ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 1

LÉONARD GUETTA

The reference for this lecture is the first chapter of [Rud76].

1. Introduction

We denote by $\mathbb{N}$ the set of *natural numbers* $\{0, 1, 2, ...\}$, $\mathbb{Z}$ the set of *integers* $\{..., -2, -1, 0, 1, 2, ...\}$ and finally $\mathbb{Q}$ the set of *rational numbers*.

As a system of numbers, $\mathbb{Q}$ has a lot of satisfactory properties, for example it is and ordered field (a notion we'll define later) and *dense*. The latter means that for two rational numbers $p < q$, there exists a rational number r such that $p < r < q$. Yet, certain numbers seem to be "missing".

Proposition 1.1. *There is no rational number x such that $x^2 = 2$.*

Proof. Exercise. □

In other words, 2 does not have a square root in $\mathbb{Q}$. However, we can approximate it "as close as possible" by rational numbers. Meaning for example that there is sequence of rational numbers

$$1.4, 1.41, 1.412, 1.4124, ...$$

which "converges to $\sqrt{2}$", even though one would need to define precisely what this means.

Intuitively speaking, $\mathbb{Q}$ have gaps and lacks the property of being a "continuum". This defect is corrected by introducing real numbers.

2. Ordered sets

Definition 2.1. An *ordered set* is a set S equipped with a binary relation $\leq$ such that for any elements x, y, z of S, the following axioms are satisfied:

- at least one of the following
$$x \leq y \text{ or } y \leq x$$
holds,
- if $x \leq y$ and $y \leq x$, then $x = y$,
- if $x \leq y$ and $y \leq z$, then $x \leq z$.

Remark 2.2. As usual, we write $x < y$ to say that $x \leq y$ and $x \neq y$. It follows that for any elements x and y in an ordered set, one, and exactly one of the following is true

$$x < y \text{ or } x = y \text{ or } x > y.$$

Example 2.3. The sets $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$, equipped with their usual order relation are ordered sets.

The sets of words in English, equipped with the lexicographical order (the "dictionnary" order) is an ordered set.

Definition 2.4. Let $(S, \leq)$ be an ordered set and $A \subset S$ a subset of S. An element $x \in S$ is an *upper bound* of A if for every element a of A, we have

$$a \leq x.$$

If such an upper bound exists, we say that A is *bounded above.*

An upper bound x of A is a *least upper bound* (or *supremum*) if for any other upper bound x' of A, we have

$$x \leq x'.$$

Proposition 2.5. *If a subset $A \subseteq S$ of an ordered set admits a least upper bound, then it is unique.*

Proof. Let x and x' be least upper bounds of $A \subset S$. Because x is a least upper bound, we have

$$x \leq x',$$

and because x' is a least upper bound, we have

$$x' \leq x.$$

It follows then from the axioms of ordered set that $x = x'$. □

As a consequence of the previous proposition, we can speak of *the* least upper bound of $A \subseteq S$ (if it exists). We use the notation $\sup A$ for the least upper bound.

Remark 2.6. Even if it exists, the least upper bound of $A \subseteq S$ is *not* necessarily an element of A. For example, if $S = \mathbb{Q}$ with the usual order and $A = \{x \in \mathbb{Q} \mid x < 0\}$, then $\sup A = 0$, but $0 \notin A$.

Definition 2.7. We say that an ordered set $(S, \leq)$ has the *least upper bound property* if every **non-empty**, bounded above, subset A of S has a least upper bound.

Proposition 2.8. *The ordered set $\mathbb{Q}$ does not have the least upper bound property.*

Proof. Consider the the subset

$$A = \{x \in \mathbb{Q} \mid x^2 < 2\}$$

of $\mathbb{Q}$. This set is not empty (for example, $0 \in \mathbb{Q}$), it is bounded above (for example, by 2). Let us prove by contradiction that it does not have a least upper bound.

Suppose that $\sup A$ exists, which we denote p for short. Let us prove that this necessarily implies that $p^2 = 2$, which would be a contradiction (since we already saw that such a rational number cannot exist). Let q be the rational number

$$q = p - \frac{p^2 - 2}{p + 2} = \frac{2p + 2}{p + 2}.$$

Then

$$q^2 = 2 + \frac{2(p^2 - 2)}{(p + 2)^2}.$$

If $p^2 < 2$, then $q > p$ and $q^2 < 2$, which contradicts the fact that p is an upper bound of A. If $p^2 > 2$, then $0 < q < p$ and $q^2 > 2$, which contradicts the fact that p is the *least* upper bound of A. Hence, we can deduce that $p^2 = 2$. $\square$

Remark 2.9. Note that by reversing all inequalities, one define the notion of *lower bound* and *greatest lower bound*, as well as *the greatest lower bound property*. (Exercise). One has then the following result.

Proposition 2.10. *An ordered set has the least upper bound property if and only if it has the greatest upper bound property.*

Proof. Exercise. $\square$

3. Real numbers

In order to state the existence theorem of the real numbers, we need to introduce the concepts of field and ordered field.

Definition 3.1. A *field* is a set F equipped with two (binary) operations $+$ and $\cdot$, and two distinguished element $0 \in F$ and $1 \in F$, such that for any element x, y, z in F, the following axioms are satisfied:

- $(x + y) + z = x + (y + z)$,
- $x + y = y + x$,
- $(x \cdot y) \cdot z = x \cdot (y \cdot z)$,
- $x \cdot y = y \cdot x$,
- $x \cdot (y + z) = x \cdot y + x \cdot z$,
- $0 \neq 1$,
- $x + 0 = 0 + x = x$,
- $1 \cdot x = x \cdot 1 = x$,
- $0 \cdot x = x \cdot 0 = 0$,
- there exists an element -1 of F such that
$$1 + (-1) = 0,$$
- if $x \neq 0$, there exists an element $\frac{1}{x}$ in F such that
$$x \cdot \frac{1}{x} = 1.$$

Remark 3.2. In a field, we usually write xy, $\frac{x}{y}$, $-x$, $x - y$ instead of $x \cdot y$, $x \cdot \frac{1}{y}$, $(-1) \cdot x$, $x + (-y)$.

Even though this is not the main focus of this course, using the previous field axioms, one can prove a lot of the usual algebraic properties of $\mathbb{Q}$ holds in any field (see for example 1.12-1.18 in [Rud76]).

Definition 3.3. An *ordered field* is a field F, equipped with a binary relation $\leq$ that makes it an ordered set, and such that the following additional axioms are satisfied:

- if $x \leq y$, then $x + z \leq y + z$,
- if $0 \leq x$ and $0 \leq y$, then $0 \leq x \cdot y$.

Elements x such that $0 < x$ (resp. $x < 0$) are called *positive* (resp. *negative*).

We can now state the main result of this lecture.

Theorem 3.4. *There exists a unique ordered field, which we denote $\mathbb{R}$, which contains $\mathbb{Q}$ as a subfield, and that satisfies the least upper bound property.*

Proof. Admitted. See appendix of first chapter of [Rud76]. □

This theorem is **fundamental**, and is the basis of all of analysis. In fact, we do not really need to know an explicit construction of the real numbers to do analysis, and can simply work abstractly with the real numbers, only using that it satisfies the previous theorem. In other words, the previous theorem could be considered as an axiom of real numbers.

Proposition 3.5. *For every real number $x > 0$ and every integer $n > 0$, there is one and only one real number y such that $y^n = x$.*

Proof. See Theorem 1.21 of [Rud76]. □

In particular, we deduce from the previous proposition that there exists a real number $\sqrt{2}$ such that $\sqrt{2}^2 = 2$.

We end this lecture with two useful properties of real numbers.

