whenever $y > c$ such that $|c-y| < \delta$ we have

$$\frac{|F(y) - F(c) - f(c)(y-c)|}{|c-y|} \le \frac{\varepsilon}{2} < \varepsilon.$$

Doing the same same argument also for $y < c$ with $|c-y| < \delta$ we conclude that for any y with $|y-c| < \delta$ we have

$$\frac{|F(y) - F(c) - f(c)(y-c)|}{|c-y|} < \varepsilon.$$

In view of Lemma 10.2, this proves differentiability of F if f is continuous and $F'(c) = f(c)$. □

[29]This is important, think of f as the constnat 1 but in c we set $f(c) = 2$. The Riemann integral doesn't care and says $F(x) = \int_a^x f(z)dz = (x-a)$. So $F'(c) = 1 \neq 2$

One consequence

Proposition 15.4. *Let $f : [a,b] \to \mathbb{R}$ be continuous. Assume that for some $c \in (a,b)$ we have that f is differentiable in $[a,b]\backslash\{c\}$, that f' is continuous in $[a,b]\backslash\{c\}$, and that $L = \lim_{x\to c} f'(x)$ exists. Then f is differentiable in c and $f'(c) = L$.*

Proof. Set
$$g(z) := \begin{cases} f'(z) & z \neq c \\ L & z = c. \end{cases}$$
By assumption g is continuous and thus Riemann integrable in $[a,b]$, Proposition 14.11. Set
$$h(x) := f(a) + \int_{[a,x]} g(z)dz.$$
By the fundamental theorem of calculus, Theorem 15.3, h is differentiable in $[a,b]$, and $h'(x) = g(x)$ for all $x \in [a,b]$. Thus $(h(x) - f(x))' = 0$ in $[a,c)$ and $(c,b]$.

For any small $\varepsilon > 0$ we can apply Proposition 12.3 in $[a, c-\varepsilon]$ and $[c+\varepsilon, b]$ and we find that there exists $D_1, D_2 \in \mathbb{R}$ such that
$$h(x) - f(x) = \begin{cases} D_1 & x \in [a, c-\varepsilon] \\ D_2 & x \in [c+\varepsilon, b] \end{cases}$$
Since $h - f$ is on the other hand continuous, we can $\varepsilon \to 0$ and find $D_1 = D_2$. Since $h(a) = f(a)$ we have $D_1 = D_2 = 0$. So $h = f$ and thus f is differentiable everywhere. □

Another application of the Fundamental theorem is the change of variables formula

Proposition 15.5. *Let $g : [a,b] \to \mathbb{R}$ be continuously differentiable*[30] *function and $f : [c,d] \to \mathbb{R}$ continuous. If $g([a,b]) \subset [c,d]$ then*
$$\int_a^b f(g(x))g'(x)\, dx = \int_{g(a)}^{g(b)} f(s)\, ds.$$

Proof. Let
$$F(x) := \int_c^x f(z)\, dz.$$
Then by the fundamental theorem, Theorem 15.3, F is continuously differentiable in $[c,d]$. Thus, $F \circ g$ is continuously differentiable in $[a,b]$, and we have
$$(F \circ g)'(x) = F'(g(x))\, g'(x) = f(g(x))\, g'(x).$$
Again by the fundamental theorem,
$$\int_a^b (F \circ g)'(x)dx = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(z)\, dz.$$

[30] recall that this means: g is differentiable and g' is continuous

That is, we have shown

$$\int_a^b f(g(x))\, g'(x)\, dx = \int_{g(a)}^{g(b)} f(z)\, dz.$$

□

Exercise 15.6. *Use the product rule for derivatives,*

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

and the fundamental theorem of calculus to show the integration by parts formula:

Let $f, g : \mathbb{R} \to \mathbb{R}$ be continuous, differentiable and with continuous derivative.

Then for any $a < b$ we have

$$\int_a^b f(x)\, g'(x) dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)\, g(x) dx$$

Exercise 15.7. [Leb, Exercises 5.3.8] *Suppose that $f : [a,b] \to \mathbb{R}$ is continuous. Suppose that $\int_a^x f(z)dz = \int_x^b f(z)dz$ for all $x \in [a,b]$. Show that $f(x) = 0$ for all $x \in [a,b]$.*

Further results. Another consequence of the fundamental theorem of calculus is the (extremely important) *Sobolev-Poincaré inequality*, in one dimension also called *Wirtinger's inequality*. (It is more complicated to prove in higer dimensions).

Proposition 15.8 (Sobolev-Poincaré inequality)**.** *Let I be an interval and $f \in C^1(\overline{I})$ with either*

(1) $f(x_0) = 0$ for some $x_0 \in I$
(2) $\int_I f(x)\, dx = 0$
(3) $\int_J f(x) dx = 0$ for some nonempty subinterval J, $\emptyset \neq J \subset I$

and let $p, q \in [1, \infty)$. Then there exists a constant $C = C(p,q)$ (C depends on J in the third case above) such that

$$\|f\|_{L^p(I)} := \left(\int_I |f(x)|^p\right)^{\frac{1}{p}} \leq C \operatorname{diam}(I)^{1+\frac{1}{p}-\frac{1}{q}} \left(\int_I |f'(x)|^q\right)^{\frac{1}{q}}.$$

Here $\operatorname{diam}(I)$ denotes the diameter of the interval, i.e. if $I = [a,b]$ then $\operatorname{diam} I = |b-a|$.

Sobolev-Poincaré are an example of a quantitative estimate of the form "oscillation $|f'|$ controls magnitude $|f|$".

Proof of Proposition 15.8. Observe that since $f \in C^1(\overline{I})$, we have that $|f|$ and $|f'|$ are Riemann integrable, Corollary 14.28. So are $|f|^p$ and $|f'|^q$ via the Riemann-Lebesgue theorem, Theorem 14.25.

We prove only the case $q = 1$, for $q > 1$ we would need *Hölder's inequality*.

(1) Assume that $f(x_0) = 0$ for some $x_0 \in I$. Then, from the fundamental theorem of calculus for any $x \in I$
$$f(x) = f(x) - f(x_0) = \int_{x_0}^{x} f'(z)\,dz.$$
Consequently,
$$|f(x)| \le \int_{[x_0,x]} |f'(z)|\,dz.$$
Since x_0 and x belong to I, we have that $[x_0, x] \subset I$, so
$$|f(x)| \le \int_I |f'(z)|\,dz.$$
thus
$$|f(x)|^p \le \left(\int_I |f'(z)|\,dz\right)^p =: \Lambda.$$
Observe that the right-hand side is a constant, and the left-hand site is an integrable function. So
$$\int_I |f(x)|^p dx \le \mathrm{diam}\,(I)\Lambda = \mathrm{diam}\,(I)\left(\int_I |f'(z)|\,dz\right)^p.$$
That is,
$$\left(\int_I |f(x)|^p\,dx\right)^{\frac{1}{p}} \le \mathrm{diam}\,(I)^{\frac{1}{p}} \int_I |f'(z)|\,dz.$$

(2) Only the first part changes in the other cases: If $\int_J f(x)dx = 0$, for some nonempty $J \subset I$ then
$$f(x) = f(x) - \frac{1}{\mathrm{diam}\,(J)}\int_J f(y)\,dy = \frac{1}{\mathrm{diam}\,(J)}\int_J (f(x) - f(y))\,dy$$
Arguing as above we find
$$|f(x)| \le \frac{1}{\mathrm{diam}\,(J)}\int_J \int_I |f'(z)|\,dz\,dy = \int_I |f'(z)|\,dz$$
The remaining argument stays the same.

□

16. Sequences of functions: Pointwise and uniform convergence

Videolink

Pointwise vs. Uniform Convergence. A relatively short (23:59) video on the topic of pointwise vs. uniform convergence of a sequence of function.
https://www.youtube.com/watch?v=McKuQEXXzH0

Videolink

An example of proving that a sequence of functions converge pointwise (13:10) https://youtu.be/gqLqkYwyq5Q

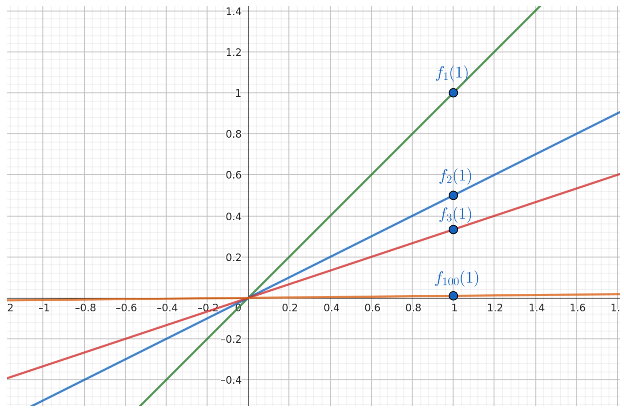

FIGURE 16.1. The sequence of functions $f_n(x) = \frac{x}{n}$, more precisely f_1 (green), f_2 (blue), f_3 (red), f_{100} (brown). If we evaluate the function sequence at any point x (here $x = 1$) we see that the function values $f_n(x)$ tend to zero as $n \to \infty$

Videolink

An example of showing that a sequence of functions converges uniformly (9:13) https://youtu.be/i_TtG4NqzBM

We have treated sequences $(x_n)_{n\in\mathbb{N}}$ and functions $f : D \to \mathbb{R}$. Now we want to treat sequences of functions, i.e. $(f_n)_{n\in\mathbb{N}}$ where $f_n : D \to \mathbb{R}$ are functions (with the same domain).

We would like to talk about *convergence* of such sequences, but what does convergence mean? Here is a first attempt, for each $x \in D$ we look at $(f_n(x))_{n\in\mathbb{N}}$ (which is a sequence in $\mathbb{R}$), and say that f_n converges if $f_n(x)$ converges for any $x \in D$. This is *pointwise convergence*, and it is a "weak convergence" (as we shall see below).

Definition 16.1 (pointwise convergence)**.** Let $(f_n)_{n\in\mathbb{N}}$, f all functions $D \to \mathbb{R}$. We say that f_n *converges pointwise* to f in D if

$$f_n(x) \xrightarrow{n\to\infty} f(x) \quad \forall x \in D.$$

Example 16.2. (1) Let $f_n(x) = 1$ be the constant function in $\mathbb{R}$. Then $\lim_{n\to\infty} f_n(x) = 1$.

(2) $f_n(x) := \frac{x}{n}$. Then $\lim_{n\to\infty} f(x_n) = 0$, cf. Figure 16.1.

(3) $f_n(x) := \sin^n(x)$ on $[0, \pi]$. Then for $x = \frac{\pi}{2}$ we have $f_n(\frac{\pi}{2}) = 1 \xrightarrow{n\to\infty} 1$. Since $|\sin(x)| < 1$ for $x \in [0, \pi]\backslash\{\pi/2\}$ we have $\lim_{n\to\infty} f_n(x) = 0$ for all such x. That is for

$$f(x) := \begin{cases} 1 & x = \frac{\pi}{2} \\ 0 & x \in [0, \pi]\backslash\{\frac{\pi}{2}\} \end{cases}$$

we have $\lim_{n\to\infty} f_n(x) = f(x)$. That is: pointwise limit of a continuous function (even smooth!) function may not be continuous, cf. Figure 16.2.

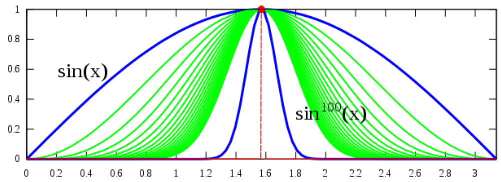

FIGURE 16.2. The sequence of functions $f_n(x) := \sin^n(x)$ on $[0, \pi]$ converges pointwise to a discontinuous function which is 1 at $\frac{\pi}{2}$ and zero everywhere else. (image: JoKalliauer/Wikipedia)

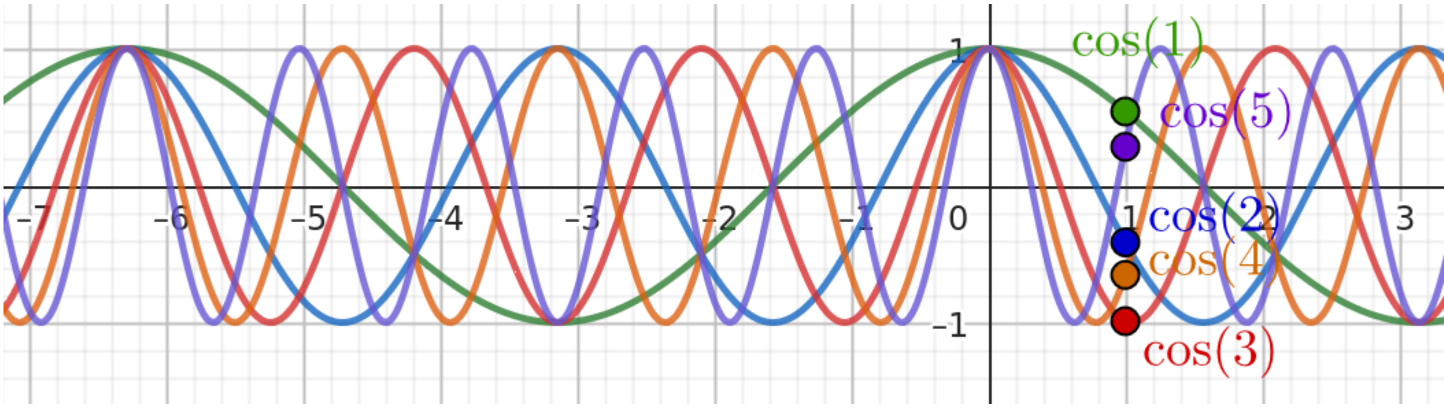

FIGURE 16.3. The sequence of functions $h_n(x) = \cos(nx)$, more precisely h_1 (green), h_2 (blue), h_3 (red), h_4 (brown), h_5 (purple). If we evaluate the function sequence at any point $x \neq 0$ (here $x = 1$) we see that the function values $h_n(x)$ don't converge as $n \to \infty$

(4) Let $f_n(x) := x^{2n}$ for $x \in [-1, 1]$. Then $\lim_{n\to\infty} f_n(x) = f(x)$.

$$f(x) := \begin{cases} 1 & x = \pm 1 \\ 0 & x \in (-1, 1) \end{cases}$$

(5) Let $f_n(x) := \frac{\sin(nx)}{n}$ in $\mathbb{R}$. Then by the squeeze theorem $\lim_{n\to\infty} f_n(x) = f(x) := 0$. Observe that f and f_n are all differentiable. However,

$$f_n'(x) = \cos(nx).$$

For $x \neq 0$ this sequence of function $\cos(nx)$ has no pointwise limit, cf. Figure 16.3. In particular $\lim_{n\to\infty} f_n'(x) \neq f'(x)$. That is: just because the convergence is pointwise, even if the functions converges to a differentiable function, the derivatives may not.

(6) Let $(q_n)_{n\in\mathbb{N}}$ be an enumeration of $\mathbb{Q} \cap [0, 1]$. Set

$$f_n(x) := \begin{cases} 1 & x = q_1, \ldots, q_n \\ 0 & \text{otherwise.} \end{cases} \tag{16.1}$$

Observe that $f_n \in \mathcal{R}[a,b]$ by Riemann-Lebesgue theorem, Theorem 14.25. But $\lim_{n\to\infty} f_n(x) = D(x) \notin \mathcal{R}[a,b]$ (The Dirichlet function from Example 14.9).

So continuity, differentiability, Riemann-integrability are not *stable* under pointwise convergence.

Later (in *Functional Analysis*) each *function space* C^0, C^1, $\mathcal{R}[a,b]$ (bit more difficult), gets associated with a *metric* in form of *norm*, which makes each space a so-called *Banach space*, so that convergence with respect to that respective metric is stable.

Here we introduce the L^∞-norm[31] which makes the set of bounded functions in D (denoted by $L^\infty(D)$) a Banach space (later: in Functional Analysis).

$$\|f\|_{L^\infty(D)} := \sup_{x\in D} |f(x)|. \tag{16.2}$$

Lemma 16.3. *$\|\cdot\|_{L^\infty(D)}$ defines a* norm *on the set $L^\infty(D)$. Namely for any $f,g \in L^\infty(D)$ and $\lambda \in \mathbb{R}$ we have*

- $\|f+g\|_{L^\infty(D)} \leq \|f\|_{L^\infty(D)} + \|g\|_{L^\infty(D)}$ *(triangle inequality)*
- $\|\lambda f\|_{L^\infty(D)} = |\lambda|\, \|f\|_{L^\infty(D)}$ *(positive homogeneity)*
- $\|f\|_{L^\infty(D)} \geq 0$. *Moreover, if* $\|f\|_{L^\infty(D)} = 0$ *then* $f \equiv 0$.

Exercise 16.4. *Prove Lemma 16.3*

The metric induced by a norm $\|\cdot\|$ is $d(f,g) := \|f-g\|$. *Uniform convergence* is convergence *in L^∞*-norm.

Definition 16.5.
- Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions, $f_n : D \to \mathbb{R}$. We say that f_n *uniformly converges* to f, if f_n converges to f *with respect to the L^∞-metric*, that is if
$$\lim_{n\to\infty} \|f_n - f\|_{L^\infty(D)} = 0.$$
that is
$$\lim_{n\to\infty} \sup_{x\in D} |f_n(x) - f(x)| = 0.$$
Cf. Figure 16.4
- We say $f : D \to \mathbb{R}$ belongs to $L^\infty(D)$, in words: f is a bounded function from D to $\mathbb{R}$, in formulas $f \in L^\infty(D)$, if $\|f\|_{L^\infty(D)} \equiv \sup_D |f| < \infty$

Observe: For uniform convergence we do *not* assume that $f_n \in L^\infty(D)$!

[31] why this letters: L^∞? Roughly it goes like follows: we know $\ell^p(\mathbb{N})$ are all *sequences* $(x_n)_{n\in\mathbb{N}}$ such that $\|(x_n)_{n\in\mathbb{N}}\|_{\ell^p} := (\sum_{n=1}^\infty |x_n|^p)^{\frac{1}{p}} < \infty$. Formally, letting $p \to \infty$, we see that ℓ^∞ should be $\|(x_n)_{n\in\mathbb{N}}\|_{\ell^\infty} = \sup_{n>0} |x_n|$. Similarly for *functions* $f : [a,b] \to \mathbb{R}$ we set $\|f\|_{L^p([a,b])} = \left(\int_{[a,b]} |f|^p\right)^{\frac{1}{p}}$ and again formally as $p \to \infty$ we see that we should set $\|f\|_{L^\infty} := \sup_x |f(x)|$.

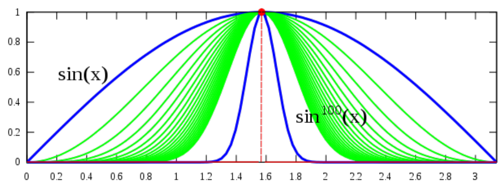

FIGURE 16.4. The sequence of functions $f_n(x) := \sin^n(x)$ on $[0, \pi]$ converges pointwise to a discontinuous function $f(x)$ with $f(\frac{\pi}{2}) = 1$ and $f(x) = 0$ for $x \neq \frac{\pi}{2}$. This convergence is not uniform, since $\|f_n - f\|_{L^\infty[0,\pi]} = 1$. (image: JoKalliauer/Wikipedia)

Example 16.6. Let $f(x) = x$. Then $f \not\in L^\infty(\mathbb{R})$.

However if we set $f_n(x) := x + \frac{1}{n}$, then

$$f(x) - f_n(x) = \frac{1}{n}$$

and thus

$$\|f - f_n\|_{L^\infty(\mathbb{R})} = \frac{1}{n} \xrightarrow{n\to\infty} 0.$$

That is also unbounded functions may converge uniformly.

Why do we call this uniform convergence?

Well, let's rewrite *pointwise convergence* $f_n(x) \to f(x)$ in ε-N form.

$$\forall x \in D,\ \forall \varepsilon > 0\ \exists N = N(\varepsilon, x) \in \mathbb{N}: \quad |f_n(x) - f(x)| < \varepsilon \quad \forall n \geq N.$$

As with continuity and uniform continuity, uniform convergence switches the order of x and N, i.e. it makes N independent of x.

Proposition 16.7. *Let $(f_n)_{n\in\mathbb{N}}$ and f be functions $D \to \mathbb{R}$. Then the following are equivalent*

(1) $\|f_n - f\|_{L^\infty(D)} \xrightarrow{n\to\infty} 0$
(2) $\forall \varepsilon > 0\ \exists N = N(\varepsilon) \in \mathbb{N}: \quad |f_n(x) - f(x)| < \varepsilon \quad \forall n \geq N\ \forall x \in D.$

In particular, uniform convergence implies pointwise convergence (but not vice versa, think of x^{2n} in $(-1, 1)$).

Proof. (1) is equivalent to saying:

for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\sup_{x\in D} |f_n(x) - f(x)| \leq \varepsilon \quad \forall n \geq N.$$

But this is equivalent to saying: for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$|f_n(x) - f(x)| \leq \varepsilon \quad \forall n \geq N \quad \forall x \in D.$$

which is (2).

The fact that we have $\leq \varepsilon$ instead of $< \varepsilon$ can be remedied by taking $\leq \frac{\varepsilon}{2} < \varepsilon$. □

Example 16.8.