Proposition 3.6. *Let x, y be two real numbers.*

(1) *If $x > 0$, there exists a natural number $n \in \mathbb{N}$ such that $y < nx$,*
(2) *If $x < y$, there exists a rational number r such that $x < r < y$.*

We refer to the first property by saying that $\mathbb{R}$ is *archimedean* and to the second property by saying that $\mathbb{Q}$ is *dense in* $\mathbb{R}$.

Proof. See Theorem 1.20 of [Rud76]. □

References

[Rud76] Walter Rudin. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1976.

Email address: l.s.guetta@uu.nl

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht

ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 2

LÉONARD GUETTA

The reference for this lecture is the third chapter of [Rud76]*, from 3.1. to 3.20., as well as Marta Pieropan's lecture notes.*

1. Convergent real sequences

Recall that a sequence $(u_n)_{n\geq n_0}$ of real numbers consists of the data of a a list of real numbers u_{n_0}, u_{n_0+1}, u_{n_0+2}, $\cdots$, indexed by natural integer $n \geq n_0$. Usually, we will have $n_0 = 0$ or $n_0 = 1$. Often we shall consider that n_0 is implicit and just write (u_n).

Definition 1.1. Let x and y be two real numbers. The *distance between x and y*, denoted by $d(x, y)$, is defined as

$$d(x, y) = |y - x|.$$

Proposition 1.2. *Let x, y, z be real numbers. The following properties holds:*

- $d(x, y) = d(y, x)$, *(Symmetry)*
- $d(x, y) = 0$ *if and only if* $x = y$, *(Separation)*
- $d(x, y) + d(y, z) \leq d(x, z)$. *(Triangle inequality)*

Proof. Exercise. □

Definition 1.3. Let (u_n) be a sequence in $\mathbb{R}$ and ℓ a real number. We say that the sequence (u_n) *converges to ℓ* if the following property holds:

For every $\epsilon > 0$, there exists an $N \in \mathbb{N}$, such that if $n \geq N$ then $d(u_n, \ell) < \epsilon$.

Intuitively, this means that u_n gets as close as we want to ℓ for n large enough.

We say that a sequence (u_n) is *convergent* or that *its limit exists* if there exists an $\ell \in \mathbb{R}$ such that (u_n) converges to ℓ, in which case we use the notation

$$\lim_{n\to+\infty} u_n = \ell.$$

When such a limit does not exist, we say that the sequence is *divergent.*

As the following result shows, *when it exists*, the limit of a sequence is unique, hence we can speak of "the" limit of a sequence.

Proposition 1.4. *Let (u_n) be a sequence in $\mathbb{R}$, and ℓ and ℓ' two real numbers. If (u_n) converges to ℓ and to ℓ', then $\ell = \ell'$.*

In order to prove this proposition, we need the following lemma.

Lemma 1.5. *Let $x \geq 0$ be a real number. If for every real number $\epsilon > 0$, we have*

$$0 \leq x < \epsilon,$$

then $x = 0$.

Proof. The hypothesis means that x is a lower bound of $\mathbb{R}_{>0} = \{\epsilon > 0 \,|\, \epsilon \in \mathbb{R}\}$. In particular $x \leq 0$, because 0 is the greatest lower bound of this subset. Since $x \geq 0$ by hypothesis, we deduce that $x = 0$. □

Proof of Proposition 1.4. Let $\epsilon > 0$. Since (u_n) converges to ℓ, there exists $N \in \mathbb{N}$ such that

$$d(u_n, \ell) < \epsilon/2$$

for all $n \geq N$. Similarly, since (u_n) converges to ℓ', there exists $N' \in \mathbb{N}$ such that

$$d(u_n, \ell') < \epsilon/2$$

for all $n \geq N'$. In particular, for all $n \geq \max(N, N')$ we have

$$d(u_n, \ell) + d(u_n, \ell') < \epsilon/2 + \epsilon/2 = \epsilon.$$

By symmetry and triangle inequality, we can deduce that for any $n \geq \max(N, N')$, we have

$$d(\ell, \ell') < d(\ell, u_n) + d(u_n, \ell) < \epsilon.$$

Hence, we have proved that for any $\epsilon > 0$, we have

$$0 \leq d(\ell, \ell') < \epsilon,$$

from which we conclude that $d(\ell, \ell') = 0$ by Lemma 1.5, and it follows by the separation property of d that $\ell = \ell'$. □

We have the following useful properties of sequences of real numbers.

Proposition 1.6. *Let (u_n) and (v_n) be* convergent *sequences in $\mathbb{R}$. The following properties hold:*

1. *if $u_n \leq v_n$ for all n, then*
$$\lim_{n\to\infty} u_n \leq \lim_{n\to+\infty} v_n,$$
2. *the sequence $(u_n + v_n)$ is convergent and we have*
$$\lim_{n\to+\infty} (u_n + v_n) = \lim_{n\to+\infty} u_n + \lim_{n\to+\infty} v_n,$$
3. *for c is a real number, then the sequence $(c \cdot u_n)$ is convergent and we have*
$$\lim_{n\to+\infty} c \cdot u_n = c \cdot \lim_{n\to+\infty} u_n$$
4. *the sequence $(u_n \cdot v_n)$ is convergent and we have*
$$\lim_{n\to+\infty} (u_n \cdot v_n) = \lim_{n\to+\infty} u_n \cdot \lim_{n\to+\infty} v_n,$$
5. *if $v_n \neq 0$ for all n and $\lim_{n\to+\infty} v_n \neq 0$, then the sequence $(\frac{u_n}{v_n})$ is convergent and*
$$\lim_{n\to+\infty} \frac{u_n}{v_n} = \frac{\lim_{n\to+\infty} u_n}{\lim_{n\to+\infty} v_n}$$

Proof. Theorem 3.3 of [Rud76]. □

Proposition 1.7 ("Sandwich" lemma)**.** *Let $(u_n), (v_n)$ and (w_n) be real sequences such that*

$$u_n \leq v_n \leq w_n$$

for all n. If u_n and w_n both converges to the same limit ℓ, then v_n also converges to ℓ.

Proof. Let $\epsilon > 0$. We know there exists and N such that for all $n \geq N$, we have

$$\ell - \epsilon < u_n < \ell + \epsilon,$$

and we know there exists an N' such that for all $n \leq N$, we have

$$\ell - \epsilon \leq w_n \leq \ell + \epsilon.$$

In particular, for all $n \geq \max(N, N')$, we have

$$\ell - \epsilon < u_n \leq v_n \leq w_n < \ell + \epsilon,$$

which proves, by definition, that (v_n) converges to ℓ. □

2. Criteria for convergence of real sequences

Definition 2.1. Let (u_n) be a sequence of real numbers. We say that (u_n) is:

- *increasing* if for all n, we have $u_n \leq u_{n+1}$,
- *bounded above* if there exists an $\alpha \in \mathbb{R}$, such that $u_n \leq \alpha$ for all n,
- *decreasing* if for all n, we have $u_n \geq u_{n+1}$,
- *bounded below* if there exists an $\alpha \in \mathbb{R}$, such that $u_n \geq \alpha$ for all n.

We say that a sequence is *bounded* if it is both bounded above and bounded below.

Lemma 2.2. *A convergent sequence is bounded.*

Proof. Exercise. □

Theorem 2.3. *Let (u_n) be a sequence of real numbers. If (u_n) is increasing and bounded above,then it is convergent. If (u_n) is decreasing and bounded below, then it is convergent.*

Proof. We only prove the first statement and leave the second one as an exercise.