- Let $f_n(x) := \frac{n}{nx+1}$ on $(0,1)$. Then $f_n(x) \xrightarrow{n\to\infty} \frac{1}{x}$ for all $x \in (0,1)$. So we have for $f(x) = \frac{1}{x}$ that $f_n(x)$ converges *pointwise* to f in $(0,1)$. However,

$$\|f_n - f\|_{L^\infty((0,1))} = \sup_{x\in(0,1)} \left|\frac{n}{nx+1} - \frac{1}{x}\right| \geq \lim_{x\to 0} \sup_{x\in(0,1)} \left|\frac{n}{nx+1} - \frac{1}{x}\right| = \infty.$$

  That is f_n does *not* converge to f uniformly.
- $f_n(x) := x^{2n}$, in $(-1,1)$. Then for $f(x) = 0$, $f_n(x) \xrightarrow{n\to\infty} f(x)$ pointwise in $(-1,1)$. However

$$\|f_n - f\|_{L^\infty(-1,1)} = 1 \not\xrightarrow{n\to\infty} 0.$$

- $f_n(x) := \frac{\sin(nx)}{n}$, then

$$\|f_n - 0\|_{L^\infty(\mathbb{R})} \leq \frac{1}{n} \xrightarrow{n\to\infty} 0.$$

Theorem 16.9. *$(L^\infty(D), \|\cdot\|_{L^\infty(D)})$ is a Banach space. That is, it is complete. That means, any Cauchy sequence $(f_n)_{n\in\mathbb{N}} \subset L^\infty(D)$ is convergent with respect to $L^\infty(D)$-norm.*

A sequence $(f_n)_{n\in\mathbb{N}} \subset L^\infty(D)$ is called Cauchy, if for any $\varepsilon > 0$ there exists $N > 0$ such that

$$\|f_n - f_m\|_{L^\infty(D)} < \varepsilon \quad \forall n, m \geq N.$$

Proof. Let $x \in D$. Then $(f_n(x))_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. Since $\mathbb{R}$ is complete, Theorem 4.4, there exists a limit value, which we call $f(x) := \lim_{n\to\infty} f(x_n)$.

$f : D \to \mathbb{R}$ is a function, and it is our candidate for uniform convergence. Let $\varepsilon > 0$. From the Cauchy sequence property, take $N = N(\varepsilon)$ such that

$$\|f_n - f_m\|_{L^\infty(D)} < \frac{\varepsilon}{2} \quad \forall n, m \geq N.$$

Then for any $x \in D$ we have

$$|f_n(x) - f_m(x)| < \frac{\varepsilon}{2} \quad \forall n, m \geq N.$$

Taking $m \to \infty$ (using the pointwise convergence of $f_m(x) \to f(x)$) we find

$$|f_n(x) - f(x)| \leq \frac{\varepsilon}{2} \quad \forall n \geq N.$$

Observe, the last estimate holds for any $x \in D$. That is,

$$\|f_n - f\|_{L^\infty(D)} = \sup_{x\in D} |f_n(x) - f(x)| \leq \frac{\varepsilon}{2} \quad \forall n \geq N.$$

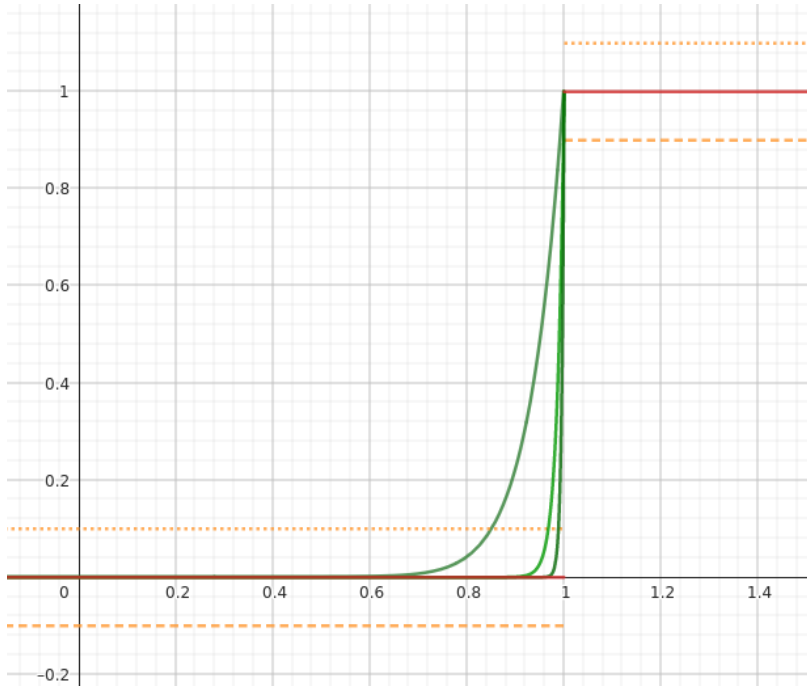

FIGURE 16.5. No Uniform convergence of the seqeunence $f_n(x) = x^n$ (green) to the (pointwise limit) function $f(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 1 \end{cases}$. The yellow dotted lines indicate the ε, and we see that the green curves f_n always traverse the yellow dotted lines

This is uniform convergence.

Lastly we need to show that $f \in L^\infty(D)$, i.e. that f is bounded. But this is easy: From the uniform convergence which we have proven already, take N such that

$$\|f_n - f\|_{L^\infty(D)} \leq 1 \quad \forall n \geq N.$$

Then

$$\|f\|_{L^\infty} \leq \|f_n - f\|_{L^\infty(D)} + \|f_N\|_{L^\infty} \leq 1 + \|f_N\|_{L^\infty}.$$

The right-hand side is finite, so f is bounded. □

Exercise 16.10. [Leb, 6.1.2]

(1) *Find the pointwise limit* $\dfrac{e^{x/n}}{n}$ *for* $x \in \mathbb{R}$.
(2) *Is the limit uniform on* $\mathbb{R}$*?*
(3) *Is the limit uniform on* $[0, 1]$*?*

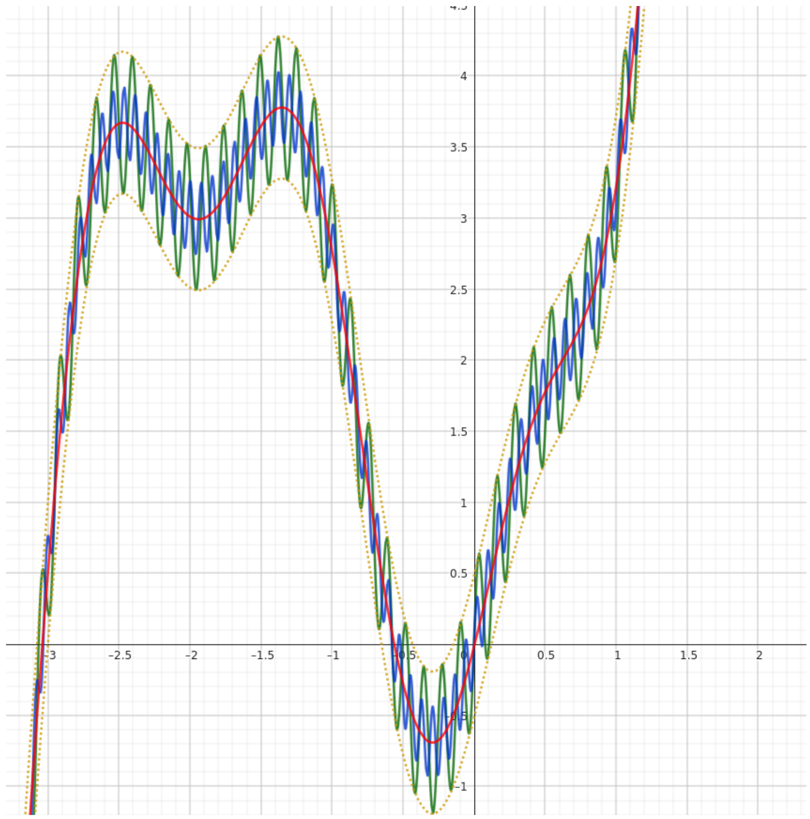

FIGURE 16.6. Uniform convergence can be pretty wild. The red line is pretty well approximated by the green (and even better by the blue) function w.r.t to uniform convergence (the yellow dotted line is the red function $\pm\varepsilon$)! We observe: uniform convergence does not control the derivative.

Exercise 16.11. [Leb, 6.1.5] *Suppose that $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ defined on some set A converge to f and g respectively uniformly on A. Show that $(f_n+g_n)_{n\in\mathbb{N}}$ converges uniformly to $f+g$ on A.*

Exercise 16.12. *Show that if $f_n : \mathbb{R} \to \mathbb{R}$ and $g_n : \mathbb{R} \to \mathbb{R}$ belong to $L^\infty(\mathbb{R})$ and converge uniformly to some f, g, then the product $f_n g_n$ converges uniformly to fg.*

Exercise 16.13. [Leb, 6.1.6] *Find an example of sequences of functions $f_n : \mathbb{R} \to \mathbb{R}$ and $g_n : \mathbb{R} \to \mathbb{R}$ that converge uniformly to some f and $g : \mathbb{R} \to \mathbb{R}$, but such that $(f_n g_n)_{n\in\mathbb{N}}$ does not converge uniformly to fg on $\mathbb{R}$.*

Hint: Which condition from Exercise 16.12 is missing?

Exercise 16.14. *Let $s < t$ and $a < b$ be finite numbers.*

Assume that we have a sequence of functions

$$f_k : [s,t] \to \mathbb{R} \quad \forall k \in \mathbb{N}$$

such that f_k uniformly converges to some $f : [s,t] \to \mathbb{R}$.

Moreover we assume $a \le f_k(r) \le b$ for all $r \in [s,t]$ and for all $k \in \mathbb{N}$.

Next, assume $F : [s,t] \times [a,b] \to \mathbb{R}$ is uniformly continuous. That is

$$\forall \varepsilon > 0 : \exists \delta > 0 \quad |F(r_1,x_1) - F(r_2,x_2)| < \varepsilon \quad \forall r_1, r_2 \in [s,t],\ x_1, x_2 \in [a,b] : \quad |r_1 - r_2| + |x_1 - x_2| < \delta.$$

Show that

$$g_k(s) := F(s, f_k(s))$$

converges uniformly to

$$g(s) := F(s, f(s)).$$

Now we turn to the properties of limits under uniform convergence.

First we treat boundedness. The example $f_n(x) := \frac{n}{nx+1} \xrightarrow{n\to\infty} \frac{1}{x}$ for $x \in (0,1)$ shows that the pointwise limit of bounded functions may not be bounded.

Proposition 16.15. *Let $f_n \in L^\infty(D)$. If $\|f_n - f\|_{L^\infty(D)} \xrightarrow{n\to\infty} 0$ then f is bounded.*

Proof. We showed this already in the last part of the proof of Theorem 16.9. □

Remark 16.16.
- Proposition 16.15 does not assume a uniform bound on f_n (i.e. we do not have $\sup_n \sup_x |f_n(x)| < \infty$ – in this case the pointwise limit would also be bounded).
- Proposition 16.15 implies: If a sequence of bounded functions converges to an unbounded function, then this convergence is not uniform.

Continuity is also stable under uniform convergence (recall that x^{2n}, $x \in [-1,1]$ gave an example of a sequence of functions continuous in $[-1,1]$ which pointwise converge to a discontinuous function).

Proposition 16.17. *Let $f_n : D \to \mathbb{R}$ be continuous and assume that $\|f_n - f\|_{L^\infty(D)} \xrightarrow{n\to\infty} 0$. Then f is continuous.*

Proof. Let $\varepsilon > 0$ and $x \in D$. Let $N \in \mathbb{N}$ such that

$$\|f_n - f\|_{L^\infty(D)} \le \frac{\varepsilon}{4} \quad \forall n \ge N.$$

For this fixed N take $\delta > 0$ such that

$$|f_N(x) - f_N(y)| \le \frac{\varepsilon}{4} \quad \forall y \in D,\ |y - x| < \delta.$$

Then we have for any $y \in D$, $|y-x| < \delta$,

$$\begin{aligned}|f(x)-f(y)| \leq &|f_N(x)-f_N(y)| + |f_N(x)-f(x)| + |f_N(y)-f(y)|\\ \leq &|f_N(x)-f_N(y)| + \|f_N - f\|_{L^\infty(D)} + \|f_N - f\|_{L^\infty(D)}\\ \leq &\frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4}\\ < &\varepsilon.\end{aligned}$$

□

Corollary 16.18. *Assume that f_n is continuous and uniformly converges to f. Then we can interchange limits (if either side exists, then the other side exists*

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

This may be false if f_n converges only pointwise to f.

Proof. Obvious since f is continuous by Proposition 16.17. □

The statement of Corollary 16.18 is false in general if we only have pointwise convergence, e.g. take $f_n(x) = \frac{1}{nx+1}$, which pointwise converges to $f(x) = \begin{cases} 1 & x = 0\\ 0 & x \neq 0.\end{cases}$ That is, $\lim_{n\to\infty}\lim_{x\to 0} f_n(x) = \lim_{n\to\infty} 1 = 1$ But $\lim_{x\to 0} f(x) = 0$.

We have seen above that the pointwise limit of integrable functions may not be integrable (think of the approximation of the Dirichlet function, Example 16.2)

Proposition 16.19. *Let $f_n \in \mathcal{R}[a,b]$, and assume that $\|f_n - f\|_{L^\infty([a,b])} \xrightarrow{n\to\infty} 0$. Then $f \in \mathcal{R}[a,b]$ and we have*

$$\lim_{n\to\infty}\int_{[a,b]} f_n = \int_{[a,b]}\lim_{n\to\infty} f_n = \int_{[a,b]} f.$$

Proof. Since f_n are Riemann integrable each f_n is bounded Lemma 14.10. By Proposition 16.15 we then have that f is bounded, so upper and lower Riemann integral exist.

Let $\varepsilon > 0$, then there exists $N \in \mathbb{N}$ such that

$$\|f_n - f\|_{L^\infty([a,b])} < \frac{\varepsilon}{2(b-a)}.$$

We then have

$$\begin{aligned}&\underline{\int_{[a,b]}} f - \overline{\int_{[a,b]}} f\\ =&\underbrace{\underline{\int_{[a,b]}} f_n - \overline{\int_{[a,b]}} f_n}_{=0} + \underline{\int_{[a,b]}} (f - f_n) - \overline{\int_{[a,b]}} f - f_n\end{aligned}$$

So we have for any $n \geq N$

$$\left|\underline{\int_{[a,b]}} f - \overline{\int_{[a,b]}} f\right| \leq \underline{\int_{[a,b]}} |f - f_n| + \overline{\int_{[a,b]}} |f - f_n|$$

Since $|f(x) - f_n(x)| \leq \|f - f_n\|_{L^\infty([a,b])}$ for any $x \in [a,b]$ we find (cf. Lemma 14.5)

$$\left|\underline{\int_{[a,b]}} f - \overline{\int_{[a,b]}} f\right| \leq 2(b-a)\|f - f_n\|_{L^\infty([a,b])} < \varepsilon.$$

This holds for any $\varepsilon > 0$, so if $\varepsilon \to 0$ we have shown

$$\left|\underline{\int_{[a,b]}} f - \overline{\int_{[a,b]}} f\right| = 0.$$

That is, f is Riemann integrable.

With this, we have for any $n \geq N$

$$\left|\int_{[a,b]} f - \int_{[a,b]} f_n\right| = \left|\int_{[a,b]} (f - f_n)\right| \leq (b-a)\,\|f - f_n\|_{L^\infty([a,b])} < \varepsilon.$$

This shows the convergence of the integrals. □

Example 16.20. Compute

$$\lim_{n\to\infty} \int_0^1 \frac{nx + \sin(nx^2)}{n}\, dx.$$

In principle this is difficult, because there is no closed form of the antiderivative of $\sin(nx^2)$. However, for $f_n(x) := \frac{nx+\sin(nx^2)}{n}$ and $f(x) = x$ we have uniform convergence, i.e.

$$\lim_{n\to\infty} \|f_n - f\|_{L^\infty([0,1])} = 0$$

Then by Proposition 16.19 we have

$$\lim_{n\to\infty} \int_0^1 \frac{nx + \sin(nx^2)}{n}\, dx = \int_0^1 \lim_{n\to\infty} \frac{nx + \sin(nx^2)}{n}\, dx = \int_0^1 x\, dx.$$

And since we know from Calculus 1, $\int_0^1 x\, dx = \frac{1}{2}$ we have computed

$$\lim_{n\to\infty} \int_0^1 \frac{nx + \sin(nx^2)}{n}\, dx = \frac{1}{2}.$$

Again, if we only have pointwise convergence, in general

$$\lim_{n\to\infty} \int_{[a,b]} f_n \neq \int_{[a,b]} f.$$

See e,g, (16.1) for an example.

Next we treat differentiability. We have already seen in Example 16.2 that even uniform convergence does not preserve differentiability; and even if the limit is differentiable, the limit of the derivative may not equal the derivative of the limit.

Example 16.21. Let for $x \in \mathbb{R}$,

$$f_n(x) := \frac{x}{1+nx^2}.$$

Then

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

Pointwise we have $\lim_{n\to\infty} f_n(x) = 0 =: f(x)$.

We also have uniform convergence, indeed for any $x \neq 0$, observe that $1 + nx^2 \geq 2\sqrt{n}|x|$ (Cauchy-Schwarz), so that

$$|f_n(x) - f(x)| \leq \frac{|x|}{2\sqrt{n}|x|} = \frac{1}{\sqrt{n}}.$$

This implies (for $x = 0$ we have $f_n(0) = 0 = f(0)$)

$$\|f_n - f\|_{L^\infty(\mathbb{R})} \leq \frac{1}{\sqrt{n}} \xrightarrow{n\to\infty} 0.$$

So f_n converges to f uniformly. However, $f'(0) = 0$, but $f_n'(0) = 1$, so we have $\lim_{n\to\infty} f_n'(0) \neq f(0)$

So we need to strengthen uniform convergence to preserve derivatives converging. The simplest version is

Proposition 16.22. *Let*[32] $f_n \in C^1([a,b])$ *and* f_n *is a Cauchy sequence with respect to the* C^1*-norm,*

$$\|g\|_{C^1([a,b])} := \|g\|_{L^\infty([a,b])} + \|g'\|_{L^\infty([a,b])}.$$

That is, assume that for any $\varepsilon > 0$ *there exists* $N \in \mathbb{N}$ *such that*

$$\|f_n - f_m\|_{L^\infty([a,b])} + \|f_n' - f_m'\|_{L^\infty([a,b])} < \varepsilon \quad \forall n, m \geq N.$$

Then there exists $g \in C^1([a,b])$ *such that*

$$\|f_n - g\|_{C^1([a,b])} \xrightarrow{n\to\infty} 0.$$

That is, f_n *and* f_n' *converge uniformly to* g *and* g'*, respectively, on* $[a,b]$*.*

Exercise 16.23. *Prove Proposition 16.22.*

Exercise 16.24. *Find an example of a sequence of function* $f_n : [0,1] \to \mathbb{R}$ *such that*

- f_n *is continuously differentiable, and* f_n' *uniformly converges to* 0
- f_n *does not converge*

Hint: Actually can you find a sequence of functions f_n *such that* $f_n' = 0$ *in* $[0,1]$ *for all* $n \in \mathbb{N}$ *but* f_n *does not converge?*

[32] Recall that we say $f \in C^k$ if f is continuous, the first derivative exists and f' is continuous, the k-th derivative exists and $f^{(k)}$ is continuous

Theorem 16.25. *Let $f_n \in C^1([a,b])$ be a sequence such that*

- *f_n' converges uniformly to some $g : [a,b] \to \mathbb{R}$, i.e.*
$$\|f_n' - g\|_{L^\infty([a,b])} \xrightarrow{n\to\infty} 0$$
- *there exists at least one single point $x_0 \in [a,b]$ such that f converges pointwise at this point, i.e. there is $y \in \mathbb{R}$ such that*
$$f_n(x_0) \xrightarrow{n\to\infty} y.$$

Then there exists $f \in C^1([a,b])$ such that f_n converges uniformly to f, i.e.
$$\|f_n - f\|_{L^\infty([a,b])}$$
and we have $f'(x) = g(x)$ for all $x \in [a,b]$.

Proof. Since g is the limit of continuous functions $f_n' \in C^0([a,b])$, and thus continuous and thus Riemann-integrable. So we can use the fundamental theorem of calculus.

Set
$$f(x) := y_0 + \int_{x_0}^{x} g(z)\, dz$$
Then we have from the fundamental theorem, Theorem 15.3,
$$f'(x) = g(x).$$
So all we need to show is that f_n uniformly converges to f. Again, by Theorem 15.3[33],
$$f_n(x) = f_n(x_0) + \int_{x_0}^{x} f_n'(z)\, dz.$$
So we have
$$f_n(x) - f(x) = f_n(x_0) - y_0 + \int_{x_0}^{x} g(z) - f_n'(z)\, dz.$$
This implies
$$\|f_n - f\|_{L^\infty([a,b])} \le |f_n(x_0) - y_0| + |b-a| \|g - f_n'\|_{L^\infty([a,b])} \xrightarrow{n\to\infty} 0.$$
That is f_n converges uniformly to f and we can conclude. □

Just as a remark, the continuous differentiability is not needed in the above theorem. Indeed we have

Theorem 16.26. *Let f_n be differentiable on (a,b) and assume*

- *f_n converges pointwise to f in (a,b)*

[33]more precisely: If we set
$$h_n(x) := f_n(x_0) + \int_{x_0}^{x} f_n'(z)\, dz$$
the Fundamental theorem tells us $h_n'(x) = f_n'(x)$ for all $x \in [a,b]$. Then the mean value theorem, Proposition 12.3, tells us that $h_n(x) - f_n(x) \equiv c$ is a constant function. But $h_n(x_0) = f_n(x_0)$, so indeed $h_n(x) - f_n(x) \equiv 0$, that is $h_n(x) = f_n(x)$ for all $x \in [a,b]$.

- f_n' *converges uniformly to* g *in* (a, b).

Then f *is differentiable on* (a, b) *and we have* $f' = g$.

Exercise 16.27. *Set*

$$f_n(x) := n\left((x + \frac{1}{n})^3 - x^3\right)$$

Show that $(f_n)_{n\in\mathbb{N}}$ *converges pointwise on* $\mathbb{R}$ *to* $f(x) = 3x^2$.