Consider the following subset of $\mathbb{R}$

$$A = \{u_n | n \geq n_0\}.$$

It is non-empty (because it contains all the values of u_n) and it is bounded above by hypothesis. In particular, since $\mathbb{R}$ has the least upper bound property, $\sup A$ exists. Let us show that (u_n) converges to $\sup A$.

Let $\epsilon > 0$. Because $\sup A$ is the *least* upper bound, then $\sup A - \epsilon$ is not an upper bound of A. This means that there exists an $N \in \mathbb{N}$ such that

$$\sup A - \epsilon < u_N \leq \sup A.$$

Since u_n is increasing, for all $n \geq N$, we have

$$\sup A - \epsilon < u_n,$$

and since $\sup A$ is an upper bound of A, for all $n \geq N$, we have

$$\sup A - \epsilon < u_n \leq \sup A,$$

and, *a fortiori*, we have

$$\sup A - \epsilon < u_n < \sup A + \epsilon,$$

which means exactly that

$$|u_n - \sup A| < \epsilon. \qquad \square$$

Definition 2.4. Let (u_n) be a *bounded* sequence of real numbers. Because this sequence is bounded, we can define sequences (v_n) and (w_n) as

$$v_n = \sup\{u_k \,|\, k \geq n\}$$

and

$$w_n = \inf\{u_k \,|\, k \geq n\}$$

The sequence (v_n) is decreasing and bounded below (why?), hence has a limit by Theorem 2.3. Dually, the sequence (w_n) is increasing and bounded above, hence has a limit. We then define

$$\limsup_{n\to+\infty} u_n := \lim_{n\to\infty} v_n$$

and

$$\liminf_{n\to+\infty} u_n := \lim_{n\to\infty} w_n.$$

Proposition 2.5. *Let (u_n) be a sequence of real numbers. Then (u_n) is convergent if and only if (u_n) is bounded and we have*

$$\limsup_{n\to\infty} u_n = \liminf_{n\to\infty} u_n,$$

in which case the limit of (u_n) is this common value.

Proof. Let's begin with the "if" part. Suppose that (u_n) is bounded, hence we can define the sequences $v_n = \sup\{u_k \,|k \geq n\}$ and $w_n = \inf\{u_k \,|k \geq n\}$ as before.

Notice now that for all n we have

$$w_n \leq u_n \leq v_n,$$

and the conclusion follows from the Sandwich Lemma (Proposition 1.7).

Now, the "only if" part. Let (u_n) be a convergent sequence of real numbers and let ℓ be its limit. We have already seen (Lemma 2.2) that a convergent sequence is bounded. Let's prove that

$$\limsup_{n\to\infty} u_n = \ell = \liminf_{n\to\infty} u_n.$$

Let $\epsilon > 0$. There exists an $N \in \mathbb{N}$ such that for all $n \geq N$, we have $d(u_n, \ell) < \epsilon$, which means exactly that

$$\ell - \epsilon < u_n < \ell + \epsilon,$$

which implies clearly that

$$\ell - \epsilon \leq \inf\{u_k \,|k \geq n\} < \ell + \epsilon$$

and

$$\ell - \epsilon < \sup\{u_k \,|k \geq n\} \leq \ell + \epsilon,$$

which concludes the proof. $\square$

3. Divergence to ∞

It is sometimes useful to consider sequences that diverges to ∞.

Definition 3.1. A real sequence (u_n) *diverges to* $+\infty$ if for all $M \in \mathbb{R}$, there exists an $N \in \mathbb{N}$, such that for all $n \geq N$, we have $u_n \geq M$. We then write

$$\lim_{n \to +\infty} u_n = +\infty.$$

Similarly, (u_n) is said to *diverges to* $-\infty$ if for all $M \in \mathbb{R}$, there exists an $N \in \mathbb{N}$, such that for all $n \geq N$, we have $u_n \leq M$. We then write

$$\lim_{n \to +\infty} u_n = -\infty.$$

Proposition 3.2. *An unbounded increasing real sequence diverges to* $+\infty$. *An unbounded decreasing real sequence diverges to* $-\infty$.

Proof. Exercise. □

In particular, given a real sequence (u_n), the limits

$$\limsup_{n \to +\infty} u_n \text{ and } \liminf_{n \to +\infty} u_n$$

are always defined (with possible values $+/-\infty$ if the sequence is unbounded), and we obtain the slightly improved version of Proposition 2.5.

Proposition 3.3. *Let* (u_n) *be a real sequence. Then* $\lim_{n \to +\infty} u_n$ *exists (with possible value* $+/-\infty$*) if and only if*

$$\limsup_{n \to +\infty} u_n = \liminf_{n \to +\infty} u_n,$$

in which case $\lim_{n \to +\infty} u_n$ *is this common value.*

References

[Rud76] Walter Rudin. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1976.

Email address: l.s.guetta@uu.nl

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht

ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 3

LÉONARD GUETTA

Recall that given a (real) sequence $(a_n)_{n\geq n_0}$, we denote by $\sum a_n$ the sequence $(s_k)_{k\geq n_0}$ of its partial sum:

$$s_k = \sum_{n=n_0}^{k} a_n$$

We say that the series $\sum a_n$ converges (resp. diverges) if this sequence converges (resp. diverges). When it converges to $\alpha \in \mathbb{R}$, we use the notation

$$\sum_{n=n_0}^{+\infty} a_n = \alpha$$

The following result is left as an exception.

Proposition 0.1. *If the series $\sum a_n$ is convergent, then $\lim_{n\to+\infty} a_n = 0$.*

1. Series of non-negative terms

Theorem 1.1. *A series of non-negative terms converges if and only if the sequence of partial sums is bounded.*

Proof. The sequence of partial sum is increasing (because the terms are non-negative), hence it is convergent if and only if it is bounded. □

Proposition 1.2 (Comparison criterion)**.** *Let Σa_n, Σb_n series of non-negative terms such that*

$$a_n \leq b_n$$

for all $n \geq N_0$, with N_0 some fixed integer.

(1) If $\sum b_n$ converges, then so does $\sum a_n$.
(2) If $\sum a_n$ diverges, then so does $\sum b_n$.

Proof. Exercise. □

Our first example of series is the geometric one.

Proposition 1.3. *Consider the series $\sum x^n$ where x is a non-negative real number.*

(1) If $0 \leq x < 1$, then $\sum x^n$ converges and we have

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

(2) If $x \geq 1$, then $\sum x^n$ diverges.

Proof. If $x \neq 1$, then we have

$$\sum_{k=1}^{n} x^n = \frac{1 - x^{n+1}}{1 - x}.$$

If $x < 1$, this sequence converges to $\frac{1}{1-x}$ and if $x > 1$, this sequence diverges. If $x = 1$, the sequence $1 + 1 + 1 + 1 + \cdots$ diverges. $\square$

The following example is also very important.

Proposition 1.4. *Consider the series* $\sum \frac{1}{n^p}$.

(1) *If $p > 1$, the series converges.*
(2) *If $p \leq 1$, the series diverges.*

In order to prove this, we need the following lemma.

Lemma 1.5. *Let $a_1 \geq a_2 \geq a_3 \geq \cdots 0$ be a non-negative decreasing sequence. The series $\sum a_n$ converges if and only if the series*

$$\sum_{k=0}^{+\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \cdots$$

converges.

Proof. Let

$$s_n = a_1 + a_2 + a_3 + \cdots + a_n$$

and

$$t_k = a_1 + 2a_2 + 4a_4 + \cdots + 2^k a_{2^k}$$

be the sequences of partial sums.

For $n < 2^k$, we have

$$\begin{aligned} s_n &\leq a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1}) \\ &\leq a_1 + 2a_2 + 4a_4 + \cdots + 2^k a_{2^k} \\ &= t_k, \end{aligned}$$

and for $n > 2^k$,

$$\begin{aligned} s_n &\geq a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \cdots + (a_{2^{k-1}+1} + \cdots a_{2^k}) \\ &\geq a_1/2 + a_2 + 2a_4 + 4a_8 + \cdots + 2^{k-1} a_{2^k} \\ &= t_k/2. \end{aligned}$$

It follows that s_n is bounded (above) if and only if t_k is bounded (above). $\square$

Proof of Proposition 1.4. If $p \leq 0$, then the divergence follows from Proposition 0.1. If $p > 0$, we can use the previous lemma, and we consider the following series

$$\sum_{k=0}^{+\infty} 2^k \cdot \frac{1}{2^{kp}} = \sum_{k=0}^{+\infty} 2^{(1-k)p}.$$

Using the comparison with the geometric series (Proposition 1.3), we see that this series is convergent if and only if $2^{1-p} < 1$, which means exactly that $1 - p < 0$. $\square$

2. Convergence criteria for series

Definition 2.1. We say that series $\sum a_n$ *converges absolutely* if the series $\sum |a_n|$ converges.

Proposition 2.2. *If $\sum a_n$ converges absolutely, then it converges.*

Proof. Notice that we have

$$0 \leq a_n + |a_n| \leq 2|a_n|.$$

By hypothesis $\sum |a_n|$ converges, hence so does $\sum 2|a_n|$. By the comparision criterion, we deduce that $\sum(a_n + |a_n|)$ converges. Finally, notice that

$$(a_n + |a_n|) - |a_n| = a_n,$$

hence we deduce that $\sum a_n$ converges. □

Remark 2.3. The converse of this proposition is false. For example, the series $\sum \frac{(-1)^n}{n}$ converges, but the series $\sum \frac{1}{n}$ diverges.

Theorem 2.4 (Root test)**.** *Let $\sum a_n$ be a series and let $\alpha = \limsup_{n\to+\infty} \sqrt[n]{|a_n|}$.*
- *(a) If $\alpha < 1$, the series converges.*
- *(b) If $\alpha > 1$, the series diverges.*
- *(c) If $\alpha = 1$, we cannot conclude.*

In order to prove this theorem, we need the following lemma.

Lemma 2.5. *Let (a_n) be a bounded real sequence and $\alpha = \limsup_{n\to+\infty} a_n$. For every $\beta > \alpha$, there exists an $N \in \mathbb{N}$ such that $a_n < \beta$ for all $n \geq N$.*

Proof. Exercise. □

Proof of the theorem. If $\alpha < 1$, we can choose $\alpha < \beta < 1$ and the previous lemme shows that there exists an $N \in \mathbb{N}$ such that

$$\sqrt[n]{|a_n|} < \beta.$$

This means that if $n \geq N$, then

$$|a_n| < \beta^n.$$

Since $0 < \beta < 1$, (a) follows from the comparison criterion (Proposition 1.2) and the convergence of the geometric series (Proposition 1.3). □

Theorem 2.6 (Ratio test)**.** *Let $\sum a_n$ be a series.*
- *(a) If $\limsup_{n\to+\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$, then the series converges.*
- *(b) If $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$, for all $n \geq N_0$ with N_0 some fixed integer, then the series diverges.*

Email address: `l.s.guetta@uu.nl`

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht

ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 4

LÉONARD GUETTA

1. Cauchy sequences in $\mathbb{R}$

Definition 1.1. Let (a_n) be a real sequence. We say that it is a *Cauchy sequence* if for all $\epsilon > 0$, there exists an $N \geq 0$ such that if $n \geq N$ and $m \geq N$, then $d(a_n, a_m) < \epsilon$.

Theorem 1.2. *A sequence in $\mathbb{R}$ is convergent if and only if it is a Cauchy sequence.*

Proof. Let us start with the "only if" part. Suppose that (a_n) is a sequence converging to $\ell \in \mathbb{R}$ and let $\epsilon > 0$. There exists an $N \in \mathbb{N}$ such that if $n \geq N$, then $d(a_n, \ell) < \epsilon/2$. Using triangle inequality and symmetry, we deduce that

$$d(a_n, a_m) \leq d(a_n, \ell) + d(\ell, a_m) < \epsilon/2 + \epsilon/2 = \epsilon,$$

for all $n, m \geq N$.

Let us now prove the "if" part. Suppose that (a_n) is a Cauchy sequence and let $\epsilon > 0$. There exists an $N \in \mathbb{N}$ such that if $n, m \geq N$, then

$$d(a_n, a_m) < \epsilon.$$

In particular, if we choose $m = N$, for all $n \geq N$, we have

$$a_N - \epsilon < a_n < a_N + \epsilon.$$

Hence, we have

$$\sup\{a_n \,|\, n \geq N\} \leq a_N + \epsilon$$

from which we deduce that

$$\limsup_{n \to +\infty} a_n \leq a_N + \epsilon$$

(because the sequence $v_n = \sup\{a_k \,|\, k \geq n\}$ is decreasing.) Similarly, we have

$$\inf\{a_n \,|\, n \geq N\} \geq a_N - \epsilon,$$

from which we deduce that

$$\liminf_{n \to +\infty} a_n \geq a_N - \epsilon.$$

Recall now that

$$\liminf_{n \to +\infty} a_n \leq \limsup_{n \to +\infty} a_n,$$

and in particular, we deduce that

$$0 \leq \limsup_{n \to +\infty} a_n - \liminf_{n \to +\infty} a_n \leq 2\epsilon.$$

Since this is true for all $\epsilon > 0$, this proves that

$$\limsup_{n\to+\infty} a_n = \liminf_{n\to+\infty} a_n,$$

which proves that (a_n) is convergent. □

Remark 1.3. In the "if" part, we have implicitly used that $\limsup a_n$ and $\liminf a_n$ are real numbers (and not ∞) in order to be able to substract one to the other ($+\infty - \infty$ is not well-defined!). This is justified by the fact that every Cauchy sequence is bounded (Exercise).

We can use this theorem to give a criterion for convergence of series.

Proposition 1.4. *A real series $\sum a_n$ is convergent if and only if for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that if $n \geq m \geq N$, then*

$$\left| \sum_{k=m}^{n} a_k \right| < \epsilon.$$

2. Metric spaces

Definition 2.1. A metric space consists of a pair (X, d), where X is a set, and d is a function

$$d\colon X \times X \to \mathbb{R}_{\geq 0},$$

such that for all $x, y, z \in X$:

(1) $d(x, y) = d(y, x)$,
(2) $d(x, y) = 0$ if and only if $x = y$,
(3) $d(x, z) \leq d(x, y) + d(y, z)$.

A function that satisfies these three axioms is called a *metric* or a *distance*.

Example 2.2. The set of real numbers $\mathbb{R}$ equipped with the usual distance function

$$d\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{\geq 0}$$
$$(x, y) \mapsto |y - x|$$

is a metric space.

Example 2.3. If (X, d) is a metric space and $A \subset X$ is a subset of X, then (A, d) is a metric space on its own right (note that we abused notation here and wrote d for the restriction of d to A). For example, $\mathbb{Q}$ equipped with the distance $(x, y) \mapsto |y - x|$ is a metric space.

Our main examples of metric spaces are euclidian spaces.