Exercise 16.28. *Set*

$$f_n(x) := 1 + x + x^2 + \ldots + x^n$$

and define

$$f(x) := \frac{1}{1-x} \quad x \in \mathbb{R} \setminus 1.$$

(1) *Show that* $(f_n)_{n\in\mathbb{N}}$ *converges pointwise to* f *on* $(-1, 1)$.
(2) *Show that* $(f_n)_{n\in\mathbb{N}}$ *does not converge uniformly all on* $(-1, 1)$
(3) *Show that* $(f_n)_{n\in\mathbb{N}}$ *does not converge at all on* $\mathbb{R} \setminus (-1, 1)$.
(4) *Show that* f_n *converes to* f *uniformly on* $[-r, r]$ *for any* $r < 1$.

Exercise 16.29. *Set*

$$f_n(x) := \frac{nx}{1 + nx^2}, \quad x \in \mathbb{R}.$$

(1) *Show that* f_n *converges pointwise on* $\mathbb{R}$
(2) *Show that* f_n *does not converge uniformly on* $\mathbb{R}$ *(Hint: consider the proposed limit function, and recall Proposition 16.17)*

Exercise 16.30. *Let* $f_n(x) := (x - \frac{1}{n})^2$ *for* $x \in [0, 1]$. *Does* f_n *converge uniformly?*

Exercise 16.31. *Assume* $f_n : \mathbb{R} \to \mathbb{R}$ *is uniformly continuous.*

Assume that f_n *uniformly converges to* $f : \mathbb{R} \to \mathbb{R}$.

Show that f *is uniformly continuous.*

Exercise 16.32. *Set*

$$f_n(x) := \frac{x^n}{n + x^n} \quad x \geq 0$$

Is f_n *pointwise or uniformly convergent?*

Exercise 16.33. *Set*

$$f_n(x) := e^{-\frac{x^2}{n}} \quad x \in \mathbb{R}$$

Is f_n *pointwise or uniformly convergent?*

17. Series of functions – The Weierstrass M-Test

In Calculus we defined for a sequence $(x_k)_{k\in\mathbb{N}}$ the notion of a *series*

$$\sum_{k=1}^{\infty} x_k := \lim_{n\to\infty} \sum_{k=1}^{n} x_k = \lim_{n\to\infty} (x_1 + x_2 + \ldots + x_n).$$

So if we have a sequence of functions $f_k : [a,b] \to \mathbb{R}$ we can define

$$\left(\sum_{k=1}^{\infty} f_k\right)(x) := \lim_{n\to\infty} \sum_{k=1}^{n} f_k(x).$$

We have seen above that pointwise convergence of this series is often not that useful, uniform convergence is much stronger (e.g. it preserves continuity in the limit!)

Theorem 17.1 (Cauchy condition)**.** *Let $f_k : D \to \mathbb{R}$ be a sequence of functions.*

The following are equivalent

(1) $\sum_{k=1}^{\infty} f_k$ converges uniformly, that is

$$\|\sum_{k=1}^{n} f_k - \sum_{k=1}^{\infty} f_k\|_{L^\infty(D)} \xrightarrow{n\to\infty} 0$$

(2) For any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for any $n > N$

$$\left|\sum_{k=n+1}^{n+L} f_k(x)\right| < \varepsilon \quad \forall L \in \mathbb{N}, x \in D.$$

Proof. This is really just the notion of Cauchy sequence (with respect to the L^∞-norm) for the sequence

$$S_n(x) := \sum_{k=1}^{n} f_k(x).$$

Indeed, by Theorem 16.9, S_n is converging uniformly in D if and only if for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for any $n, m > N$ we have

$$\|S_n - S_m\|_{L^\infty(D)} < \varepsilon.$$

Assuming that $m > n$ we set $L := m - n$ and then this is the same as to say

$$\|S_n - S_{n+L}\|_{L^\infty(D)} < \varepsilon.$$

which is exactly what (2) is requesting. □

A more checkable test is called the Weierstrass M-test, it shows absolute convergence through the dominated convergence theorem for series.

Theorem 17.2 (Weierstrass M-test)**.** *Let $f_k : D \to \mathbb{R}$ be a sequence of functions and let $(M_k)_{k\in\mathbb{N}}$ be sequence of nonnegative numbers such that*

(1) $|f_k(x)| \leq M_k$ for all $x \in D$ and
(2) $\sum_{k=1}^{\infty} M_k < \infty$ is convergent (thus: absolutely convergent)

Then $\sum_{k=1}^{\infty} f_k$ converges uniformly in D.

Proof. We have for any $x \in D$ and $n \in \mathbb{N}$, and $L \in \mathbb{N}$,

$$\left|\sum_{k=n+1}^{n+L} f_k(x)\right| \leq \sum_{k=n+1}^{n+L} M_k.$$

Since $\sum_{k=1}^{\infty} M_k$ is convergent (and $M_k \geq 0$: absolutely convergent) for any $\varepsilon > 0$ there exists and $N \in \mathbb{N}$ such that

$$\sum_{k=n+1}^{n+L} M_k < \varepsilon \quad \forall n > N, L \in \mathbb{N}.$$

Thus by Theorem 17.1, $\sum_{k=1}^{\infty} f_k$ converges uniformly. □

Videolink
A lecture on the Weierstrass M-test, https://www.youtube.com/watch?v=WzLTgErxep4

The real example of this approach are power series, but here we look at some examples.

Example 17.3.
- $e^x := \sum_{k=0}^{\infty} \frac{x^k}{k!}$ converges uniformly in every interval $[a, b]$ and is $C^\infty((a, b))$, with $(e^x)' = e^x$.

  Indeed, we have $\left|\frac{x^k}{k!}\right| \leq M_k := \frac{\max\{|a|,|b|\}^k}{k!}$, and we know that $\sum_{k=1}^{\infty} M_k = e^{\max\{|a|,|b|\}} < \infty$. So e^x converges uniformly.

  Now let
  $$g_n(x) := \sum_{k=0}^{n} \frac{x^k}{k!}.$$
  This is a C^∞-function for each n, and since g_n converges uniformly to e^x we obtain that e^x is continuous.

  But also
  $$g_n'(x) = \sum_{k=1}^{n} \frac{kx^{k-1}}{k!} = \sum_{k=1}^{n} \frac{x^{k-1}}{(k-1)!} = \sum_{\tilde{k}=0}^{n-1} \frac{x^{\tilde{k}}}{\tilde{k}!}$$
  is a continuous function, converges uniformly to e^x.

  By Proposition 16.22 this implies that e^x is differentiable and
  $$(e^x)' = (\lim_{n\to\infty} g_n(x))' = \lim_{n\to\infty} g_n'(x) = e^x.$$
  By induction we conclude that $e^x \in C^\infty((a, b))$. This holds for any $-\infty < a < b < \infty$, and we thus say $e^x \in C^\infty(\mathbb{R})$.

- $\sum_{n=1}^{\infty} \frac{\sin(nx)}{n^2}$ is uniformly convergent on $\mathbb{R}$ since
$$\left|\frac{\sin(nx)}{n^2}\right| \leq \frac{1}{n^2}.$$
So we can apply the Weierstrass M-test.
- $\sum_{n=1}^{\infty} \frac{2x}{1+n^6x^2}$ converges uniformly in $\mathbb{R}$. Indeed, observe that since $2ab \leq a^2+b^2$ we have
$$2n^3|x| \leq 1+n^6x^2.$$
Consequently,
$$\left|\frac{2x}{1+n^6x^2}\right| \leq \frac{2|x|}{2n^3|x|} = \frac{1}{n^3}.$$
so we can apply the Weierstrass M-test.
- $\sum_{n=1}^{\infty} \frac{x}{n^2}$ converges uniformly in any finite interval (a,b), by Weierstrass M-test, since
$$\left|\frac{x}{n^2}\right| \leq \frac{\max\{|a|,|b|\}}{n^2}.$$
However it does not converge uniformly in $\mathbb{R}$. Indeed,
$$S_L(x) := \sum_{n=1}^{L} \frac{x}{n^2}$$
is not a Cauchy sequence in $L^\infty(\mathbb{R})$, because it is not even bounded, $\|S_L\|_{L^\infty(\mathbb{R})} = \infty$.
- Let $f(x) := \sum_{k=1}^{\infty} \frac{1}{1+k^2x}$. The series converges uniformly on $[a,\infty)$ for any $a>0$ but only pointwise on $(0,\infty)$. In particular f is continuous on $(0,\infty)$.

  To show that the series converges uniformly on $[a,\infty)$ (and thus, by Proposition 16.17, f is continuous) we use the Weierstrass M-test. Observe that
$$\frac{1}{1+k^2x} \leq \frac{1}{1+k^2a} =: M_k$$
and $\sum_{k=1}^{\infty} M_k < \infty$ by Calculus 2 arguments: indeed, since $a>0$ we have $\lim_{k\to\infty} \frac{k^2}{1+k^2a} = \frac{1}{a} < \infty$ and thus by the Limit comparison test for series, $\sum_{k=1}^{\infty} M_k < \infty$.

  In particular the series converges *pointwise* in $(0,\infty)$, since for any $x \in (0,\infty)$ it converges uniformly in $[x/2,\infty)$.

  However, the series does not converge uniformly on $(0,\infty)$. Indeed, we will show that it does not satisfy the Cauchy condition, see Theorem 17.1. Let
$$f_n(x) := \sum_{k=1}^{n} \frac{1}{1+k^2x}.$$
If f_n was uniformly converging in $(0,\infty)$, by Theorem 17.1
$$\forall \varepsilon>0 \quad \exists N \in \mathbb{N}: \quad \forall n>m\geq N: \sup_{x\in(0,\infty)} |f_n(x)-f_m(x)| < \varepsilon \tag{17.1}$$

Observe that

$$|f_n(x) - f_m(x)| = \sum_{k=m+1}^{n} \frac{1}{1+k^2x}.$$

Now let $x := \frac{1}{n^2} \in (0,1)$ then $k^2x \leq 1$ for all $k \leq n$, and thus

$$|f_n(x) - f_m(x)| = \sum_{k=m+1}^{n} \frac{1}{1+k^2x} \geq \sum_{k=m+1}^{n} \frac{1}{1+1} = \frac{n-m-1}{2}.$$

In particular,

$$\sup_{x\in(0,1)} |f_n(x) - f_m(x)| \geq \frac{n-m-1}{2}.$$

This is a contradiction to (17.1). Indeed, for $\varepsilon := 1$ and any $N \in \mathbb{N}$ we can choose $m := N+1$ and $n := N+6$ and have

$$\sup_{x\in(0,1)} |f_n(x) - f_m(x)| \geq \frac{n-m-1}{2} = \frac{N+6-N-2}{2} = 2 > 1 = \varepsilon.$$

So (17.1) is not satisfied, i.e. the series is not uniformly convergent.