Definition 2.4. Let $k \geq 0$. We denote by $\mathbb{R}^k$ the set of ordered k-uples

$$\mathbf{x} = (x_1, x_2, \cdots, x_k)$$

where each x_i is a real number. Elements of $\mathbb{R}^k$ are usually called *points* (or sometimes *vectors*) and the x_i's are called the *coordinates*. If $\mathbf{x}$ and $\mathbf{y}$ are points in $\mathbb{R}^k$ and α is a real number, we define

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \cdots, x_k + y_k)$$

and

$$\alpha \cdot \mathbf{x} = (\alpha x_1, \cdots, \alpha x_k).$$

These operations satisfy several axioms (commutativity and associativity of $+$, associativity of $\cdot$ and distributivity of $\cdot$ over $+$) that makes $\mathbb{R}^k$ a *real vector space*. Moreover, we also have the *inner product* (or *scalar product*) of $\mathbf{x}$ and $\mathbf{x}$ as

$$(\mathbf{x}|\mathbf{y}) = \sum_{i=1}^{k} x_i y_i.$$

and the (euclidian) norm of $\mathbf{x}$ as

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}|\mathbf{x})} = \sqrt{\sum_{i=1}^{k} x_i^2}.$$

Lemma 2.5. *Let* $\mathbf{x}, \mathbf{y}$ *be points in* $\mathbb{R}^k$ *and* $\alpha \in \mathbb{R}$. *Then*

(1) $\|\mathbf{x}\| \geq 0$,
(2) $\|\mathbf{x}\| = 0$ *if and only if* $\mathbf{x} = \mathbf{0}$,
(3) $\|\alpha \cdot \mathbf{x}\| = |\alpha| \, \|\mathbf{x}\|$,
(4) $|(\mathbf{x}|\mathbf{y})| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$,
(5) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Proof. The first three properties are obvious. For the fourth one, let t be a real number. We have

$$\begin{aligned} \|\mathbf{x} + t \cdot \mathbf{y}\|^2 &= (\mathbf{x} + t \cdot \mathbf{y} | \mathbf{x} + t \cdot \mathbf{y}) \\ &= \|\mathbf{x}\|^2 + t^2 \|\mathbf{y}\|^2 + 2t \, (\mathbf{x}|\mathbf{y}). \end{aligned}$$

Since $\|\mathbf{x} + t \cdot \mathbf{y}\|^2 \geq 0$, we deduce that $\|\mathbf{x}\|^2 + t^2 \|\mathbf{y}\|^2 + 2t \, (\mathbf{x}|\mathbf{y}) \geq 0$. By considering this as a polynomial of degree 2 in t, this implies that the discriminant

$$\Delta = 4 \, (\mathbf{x}|\mathbf{y})^2 - 4 \, \|\mathbf{x}\|^2 \, \|\mathbf{y}\|^2$$

is non-positive, which implies the desired inequality.

For the last assertion, notice that

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2 \, (\mathbf{x}|\mathbf{y}) \\ &\leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2 \, \|\mathbf{x}\| \, \|\mathbf{y}\| \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

□

From this lemma, we immediatly deduce the following.

Proposition 2.6. *The function*

$$\begin{aligned} \mathbb{R}^k \times \mathbb{R}^k &\to \mathbb{R}_{\geq 0} \\ (\mathbf{x}, \mathbf{y}) &\mapsto \|\mathbf{x} - \mathbf{y}\| \end{aligned}$$

is a metric on $\mathbb{R}^k$.

From now on, unless explicitly stated otherwise, we shall *always* consider that $\mathbb{R}^k$ is equipped with this metric, hence making it a metric space.

Email address: `l.s.guetta@uu.nl`

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht

ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 5

LÉONARD GUETTA

1. Sequences in metric spaces

Definition 1.1. Let (X, d) be a metric space. A sequence (x_n) in X converges to ℓ in X, if for all $\epsilon > 0$, there exists an $n \in \mathbb{N}$ such that if $n \geq N$, then $d(x, \ell) < \epsilon$.

Proposition 1.2. *When it exists, the sequence of a limit is unique.*

As it turns out, the convergence in $\mathbb{R}^k$ can be tested pointwise.

Proposition 1.3. *Let $(\mathbf{x}_n)$ be a sequence in $\mathbb{R}^k$, (i.e. each $\mathbf{x}_n = (x_{1,n}, x_{2,n}, \cdots, x_{k,n})$ is in $\mathbb{R}^k$). This sequence is convergent in $\mathbb{R}^k$ if and only if each sequence $(x_{i,n})_n$ for $1 \leq i \leq n$ is convergent. Moreover, if ℓ_i is the limit of $(x_{i,n})$, then*

$$\lim_{n \to +\infty} \mathbf{x}_n = (\ell_1, \ell_2, \cdots, \ell_k).$$

Sketch of proof. Notice that for a point $\mathbf{x} = (x_1, \cdots, x_k)$ of $\mathbb{R}^k$, we have

$$|x_i| \leq \|\mathbf{x}\| \leq \sum_{i=1}^{k} |x_i|,$$

for any $1 \leq i \leq k$. □

2. Open and closed subsets

Definition 2.1. Let (X, d) be a metric space, x an element of X and $r > 0$ a real number. We call *(open) neighborhood of x of radius r*, the set

$$N_r(x) = \{y \in X \mid d(x, y) < r\}.$$

Example 2.2. If $X = \mathbb{R}^2$, then $N_r(\mathbf{x})$ is nothing but the (open) disk of radius r and center $\mathbf{x}$.

Definition 2.3. Let (X, d) be a metric space, $E \subseteq X$ a subset of X and x an element of X.

(1) x is an *adherent point of E* if for all $r > 0$, the subset
$$E \cap N_r(x)$$
is non-empty. In other words, every open neighborhood of x contains at least an element of E,

(2) x is an *accumulation point of E* (also called *limit point of E*) if for all $r > 0$, the subset
$$E \cap (N_r(x) - \{x\})$$
is non-empty. In other words, every open neighborhood of x contains at least an element of E different from x.

(3) x is an *isolated point of* E if there exists an $r > 0$, such that

$$E \cap N_r(x) = \{x\}.$$

In other words, there exists a neighborhood of x that intersects E only at x,

(4) x in an *interior point of* E if there exists an $r > 0$ such that

$$N_r(x) \subseteq E.$$

In other words, there exists a neighborhood of x included in E.

Remark 2.4. Any element $x \in E$ is an adherent point of E. The converse is not true. For example, if $X = \mathbb{R}$ and $E = [0, 1)$, then 1 is an adherent point of E.

Not every element $x \in E$ is an accumulation point. It is if and only if it is not an isolated point (Exercise).

Finally, note that an isolated point of E is necessarily an element of E.

Definition 2.5. Let (X, d) be a metric space. A subset $E \subseteq X$ is *closed* if it contains all its adherent points.

Proposition 2.6. *A subset E of a metric space (X, d) is closed if and only if for every sequence (x_n) of elements of E which is convergent in X, the limit is in E*

$$\lim_{n \to +\infty} x_n \in E.$$

Proof. Let's start with the "only if" part. Let (x_n) be a sequence of elements of E, which is convergent to $\ell \in X$. Let us show that if E is closed, then $\ell \in E$. By definition, for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $n \geq N$, then $x_n \in N_\epsilon(\ell)$. This shows in particular that ℓ is adherent to E, and thus is an element of E because it is closed.

Now, the "if" part. Let x be an adherent point of E. For any $n > 0$, let $r_n = \frac{1}{n}$. Since x is an adherent point of E, contains $N_{r_n}(x)$ at least a point of E. For each $n > 0$, choose such an $x_n \in N_{r_n}(x) \cap E$. It is easy (exercise) to see that (x_n) converges to x, and therefore, by hypothesis, $x \in E$. □

Example 2.7. It follows from the previous proposition that the subset $[a, b]$ or $\mathbb{R}$, where a and b are real numbers, is closed. The same is true for intervals of the form $(-\infty, a]$ or $[a, +\infty)$.