Exercise 17.4. *Use the Weierstrass M-test to show that each of the following series converge uniformly in the given domain:*

(1) $\sum_{k=1}^{\infty} \frac{x^k}{k^2}$ *for* $x \in [-1,1]$
(2) $\sum_{k=1}^{\infty} \frac{1}{x^k}$ *for* $x \in [2,\infty)$
(3) $\sum_{k=1}^{\infty} \frac{x^k}{x^k+1}$ *for* $-\frac{1}{2} < x < \frac{1}{2}$

Exercise 17.5. *Prove that the following series*

$$f(x) := \sum_{n=1}^{\infty} \frac{n^2+x^4}{n^4+x^2}$$

converges to a continuous function $f : \mathbb{R} \to \mathbb{R}$.

Hint: Use Theorem 17.2 and Proposition 16.17 to show that $f : [-R,R] \to \mathbb{R}$ *is continuous for any* $R > 0$.

Exercise 17.6. *Show that the series*

$$f(x) := \sum_{n=1}^{\infty} \frac{\sin(nx)}{n^p}$$

defines a continuous function on $\mathbb{R}$.

Exercise 17.7. *Show that the series*

$$f(x) := \sum_{n=1}^{\infty} \frac{\cos(2^n x)}{3^n}$$

defines a continuous function on $\mathbb{R}$. *Show that it is actually continuously differentiable and compute its derivative.*

Exercise 17.8. *Consider the series*

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

(1) *Show that for any $R > 0$*

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!} = e^x \quad \text{with uniform convergence in } [-R, R]$$

(2) *Show that the above convergence is not uniform in $\mathbb{R}$.*

Exercise 17.9. *Consider the series*

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

Show that f is continuous on $(1, \infty)$. Show that the convergence of the series is not uniform on $(1, \infty)$.

Exercise 17.10. *Assume that we have a sequence $(a_n)_{n=1}^{\infty} \subset \mathbb{R}$ so that the series*

$$\sum_{n=1}^{\infty} a_n x^n$$

converges uniformly on $\mathbb{R}$.

Show that all but finitely many sequence elements of $(a_n)_{n\in\mathbb{N}}$ are zero, i.e. show that there exists $N > 0$ such that $a_n = 0$ for all $n \geq N$.

18. Power series

A very important class of function series are so-called power series (this importance becomes clearer below with Taylor's theorem).

The idea that $f_n(x) = a_n(x - x_0)^n$, that is we are interested in convergence of

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n.$$

Here (and henceforth) in a sum we use the convention $(x - x_0)^0 = 1$ even for $x = x_0$.

By a change of variables $y := x - x_0$ the question of convergence is about

$$\sum_{n=0}^{\infty} a_n y^n.$$

Easy case: this sum converges at $y = 0$.

Definition 18.1. Let $(a_n)_{n\in\mathbb{N}_0}$ be a sequence. The *radius of convergence* of the power series

$$\sum_{n=0}^{\infty} a_n y^n$$

is

$$\sup\left\{r \geq 0 : \sum_{n=0}^{\infty} |a_n| r^n < \infty\right\}$$

Example 18.2.

- The radius of convergence might be 0. Take for example $a_n := n^n$. Then for any $r > 0$

$$\sum_{n=1}^{\infty} n^n r^n = \infty,$$

since for n suitably large $nr \geq 1$, so $(nr)^n \geq 1$.

- The radius of convergence might be ∞. The trivial case is $a_n = 0$. But also more interestingly,

$$a_n = \frac{1}{n!}$$

Because then $\sum_{n=1}^{\infty} a_n r^n = e^r$ which is finite for any $r < \infty$.

- Assume that there exists $r_0 > 0$ such that $\Lambda := \sup_{n\in\mathbb{N}} a_n r_0^n < \infty$. Then for any $r < r_0$ we have

$$a_n r^n = a_n\, r_0^n \left(\frac{r}{r_0}\right)^n$$

so that for $\theta := \left(\frac{r}{r_0}\right) < 1$ we have

$$|a_n r^n| \leq \Lambda \theta^n$$

By dominated convergence (observe $\theta < 1$)

$$\sum_{n=0}^{\infty} |a_n| r^n \leq \Lambda \sum_{n=0}^{\infty} \theta^n < \infty.$$

Theorem 18.3. *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence with radius of convergence $R > 0$. Let $x_0 \in \mathbb{R}$*

(1) for any $r < R$ the sequence

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

converges uniformly in $[x_0 - r, x_0 + r]$.

(2) for any $x \notin [x_0 - R, x_0 + R]$ the sequence

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

does not converge.

Proof. (1) Let $r < R$. Since R is the supremum over all ρ such that $\sum_{n=1}^{\infty} |a_n|\rho^n$ converges, there must be some $\rho > r$ such that

$$\sum_{n=0}^{\infty} |a_n|\rho^n < \infty.$$

In particular this implies that $\lim_{n\to\infty} |a_n|\rho^n = 0$, and thus $\Lambda := \sup_{n\in\mathbb{N}} |a_n|\rho^n < \infty$.

For any $x \in [x_0 - r, x_0 + r]$ we then have

$$|a_n(x-x_0)^n| \leq |a_n|r^n = \underbrace{|a_n|\rho^n}_{\leq \Lambda} \left(\frac{r}{\rho}\right)^n$$

Set $\theta := \frac{r}{\rho} \in (0,1)$. Then for $M_n := \Lambda\theta^n$ we have $\sum_{n=0}^{\infty} M_n < \infty$ and

$$|a_n(x-x_0)^n| \leq M_n \quad \forall x \in [x_0 - r, x_0 + r]$$

By the Weierstrass M-test, the sequence $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ converges uniformly in $[x_0 - r, x_0 + r]$.

(2) Let now $x \in \mathbb{R}$ and assume that

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n$$

converges. This implies that $\lim_{n\to\infty} a_n(x-x_0)^n = 0$, that is $\Lambda := \sup_n |a_n||x-x_0|^n < \infty$. Now for any $0 < \rho < |x-x_0|$ we then have

$$\sum_{n=0}^{\infty} a_n\rho^n = \sum_{n=0}^{\infty} \underbrace{a_n|x-x_0|^n}_{\leq \Lambda} \left(\underbrace{\frac{\rho}{|x-x_0|}}_{\leq 1}\right)^n < \infty$$

But this means that (by the argument in (1)) the radius of convergence $R \geq \rho$.

So we have shown: for any $\rho < |x-x_0|$ that $R \geq \rho$. This implies that $R \geq |x-x_0|$.

In other words, if $|x-x_0| > R$ then $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ cannot converge.

□

Example 18.4.

- In Theorem 18.3 we do not know what happens for $x = x_0 \pm R$ if R is the radius of convergence. Indeed, for $a_n = 1$ the radius of convergence is 1 and for $x = x_0 + 1$ we have divergence, but for $x := x_0 - 1$ we have convergence
- As we learned in Calculus, the radius of convergence can be found via the ratio test, it is the largest number $R \geq 0$ such that

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} R \leq 1$$

Because then by the ratio test if $|x-x_0| < R$,

$$\lim_{n\to\infty} \frac{|a_{n+1}||x-x_0|^{n+1}}{|a_n||x-x_0|^n} = \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}|x-x_0| < 1.$$

Alternatively we can use the root test,

$$\lim_{n\to\infty} \sqrt[n]{|a_n|} R \le 1.$$

Observe that the formal derivative of a power series is

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

is

$$\sum_{n=1}^{\infty} n\, a_n (x - x_0)^{n-1}$$

Good news is that the radius of convergence does not change if we take derivatives.

Proposition 18.5. *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence with radius of convergence $R > 0$, and for some $x_0 \in \mathbb{R}$ set*

$$f(x) := \sum_{n=0}^{\infty} a_n (x - x_0)^n \quad x \in (x_0 - R, x_0 + R).$$

Then f is differentiable in $(x_0 - R, x_0 + R)$ and we have

$$f'(x) = \sum_{n=1}^{\infty} n\, a_n (x - x_0)^{n-1} \quad x \in (x_0 - R, x_0 + R).$$

which has the same radius of convergence. In particular $f \in C^\infty((x_0 - R, x_0 + R))$.

Proof. We only need to show that the radius of convergence of $(na_n)_{n\in\mathbb{N}}$ is the same as the radius of convergence $(a_n)_{n\in\mathbb{N}}$. Everything else (in particular the formula for f') then follows from the uniform convergence, Theorem 18.3, and the interchange of limits Proposition 16.22.

To obtain that the radius of convergence of $(a_n)_n$ and $(na_n)_n$ we argue as follows.

Let R denote the radius of convergence of a_n, and let R' the radius of convergence of na_n.

Firstly we show $\underline{R' \le R}$. This is easy: since $|a_n| \le n|a_n|$ (for $n \ge 1$) we have $\sum_n n|a_n|r^n < \infty$ then also $\sum_n |a_n|r^n < \infty$.

So let $r < R'$, then by definition of the radius of convergence for $(na_n)_{n\in\mathbb{N}}$, which we called R', we have $\sum_n n|a_n|r^n < \infty$, thus $\sum_n |a_n|r^n < \infty$. But this implies that the radius of convergence $(a_n)_{n\in\mathbb{N}}$, which we called R, must satisfy $R > r$.

So we have

$$r < R \quad \forall r < R'.$$

This readily implies that $R' \le R$ (exercise: argue by contradiction!).

It remains to show that $R' \ge R$.

For this take any $r, \rho > 0$ such that $r < \rho < R$.

Since R is the radius of convergence of a_n and $\rho < R$, we have

$$\sum_{n=0}^{\infty} |a_n|\rho^n < \infty.$$

As before we conclude that

$$\Lambda := \sup_n |a_n|\rho^n < \infty.$$

Set $\theta := \left(\frac{r}{\rho}\right) < 1$ (since $r < \rho$). Observe that $\theta \in (0,1)$ implies that

$$\lim_{n\to\infty} n\theta^{\frac{n}{2}} = 0.$$

Thus

$$\lambda := \sup_n n\theta^{\frac{n}{2}} < \infty.$$

So we have

$$n|a_n|r^n = |a_n|\rho^n n \left(\frac{r}{\rho}\right)^n = \underbrace{|a_n|\rho^n}_{\leq \Lambda} \underbrace{n\;(\theta)^{\frac{n}{2}}}_{\leq \lambda} \theta^{\frac{n}{2}}$$

That is,

$$n|a_n|r^n \leq \Lambda\,\lambda\,\theta^{\frac{n}{2}}.$$

Thus, since $\theta < 1$, so $\theta^{\frac{1}{2}} = \sqrt{\theta} < 1$

$$\sum_{n=0}^{\infty} n|a_n|r^n \leq \Lambda\,\lambda \sum_{n=0}^{\infty} \theta^{\frac{n}{2}} = \Lambda\,\lambda \sum_{n=0}^{\infty} (\sqrt{\theta})^n < \infty$$

Thus the radius of convergence R' of $n|a_n|$ satisfies $R' \geq r$.

That is, we have shown: For any $r < \rho < R$ we have $R' \geq r$. Again this implies $R' \geq R$.