Proposition 2.8. *Let E be a subset of $\mathbb{R}$ which is bounded above (resp. bounded below). If E is closed, then* $\sup E$ *(resp.* $\inf E$*) is an element of E.*

Proof. We do the bounded above case, this other one being similar. Let us show that $\alpha = \sup E$ is an adherent point. Because α is an upper bound of E, for every $x \in E$, we have $x \leq \alpha$. Now, for any $r > 0$, there exists an element of E in $(\alpha - r, \alpha]$, because otherwise, this would mean that all elements of E are smaller than $\alpha - r$, which would contradict the fact that α is the least upper bound. □

Definition 2.9. Let (X, d) be a metric space. A subset $E \subseteq X$ is *open* if for every element x of E is an interior point of E.

Recall that for a subset $E \subseteq X$, the complementary of E (in X) is the following subset of X:

$$E^c := \{x \in X \,|\, x \notin E\}$$

Proposition 2.10. *Let (X, d) be a metric space and $E \subseteq X$. Then E is open if and only if its completary is closed. Dually E is closed if and only if its complementary is open.*

Proof. Suppose that E is open, and let x be an adherent point to E^c. By definition, for all $r > 0$, $N_r(x)$ contains at least a point of E^c. This implies that $x \in E^c$, because otherwise, x would be an element of E, and since E is open, $N_r(x)$ would be included in E and could not contain a point of E^c. Conversely, suppose that E^c is closed and let $x \in E$. Since E^c is closed, x cannot be an adherant point of E^c, otherwise it would be in E^c. This means that there exists $r > 0$ such that $N_r(x) \subseteq E$.

The other statement follows immediatly from the fact that $(E^c)^c = E$. □

Proposition 2.11. *Let (X, d) be a metric space. The following holds:*

(1) $\emptyset$ and X are both open and closed subsets of X,
(2) if $(E_i)_i$ is an arbitrary family of open sets of X, then $\cup_i E_i$ is open,
(3) if $(E_i)_i$ is an arbitrary family of closed sets of X, then $\cap_i E_i$ is closed,
(4) if $(E_1, E_2, \cdots, E_n)$ is a finite *family of open sets of X, then $\cap_{i=1}^n E_i$ is open,*
(5) if $(E_1, E_2, \cdots, E_n)$ is a finite *family of closed sets of X, then $\cup_{i=1}^n E_i$ is closed.*

Proof. (1) Since $\emptyset$ contains no points, it satisfies trivially the conditions to be open and closed. The statement about X follows from the fact that $\emptyset^c = X$.

(2) Let x be an element of $\cup_i E_i$. By definition, there exists i_0 such that $x \in E_{i_0}$. Since E_{i_0} is open, there exists $r > 0$ such that $N_r(x) \subseteq E_{i_0} \subset \cup_i E_i$.

(3) This follows from the previous point, Proposition 2.10 and the fact that $(\cap_i E_i)^c = \cup_i E_i^c$.

(4) Let x be an element of $\cap_{i=1}^n E_i$. By definition, x is an element of all E_i. Since each E_i is open, for all $1 \leq i \leq n$, there exists $r_i > 0$ such that $N_{r_i}(x) \subseteq E_i$. Now if we set $r = \min\{i_1, \cdots, i_n\}$, then $r > 0$ and we have

$$N_r(x) \subseteq \cap_{i=1}^n E_i.$$

(5) This follows from the previous point, Proposition 2.10 and the fact that $(\cup_i E_i)^c = \cap_i E_i^c$.

□

Recall that any subset Y of a metric space (X, d) can also be seen as a metric space in itself. Note, however, that a subset $E \subseteq Y \subseteq X$ can be closed (resp. open) in Y without being closed (resp. open) in X. To avoid any confusion, we can use the terminology *closed (resp. open) relative to Y* for subsets of Y which are closed in Y, when considered as a metric space in itself.

Proposition 2.12. *Let (X,d) be a metric space and $Y \subseteq X$. A subset $E \subseteq Y$ of Y is open (resp. closed) relative to Y if and only if there exists an open (resp. closed) subset $F \subset X$ which is open (resp. closed) in X and such that*

$$E = Y \cap F.$$

Proof. For a point $y \in Y$ and $r > 0$, let us write $N_r^Y(y)$ for the neighborhood (of radius r) of y in Y and $N_r^X(y)$ for neighborhood (of radius r) of y in X. In other words, we have

$$N_r^Y(y) = \{z \in Y \mid d(y,z) < r\}$$

and

$$N_r^X(y) = \{z \in X \mid d(y,z) < r\}.$$

Notice that we have $N_r^Y(y) = N_r^X(x) \cap Y$. Let's now prove the "if" part of the proposition, in the open case and then the closed case. Let $E = Y \cap F$ with F open and let $x \in E$. In particular, $x \in F$ and since F is open in X, there exists an $r > 0$ such that $N_r^X(x) \subseteq F$, from which we deduce that $N_r^Y(x) = N_r^X(x) \cap Y \subseteq F \cap Y$, which proves that E is closed relative to Y. Let's do now the case F closed. Let $x \in Y$ be an adherent point to $Y \cap F$. This means that for all $r > 0$,

$$N_r^Y(x) \cap E = N_r^X(x) \cap Y \cap E = N_r^X(x) \cap Y \cap F \neq \emptyset.$$

In particular, $N_r^X(x) \cap F \neq \emptyset$, which means that x is adherent to F (in X). Since F is closed (in X), x is an element of F, and by hypothesis it is an element of Y, hence it is an element of E, which proves that E is closed (in Y).

Now the "only if" part. Suppose that E is open in Y. For every x in E, there exists $r_x > 0$, such that $N_{r_x}^Y(x) \subseteq E$. Define F as

$$F := \bigcup_{x \in E} N_{r_x}^X(x),$$

which is open in X as a union of open subsets of X. Moreover, we have

$$Y \cap F = \bigcup_{x \in E} (Y \cap N_{r_x}^X(x)) = \bigcup_{x \in E} N_{r_x}^Y(x) = E.$$

Let's now do the case when E is closed in Y. Then $Y - E$ is open in Y (because it is the complementary of E in Y), and from the previous proposition $Y - E = Y \cap F$ with F open in X. This implies that $E = Y \cap F^c$, which proves the assertion. □

Nevertheless, in the case that Y is closed or open in X, we have the following corollary.

Corollary 2.13. *Let (X,d) be a metric space and $Y \subset X$ an* open *(resp.* closed*) subset of X. Then a subset $E \subseteq Y$ is open (resp. closed) relative to Y if and only if it is open (resp. closed) in X.*

Proof. Exercise. □

Email address: `l.s.guetta@uu.nl`

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht

ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 6

LÉONARD GUETTA

1. Compact subsets

Definition 1.1. Let (X,d) be a metric space and $E \subseteq X$ be a subset. We say that E is *compact* if every infinite subset of E has a limit point in E.

Example 1.2. If E is finite then it is trivially compact. We shall see soon many examples of compact subset of $\mathbb{R}^k$ which are infinite.

Remark 1.3. Contrary to open or closed subsets, the property of being compact for a subset does not depend on the space in which it is embedded. In particular, we can just unambiguously speak of a compact metric space.

We are going to reformulate the definition of compact subsets using sequences. In order to do that, we need the following definition.

Definition 1.4. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a metric space. A *subsequence* of $(x_n)_{n\in\mathbb{N}}$ is a sequence of the form $(x_{n_k})_{k\in\mathbb{N}}$, where

$$n_0 < n_1 < n_2 < \cdots$$

is a *strictly* increasing sequence of natural integers. If a subsequence converges, we call the limit of this subsequence a *subsequential limit of* (x_n).

Example 1.5. For any sequence (x_n), we can consider the subsequence $(x_{2p})_p$ of even indices, and the subsequence $(x_{2p+1})_p$ of odd indices.

Proposition 1.6. *A sequence (x_n) in a metric space is convergent to ℓ if and only if every subsequence converges to ℓ.*

Proof. Exercise. □

Remark 1.7. Note that a sequence can be divergent but still have certain subsequences that are convergent. For example, for the sequence $(u_n = (-1)^n)$, the subsequence (u_{2k}) converges to 1 and the subsequence (u_{2k+1}) converges to -1.

Lemma 1.8. *A sequence (x_n) in a metric space (X,d). A point $x \in X$ is a subsequential limit of (x_n), if and only if for all $\epsilon > 0$, there exists infinitely many $n \in \mathbb{N}$ such that $x_n \in N_\epsilon(x)$.*

Proof. "If" part. We are going to construct recursively a subsequence of (x_n) which converges to x. For all $k \geq 1$, let $\epsilon_k = \frac{1}{k}$. First choose any n_1 in $\mathbb{N}$ such that $x_{n_1} \in N_{\epsilon_1}(x)$ (which is possible by hypothesis). Suppose now that we have constructed $n_1 < n_2 < \cdots < n_k$ such that $x_{n_k} \in N_{\epsilon_k}(x)$. Because we know that there exists infinitely many $n \in \mathbb{N}$ such that $x_n \in N_{\epsilon_{k+1}}(x)$, there exists at least one $n_{k+1} > n_k$ such that $x_{n_{k+1}} \in N_{\epsilon_{k+1}}(x)$. Repeating this process, this construct a subsequence $(x_{n_k})_k$ of (x_n) such that for all $k > K$, $x_{n_k} \in N_{\epsilon_K}(x)$, which shows that (x_{n_k}) converges to x.

"Only if" part. Let (x_{n_k}) be a subsequence converging to x. For all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that if $k > N$ then $x_{n_k} \in N_\epsilon(x)$. In particular for any $n = n_k$ with $k > N$, we have $x_n \in N_\epsilon(x)$, which is infinitely many of such n.

□

Theorem 1.9. *Let (X, d) be a metric space and $E \subset X$. Then E is a compact if and only if every sequence (x_n) in E admits a subsequential limit in E.*

Proof. "If" part. Let $F \subseteq E$ be an infinite subset of E. Because this subset is infinite, there exists a sequence (x_n) of elements of F such that if $n \neq m$, then $x_n \neq x_m$. By hypothesis, this sequence admits a subsequential limit $x \in E$. Let us show that x is an accumulation point of F. By Lemma 1.8, this means that for all ϵ there exists infinitely many Suppose that E is compact and let (x_n) be a sequence in E. Consider the subset $F = \{x_n \mid n \in \mathbb{N}\}$ of E. There are two cases to consider. First if F is finite $F = \{y_1, y_2, \cdots, y_k\}$, then necessarily A FINIR. If F is infinite, then, because E is compact, it admits an accumulation point in E, which means exactly that (x_n) admits a subsequential limit in E by Lemma 1.8.

□

Proposition 1.10. *Every compact subset of a metric space is closed.*

Proof. Let (x_n) be a sequence of K which is convergent to $\ell \in X$. We need to show that $\ell \in K$. Because K is compact, we know that (x_n) has a sequential limit in K, which must be equal to ℓ by Proposition 1.6, hence $\ell \in K$. □

Proposition 1.11. *Every closed subset of a compact is compact.*

Proof. Let $F \subset K \subset X$, with (X, d) metric space, K compact and F closed in K (or equivalently in X, since K is necessarily closed by Proposition 1.10). Let (x_n) be a sequence in F. It is also a sequence in K, and thus admits a subsequential limit in K since K is compact. But because F is closed, this subsequential limit is also in F, which proves that F is compact. □

Definition 1.12. Let (X, d) be a metric space. A subset $E \subseteq X$ is bounded if there exists and $M \in \mathbb{R}_{\geq 0}$ such that

$$d(x, y) \leq M$$

for all $x, y \in E$.

Example 1.13. A subset $E \subseteq \mathbb{R}^k$ is bounded if and only if there exists $M_1, M_2, \cdots, M_k$ such that $E \subseteq [-M_1, M_1] \times [-M_2, M_2] \times \cdots \times [-M_k, M_k]$ (Exercise). In particular, for $k = 1$, we recover the usual definition of boundedness.

Theorem 1.14. *Let (X, d) be a metric space and $K \subseteq X$. The following holds:*

(1) If K is compact, then K is closed and bounded.

(2) *If $X = \mathbb{R}^k$, then the converse holds, that is, K is compact* if and only if *it is closed and bounded.*

Sketch of proof. □

Corollary 1.15. *Every bounded sequence of $\mathbb{R}^k$ has a subsequential limit.*

2. Complete metric spaces

Definition 2.1. A metric space (X, d) is *complete* if and only if every Cauchy sequence in X is convergent.

Example 2.2. We have already seen that $\mathbb{R}$ is convergent. More generally, we have the following result.

Proposition 2.3. *For any $k \geq 1$, the metric space $\mathbb{R}^k$ is complete.*

Proof. It is easy to see that a sequence $(\mathbf{x}_n)$ in $\mathbb{R}^k$ (i.e. where each $\mathbf{x}_n = (x_{1,n}, x_{2,n}, \cdots, x_{k,n})$ is a point in $\mathbb{R}^k$) is Cauchy if and only if each of the sequence

$$(x_{i,n})_n$$

is Cauchy, for $1 \leq i \leq k$. Using that $\mathbb{R}$ is complete and thus each of this sequence is convergent, it follows that $(\mathbf{x}_n)$ is convergent, which proves that $\mathbb{R}^k$ is complete. □

Proposition 2.4. *Every closed subset $F \subseteq X$ of a complete metric space is complete (when F is equipped with the metric induced by X).*

Proof. Exercise. □

Proposition 2.5. *Every compact metric space is complete.*

Proof. Exercise. □

Email address: `l.s.guetta@uu.nl`

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht

ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 7

LÉONARD GUETTA

1. Continuity

Definition 1.1. Let X and Y be metric spaces, x_0 an element of X and $f\colon X \to Y$ a function. We say that f *is continuous at* x_0 if for all $\epsilon > 0$, there exists $\delta > 0$ such that if $d_X(x_0, x) < \delta$, then $d_Y(f(x_0), f(x)) < \epsilon$.

Proposition 1.2. *A function $f\colon X \to Y$ is continuous at $x_0 \in X$ if and only if for every sequence (x_n) in X which converges to x_0, the sequence $(f(x_n))$ converges to $f(x_0)$.*

Proof. Suppose that f is continuous at x_0 and let (x_n) be a sequence converging to x_0. For any $\epsilon > 0$, there exists $\delta > 0$ such that if $d_X(x_0, x) < \delta$ then $d_Y(f(x_0), f(x)) < \epsilon$, and there exists $N \in \mathbb{N}$ such that if $n \geq N$, then $d_X(x_0, x_n) < \delta$. In particular, if $n \geq N$ then $d_Y(f(x_n), f(x_0)) < \epsilon$, which proves that $(f(x_n))$ converges to $f(x_0)$.