Indeed, if $R < R'$ then we could find and $r \in (R, R')$ such that $\sum_{n=0}^{\infty} |a_n|r^n < \infty$ which is impossible by the definition of R (and Theorem 18.3(2)). ☐

Exercise 18.6. *Let $(a_n)_n$ be a sequence with radius of convergence $R > 0$ and set*

$$f(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n \quad x \in (x_0 - R, x_0 + R)$$

Assume this series has radius $R > 0$. Show that $a_n = \frac{f^{(k)}(x_0)}{k!}$.

19. Taylor's Theorem

Recall the mean value theorem, Theorem 12.2. It stated that (under differentiability assumptions on f), for any x and y we find c such that.

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

FIGURE 19.1. Brook Taylor, 1685 - 1731. English, Mathematician

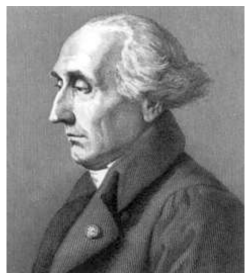

FIGURE 19.2. Joseph-Louis Lagrange, 1736 - 1813. Franco-Italian, Mathematician, Astronomer. Taylor's theorem was proposed by Taylor, but really proven by Lagrange.

Equivalently we could write

$$f(y) = f(x) + f'(c)(y - x).$$

We can interpret this as "$f(y) \approx f(x)$" up to the term $f'(c)(x - y)$ (which is a term which growth like $|x - y|$ (if f' is a bounded function). That is, the mean value theorem provides an approximation of $f(y)$ by a constant function $g = g(x)$.

Taylor's theorem Theorem 19.2 generalizes this to higher order polynomials; it approximates (suitably differentiable functions) with the *Taylor polynomial*

Definition 19.1. For a function f defined in an open neighborhood of x_0 the *n-th Taylor polynomial* of f at x_0 is given by

$$\begin{aligned} P_n(x) :=& \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \\ =& f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \frac{f^{(3)}(x_0)}{6}(x - x_0)^3 + \ldots + \frac{f^{(n)}}{n!}(x - x_0)^n. \end{aligned}$$

In this (and power series) we define $0^0 = 1$, that is $P_n(x_0) = f(x_0)$.

The mean value theorem, Theorem 12.2, is then the statement that for any x there exists c such that

$$f(x) - P_0(x) = f'(c)(x - x_0).$$

Taylor's theorem, generalizes this to arbitrary n.

Theorem 19.2 (Taylor)**.** *Let $f \in C^n([a,b])$ be a function with n continuous derivatives on $[a,b]$ such that $f^{(n)}$ is differentiable on (a,b). For any $x_0, x \in [a,b]$ there exists $c \in [x_0, x]$ such that*

$$f(x) = P_n(x) + R_n(x)$$

where the remainder term R_n is given by (observe: c depends on each choice of x)

$$R_n(x) := \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}.$$

Example 19.3. Let $f(x)$ be a polynomial of degree n, then $P_n(x) = f(x)$ (this follows from Taylor's theorem, because the $(n+1)$-st derivative of f vanishes.

Proof. Fix $x_0, x \in [a,b]$. For $x = x_0$ there is nothing to show, so we assume $x \neq x_0$. Define $M \in \mathbb{R}$ such that.

$$f(x) = P_n(x) + M(x-x_0)^{n+1}. \tag{19.1}$$

That is,

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

We need to show that M has the form $\frac{f^{(n+1)}(c)}{(n+1)!}$ (for some c).

For this we set

$$g(y) := f(y) - P_n(y) - M(y-x_0)^{n+1}.$$

By the choice of M we have $g(x) = 0$. Moreover, by the choice of P_n we have $g(x_0) = g'(x_0) = \ldots = g^{(n)}(x_0) = 0$.

So we apply the mean value theorem, Theorem 12.2/ Rolle's theorem, Theorem 12.1. We obtain $x_1 \in (x_0, x)$ such that

$$g'(x_1) = 0.$$

Since $g'(x_0) = 0$ we apply the mean value theorem, Theorem 12.2/ Rolle's theorem, Theorem 12.1 again, and we obtain $x_2 \in (x_0, x_1)$ such that

$$g'(x_2) = 0.$$

etc. and we find $c := x_{n+1} \in (x_0, x_n) \subset \ldots \subset (x_0, x)$ such that

$$g^{(n+1)}(c) = 0. \tag{19.2}$$

On the other hand we have from the definition of g,

$$g^{(n+1)}(c) = f^{(n+1)}(c) - 0 - (n+1)!M. \tag{19.3}$$

Together (19.2) and (19.3) imply

$$M = \frac{f^{(n+1)}(c)}{(n+1)!}.$$

Plugging this into (19.1) we have shown that

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

This concludes the proof. □

The following is a typical application of Taylor's theorem. It says how good we can approximate a $C^{n+1}([a, b])$-function by a polynomial of degree n.

Corollary 19.4. *Let $f \in C^{n+1}([a, b])$. Then there exists a constant $C = C(f) > 0$ such that for any $x_0 \in [a, b]$ the*

$$|f(x) - P_n(x)| \leq C\, |x - x_0|^{n+1}.$$

The result of Corollary 19.4 is often written in the so-called O or o-notation.

Here we write a quantity as $O(t)$ (we usually don't care about the precise formulation) if said quantity satisfies such that $\limsup_{t \to 0^+} \frac{|O(t)|}{t} < \infty$. In this sense Corollary 19.4 can be written as: whenever $f \in C^{n+1}([a, b])$ then

$$f(x) = P_n(x) + O(|x - x_0|^{n+1}) \quad \text{as } x \to x_0.$$

A quantity is noted as $o(t)$ if $\limsup_{t \to 0^+} \frac{|O(t)|}{t} = 0$. Actually one can show that whenever $f \in C^{n+1}([a, b])$ then

$$f(x) = P_{n+1}(x) + o(|x - x_0|^{n+1}) \quad \text{as } x \to x_0.$$

Proof. We reprove Taylor's theorem, using the fundamental theorem of calculus, Theorem 15.2. If $f \in C^1$ then

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(z_1)\, dz_1$$

Now we make a trick

$$\int_{x_0}^{x} f'(z)dz = f'(x_0)(x - x_0) + \int_{x_0}^{x} (f'(z_1) - f'(x_0))\, dz_1$$

We can use the the fundamental theorem again, for $f'(z_1) - f'(x_0)$ (if $f \in C^2$), and we get

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \int_{x_0}^{x} \int_{x_0}^{z_1} f''(z_2) dz_2\, dz_1$$

Again we do our trick (with a short calculation)

$$\int_{x_0}^{x} \int_{x_0}^{z_1} f''(z_2) dz_2\, dz_1 = f''(x_0) \frac{1}{2}(x - x_0)^2 + \int_{x_0}^{x} \int_{x_0}^{z_1} f''(z_2) - f''(x_0) dz_2\, dz_1$$

Repeating this $n + 1$ times, we have

$$f(x) - P_{n+1}(x) = \int_{x_0}^{x} \int_{x_0}^{z_1} \ldots \int_{x_0}^{z_{n+1}} \left(f^{(n+1)}(z_{n+1}) - f^{n+1}(x_0)\right) dz_{n+1}\, dz_n \ldots dz_1$$

Next, if $f \in C^{n+1}$ then for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|\left(f^{(n+1)}(z) - f^{n+1}(x_0)\right)| < \varepsilon \quad \forall z : |z - x_0| < \delta.$$

Observe that if $x \in \mathbb{R}$ with $|x - x_0| < \delta$ then for all z_i in the integral above, $z_i \in [x, x_0]$. So we obtain for all x such that $|x - x_0| < \delta$,

$$|f(x) - P_{n+1}(x)| \leq |x_0 - x|^{n+1}\varepsilon.$$

Which is what we wanted. □

Remark 19.5. With the argument above, one can also obtain another representation of the remainder term in Taylor's theorem, namely

$$R_n(x) = \int_a^x \frac{f^{n+1}(t)}{n!}(x-t)^n\, dt.$$

The additional technique is the interchange of integrals, called *Fubini's theorem*, which we shall not prove here.

Example 19.6. (1) For $f(x) = e^x$ we have $f^{(k)}(0) = e^0 = 1$. So for $c = 0$ we by get Taylor's theorem

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + o(x^n).$$

(2) For $f(x) = \sin(x)$ we have $f^{(k)}(0) = 0$ for k even and $f^{(k)} = (-1)^k$ for k odd. So for $c = 0$ we by get Taylor's theorem

$$\sin(x) = \sum_{k=1}^{n} (-1)^{k-1} \frac{x^{2k-1}}{(2k-1)!} + o(x^{2n-1}).$$

Cf. Figure 19.3.

Now that we have Taylor's theorem, if $f \in C^\infty([a,b])$ we can formally write

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

However, as we have learned in Section 18 the right-hand side has no reason to converge to anything! Functions for which the right-hand side converges (with a convergence radius > 0) around a point x_0 are called *analytic*. These are essentially infinite polynomials, which have many fascinating properties (For more on analytic functions see [KP02]) They appear again e.g. in complex analysis with holomorphic functions which are analytic everywhere.

Example 19.7.

- Assume that f is analytic in some $(x_0 - R, x_0 + R)$, i.e.

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

Then for any $c \in (x_0 - R, x_0 + R)$ we have

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!}(x - c)^k.$$

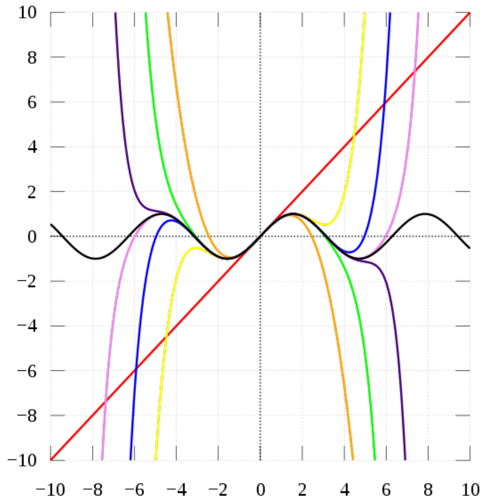

FIGURE 19.3. As the degree of the Taylor polynomial rises, it approaches the correct function. This image shows $\sin x$ and its Taylor approximations, polynomials of degree 1, 3, 5, 7, 9, 11 and 13. Source: Ikamusume-Fan/Wikipedia

From this, in turn, we can obtain that f has at most finitely many roots in $[x_0 - r, x_0 + r]$ for any $r < R$ or it is constant.

Indeed, assume not then there are countably many roots $(c_i)_{i=1}^N$ which have a limit c, and by continuity $c \in [x_0 - r, x_0 + r]$ and $f(c) = 0$ so

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(c)}{k!}(x-c)^k.$$

Now, if $f \not\equiv 0$ there must be the first k_0 such that $f^{(k_0)}(c) \neq 0$. Then we have

$$f(x) = \sum_{k=k_0}^{\infty} \frac{f^{(k)}(c)}{k!}(x-c)^k = (x-c)^{k_0} \underbrace{\sum_{k=k_0}^{\infty} \frac{f^{(k)}(c)}{k!}(x-c)^{k-k_0}}_{\neq 0 \text{ in } x=c}$$

But this representation of f shows that $f(x) \neq 0$ for $x \neq c$ and $x \approx c$ (since $|x-c| \ll 1$ for $x \approx c$). But then c cannot be the limit of values c_i with $f(c_i) = 0$.

- The above shows that the function from Example 13.2

$$f(x) = \begin{cases} 0 & |x| \geq 1 \\ e^{\frac{1}{x^2-1}} & |x| < 1. \end{cases}$$

cannot be analytic around $x = \pm 1$.

FIGURE 20.1. Émile Picard, 1856-1941. French mathematician.

FIGURE 20.2. Émile Picard, 1870-1946. Finnish mathematician. Not related to Jean-Luc.

20. Picard-Lindelöf or Cauchy-Lipschitz Theorem for ordinary differential equations

The Picard-Lindelöf or Cauchy-Lipschitz theorem is the fundamental theorem of ordinary differential equations. It is a consequence of the Banach Fixed Point theorem, Section 6, although we give a direct proof here based on what is called the Picard iteration.

Theorem 20.1. *Let $[t_0, t_1]$ and $[p_0, p_1]$ be two intervals in $\mathbb{R}$. Suppose that*

$$F : [t_0, t_1] \times [p_0, p_1] \to \mathbb{R}$$

- *is continuous in both variables, i.e.*

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : |F(s,x) - F(t,y)| < \varepsilon \quad \forall s,t \in [t_0,t_1], \quad x,y \in [p_0,p_1] : \quad |s-t|+|x-y| < \delta.$$

- *and Lipschitz continuous in the second variable, that is there exists $L \in \mathbb{R}$ such that*

$$|F(s,x) - F(s,y)| \leq L|x-y| \quad \text{for all } s \in [t_0,t_1], \quad x,y \in [p_0,p_1]. \tag{20.1}$$

Fix $\bar{t} \in (t_0, t_1)$ and $\bar{x} \in (p_0, p_1)$. There exists a small $\theta > 0$ such that

(1) there is $f : (\bar{t} - \theta, \bar{t} + \theta) \to \mathbb{R}$ differentiable that solves the Initial Value Problem

$$\begin{cases} f'(t) = F(t, f(t)) & x \in (\bar{t} - \theta, \bar{t} + \theta) \\ f(\bar{t}) = \bar{x}. \end{cases} \tag{20.2}$$

(2) *The solution is unique, i.e. if there exists a differentiable* $\tilde{f} : (\bar{t} - \theta, \bar{t} + \theta) \to \mathbb{R}$ *satisfying*

$$\begin{cases} \tilde{f}'(t) = F(t, \tilde{f}(t)) & t \in (\bar{t} - \theta, \bar{t} + \theta) \\ \tilde{f}(\bar{t}) = \bar{x}. \end{cases}$$

Then $\tilde{f}(t) = f(t)$ *for all* $t \in (\bar{t} - h, \bar{t} + \theta)$.

Our first step is the realization that the ODE (20.2) is equivalent to an integral equation

Lemma 20.2. *Assume* $f : (\bar{t} - \theta, \bar{t} + \theta) \to \mathbb{R}$ *is continuous. Then the following are equivalent*

(1) f *is differentiable, and* f *solves* (20.2)
(2) f *solves*

$$f(t) = \bar{x} + \int_{\bar{t}}^{t} F(s, f(s))\, ds \quad \forall t \in (\bar{t} - \theta, \bar{t} + \theta). \tag{20.3}$$

Proof. Observe that by assumption $s \mapsto F(s, f(s))$ is continuous and thus Riemann integrable. Now the claim is a consequence of the Fundamental Theorem of Calculus Theorem 15.2 and Theorem 15.3. □

For this integral equation we now prove existence and uniqueness using *Picard iterates*:

Proof of existence. We define the Picard iterates:

$$f_0(t) := \bar{x},$$

and

$$f_k(t) := \bar{x} + \int_{\bar{t}}^{t} F(s, f_{k-1}(s)) ds. \tag{20.4}$$

We do so on some closed interval $[\bar{t} - \theta, \bar{t} + \theta]$, where θ will be chosen later (but it is so small that $[\bar{t} - \theta, \bar{t} + \theta] \subset (t_0, t_1)$.

Exercise 20.3. *Assume* F *is as in Theorem 20.1. Show that*

$$\sup_{s \in [t_0, t_1], x \in [p_0, p_1]} |F(s, x)| < \infty$$

When defining f_k in (20.4) we have implicitely assumed that $f_{k-1}([\bar{t} - \theta, \bar{t} + \theta]) \subset [p_0, p_1]$ (so that $F(s, f_{k-1}(s))$ is defined). This is justified by the following inductive argument. Assume $f_{k-1}([\bar{t} - \theta, \bar{t} + \theta]) \subset [p_0, p_1]$. Then f_k is well-defined in (20.4) and we have

$$|f_k(t) - \bar{x}| \leq |t - \bar{t}| \sup_{s \in [t_0, t_1], x \in [p_0, p_1]} |F(s, x)| \leq \theta \sup_{s \in [t_0, t_1], x \in [p_0, p_1]} |F(s, x)|.$$

Recall that $\bar{x} \in (p_0, p_1)$. Using Exercise 20.3 by choosing θ small enough (but independent of k!) we ensure that $f_k(t) \in (p_0, p_1)$.

Now we see that f_k is continuous, and thus uniformly continuous and in particular bounded on $[\bar{t}-\theta, \bar{t}+\theta]$. We also observe

$$\begin{aligned}|f_k(t) - f_{k-1}(t)| &\leq \int_{\bar{t}}^{t} \left(F(s, f_{k-1}(s)) - F(s, f_{k-2}(s))\right) ds \\ &\overset{(20.1)}{\leq} \underbrace{|t-\bar{t}|}_{\leq \theta} L \sup_{s\in[\bar{t}-\theta,\bar{t}+\theta]} |f_{k-1}(s) - f_{k-2}(s)|\end{aligned} \tag{20.5}$$

Thus,

$$\sup_{t\in[t_0-\theta,t_0+\theta]} |f_k(t) - f_{k-1}(t)| \leq \theta L \sup_{t\in[t_0-\theta,t_0+\theta]} |f_{k-1}(t) - f_{k-2}(t)|$$

Recall that in (16.2) we called this the L^∞-norm, so we could write equivalently

$$\|f_k - f_{k-1}\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])} \leq \theta L \|f_{k-1} - f_{k-2}\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])}$$

Now we choose $\theta > 0$ possibly even smaller such that $\theta L < \frac{1}{2}$. Then we have

$$\|f_k - f_{k-1}\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])} \leq \frac{1}{2} \|f_{k-1} - f_{k-2}\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])}$$

We can now argue similiar to the contraction principle, Exercise 4.10: Namely by iteration we find from the above

$$\|f_k - f_{k-1}\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])} \leq 2^{1-k} \|f_1 - f_0\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])} \quad \forall k \geq 1.$$

Then for $\ell > k$ we have

$$\begin{aligned}\|f_k - f_\ell\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])} &\leq \sum_{i=k+1}^{\ell} \|f_i - f_{i-1}\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])} \\ &\leq \sum_{i=k+1}^{\ell} 2^{1-i} \|f_1 - f_0\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])} \\ &\leq \left(\sum_{i=k+1}^{\infty} 2^{1-i}\right) \|f_1 - f_0\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])} \\ &\leq 2^{-k} 2 \, \|f_1 - f_0\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])}.\end{aligned}$$

Reversing the role of ℓ and k we find that for any $N \geq 1$

$$\|f_k - f_\ell\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])} \leq 2^{-N} 2 \, \|f_1 - f_0\|_{L^\infty([\bar{t}-\delta,\bar{t}+\delta])} \quad \forall k, \ell \geq N$$

This says that $(f_k)_{k\in\mathbb{N}}$ is a Cauchy sequence in L^∞, and by Theorem 16.9 and Proposition 16.17, there exists $f : [\bar{t}-\theta, \bar{t}+\theta]$ such that f_k converges uniformly to f in $[\bar{t}-\theta, \bar{t}+\theta]$.

But then in particular, $F(s, f_{k-1}(s))$ converges uniformly to $F(s, f(s))$ in $[\bar{t}-\theta, \bar{t}+\theta]$, Exercise 16.14. Using also that uniform convergence implies convergence of the integral, Proposition 16.19, we have for any $t \in [t-\theta, t+\theta]$

$$f(t) = \lim_{k\to\infty} f_k(t) \overset{(20.4)}{=} \lim_{k\to\infty} \left(\bar{x} + \int_{\bar{t}}^{t} F(s, f_{k-1}(s)) ds\right) = \bar{x} + \int_{\bar{t}}^{t} F(s, f(s)) ds$$

That is

$$f(t) = \int_{\bar{t}}^{t} F(s, f(s))ds$$

and by Lemma 20.2 we have found that f solves (20.2) in $(\bar{t} - \theta, \bar{t} + \theta)$. □

Proof of uniqueness. As for uniqueness, assume that we have f and $\tilde{f} : (\bar{t} - \theta, \bar{t} + \theta)$ are differentiable and solve (20.2).

Let $\mu < \theta$ be arbitary.

Then we have in view of (20.2),

$$f(t) = \int_{\bar{t}}^{t} F(s, f(s))ds, \quad \tilde{f}(t) = \int_{\bar{t}}^{t} F(s, \tilde{f}(s))ds \quad \forall t \in [\bar{t} - \mu, \bar{t} + \mu]$$

Thus, arguing as in (20.5) for any $t \in [\bar{t} - \mu, \bar{t} + \mu]$

$$|f(t) - \tilde{f}(t)| \leq \int_{\bar{t}}^{t} \left(F(s, f(s)) - F(s, \tilde{f}(s))\right) ds \overset{(20.1)}{\leq} \mu L \sup_{s \in [\bar{t} - \mu, \bar{t} + \mu]} |f_{k-1}(s) - f_{k-2}(s)|$$

Taking the supremum over all $t \in [\bar{t} - \mu, \bar{t} + \mu]$ we find

$$\|f - \tilde{f}\|_{L^\infty([\bar{t}-\mu,\bar{t}+\mu])} \leq \mu L \|f - \tilde{f}\|_{L^\infty([\bar{t}-\mu,\bar{t}+\mu])}.$$

But recall that we have $\mu < \theta$ and we assumed $\theta L < \frac{1}{2}$, so we find

$$\|f - \tilde{f}\|_{L^\infty([\bar{t}-\mu,\bar{t}+\mu])} \leq \frac{1}{2} \|f - \tilde{f}\|_{L^\infty([\bar{t}-\mu,\bar{t}+\mu])}.$$

But then $\|f - \tilde{f}\|_{L^\infty([\bar{t}-\mu,\bar{t}+\mu])} = 0$, and thus $f = \tilde{f}$ in $[\bar{t} - \mu, \bar{t} + \mu]$. Since this holds for any $\mu < \theta$ we conclude that $f = \tilde{f}$ in $(\bar{t} - \theta, \bar{t} + \theta)$. □

Email address: armin@pitt.edu