For the converse, let us prove the contraposive, that is if there exists a sequence (x_n) converging to x_0 such that $f(x_n)$ does not converge to $f(x_0)$, then f is not continuous at x_0. The fact that $f(x_n)$ does not converge to $f(x_0)$ means that there exists $\epsilon > 0$, such that for all $N \in \mathbb{N}$, there exists $n \geq N$ such that $d_Y(f(x_0), f(x_n)) \geq \epsilon$. Now, because (x_n) is converging to x_0, we know that for all $\delta > 0$, there exists $N_\delta \in \mathbb{N}$ such that if $n \geq N_\delta$, then $d_X(x_0, x_n) < \delta$. In particular, this proves that for all δ, there exists at least one x_n such that $d_X(x_0, x_n) < \delta$, but $d_Y(f(x_0), f(x)) \geq 0$, which proves that f is not continuous at x_0. □

Corollary 1.3. *Let $f\colon X \to Y$ and $g\colon Y \to Z$ be functions, with X, Y and Z metrics spaces. If f is continuous at $x_0 \in X$ and g is continuous at $f(x_0) \in Y$, then $g \circ f$ is continuous at x_0.*

Definition 1.4. Let X and Y be two metric spaces. A function $f\colon X \to Y$ is *continuous* if it is continuous at every $x_0 \in X$.

Recall that given a function $f\colon X \to Y$ and $E \subseteq Y$, we denote by $f^{-1}(E)$ the subset of X defined as

$$f^{-1}(E) := \{x \in X \mid f(x) \in E\}$$

Proposition 1.5. *Let X and Y be metric spaces and $f\colon X \to Y$ a function. The following are equivalent:*

1. *f is continuous,*
2. *for every open subset O of Y, the subset $f^{-1}(O)$ of X is open,*
3. *for every closed subset F of Y, the subset $f^{-1}(F)$ of X is closed.*

Proof. First notice that (2) and (3) are equivalent because for any subset $E \subseteq Y$, we have $f^{-1}(E^c) = (f^{-1}(E))^c$.

Let us prove that (1) $\Rightarrow$ (2). Suppose that f is continuous and let O be an open subset of Y. We want to show that $f^{-1}(O)$ is open, that is, that every point of $f^{-1}(O)$ is interior to $f^{-1}(O)$. Let $x \in f^{-1}(O)$, that is $f(x) \in O$. Since O is open, there exists $\epsilon > 0$ such that $N_\epsilon(f(x)) \subseteq O$, and because f is continuous, there exists $\delta > 0$ such that if $x' \in N_\delta(x)$ then $f(x') \in N_\epsilon(f(x))$. This implies that $f(N_\delta(x)) \subseteq O$, and so that $N_\delta(x) \subseteq f^{-1}(O)$, hence x is interior to $f^{-1}(O)$.

Let us now prove that (2) $\Rightarrow$ (1). Let $x \in X$ and let $\epsilon > 0$. The subset $N_\epsilon(f(x))$ of Y is open (see exercises from week 5), and so $f^{-1}(N_\epsilon(f(x))$ is open. Since $x \in f^{-1}(N_\epsilon(f(x)))$, there exists $\delta > 0$, such that $N_\delta(x) \subseteq f^{-1}(N_\epsilon(f(x)))$, which means exactly the continuity at x. □

We end this section with the relation between composition of functions and continuity.

Proposition 1.6. *Let $f\colon X \to Y$ and $g\colon Y \to Z$ be functions with X, Y and Z metric spaces. If f is continuous at $x_0 \in X$ and g is continuous at $f(x_0) \in Y$, then $g \circ f$ is continuous at x_0.*

In particular, if f and q are continuous, then $g \circ f\colon X \to Y$ is continuous.

Proof. Let (x_n) be a sequence in X converging to x_0. Since f is continuous, then $f(x_n)$ converges to $f(x_0)$, and since g is continuous, then $g(f(x_n))$ converges to $g(f(x_0))$, which shows the continuity of $g \circ f$ at x_0. □

2. Numerical functions

Lemma 2.1. *Let (X, d) be a metric space. A function*

$$\mathbf{f} = (f_1, f_2, \cdots, f_n)\colon X \to \mathbb{R}^k$$

is continuous if and only if $f_i\colon X \to \mathbb{R}$ is continuous for each $1 \leq i \leq k$.

Proof. This follows easily from Proposition 1.2 and the fact that a sequence $(\mathbf{x}_n)$ in $\mathbb{R}^k$ converges if and only if each of its coordinate sequences converge (Proposition 1.3 of Week 5's lecture). □

Proposition 2.2. *Let (X, d) be a metric space and let $\mathbf{f}, \mathbf{g}\colon X \to \mathbb{R}^k$ be functions. The following hold:*

(1) if $\mathbf{f}$ and $\mathbf{g}$ are continuous then $\mathbf{f} + \mathbf{g}$ is continuous,
(2) if $\mathbf{f}$ and $\mathbf{g}$ are continuous then $(\mathbf{f}|\mathbf{g})$ is continuous.

Moreover, in the case that $k = 1$ and $g(x) \neq 0$ for all $x \in X$, we have

(3) if f and g are continuous, then $\frac{f}{g}$ is continuous.

Proof. Left as an exercise. □

Definition 2.3. A function $f\colon \mathbb{R}^k \to \mathbb{R}$ is called *monomial* if it is of the form

$$f\colon \mathbb{R}^k \to \mathbb{R}$$

$$(x_1, \cdots, x_k) \mapsto x_{i_1} x_{i_2} \cdots x_{i_n},$$

where $i_1, i_2, \cdots, i_n$ are integers between 1 and k (included).

A function $f\colon \mathbb{R}^k \to \mathbb{R}$ is called *polynomial* if it is a sum of monomial functions.

Proposition 2.4. *A polynomial function* $f\colon \mathbb{R}^k \to \mathbb{R}$ *is continuous.*

Proof. This follows immediatly from Proposition 2.2 (a) and (b). □

Proposition 2.5. *The following real functions are continuous:*

- $\mathbb{R}_{\geq 0} \to \mathbb{R}, x \mapsto \sqrt[n]{x}$ *with* n *even,*
- $\mathbb{R} \to \mathbb{R}, x \mapsto \sqrt[n]{x}$ *with* n *odd,*
- $\mathbb{R} \to \mathbb{R}, x \mapsto \exp(x)$*,*
- $\mathbb{R}_{\geq 0} \to \mathbb{R}, x \mapsto \ln(x)$*.*

Proof. Admitted. See chapter ?? of Rudin. □

We finish this section with a very important theorem on real functions.

Theorem 2.6 (Intermediate-value theorem)**.** *Let* $f\colon [a, b] \to \mathbb{R}$ *be a continuous function such that* $f(a) < f(b)$*. For all* $f(a) < c < f(b)$*, there exists* $a \leq x \leq b$ *such that* $f(x) = c$*.*

Proof. Let's start with the case $f(a) < 0$, $f(b) > 0$ and $c = 0$. We need to show that there exists $a \leq x \leq b$ such that $f(x) = 0$. □

Email address: l.s.guetta@uu.nl

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht

ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 9

LÉONARD GUETTA

1. Continuous functions and compactness

Definition 1.1. A function $f\colon X \to \mathbb{R}$ is *bounded* if there exists an $M \in \mathbb{R}_{\geq 0}$ such that

$$|f(x)| \leq M$$

for all $x \in X$.

Recall that given a function $f\colon X \to Y$ and $E \subseteq X$, we denote by $f(E)$ the subset of Y defined as

$$f(E) := \{y \in Y \mid \text{ there exists } x \in X \text{ such that } f(x) = y\}.$$

Proposition 1.2. *Let X and Y be metric spaces and $f\colon X \to Y$ a continuous function. If K is a compact subset of X, then $f(K)$ is a compact subset of Y.*

Proof. Let (y_n) be a sequence in $f(K)$. By definition, this means that each y_n is of the form $y_n = f(x_n)$ with $x_n \in K$. Since K is compact, there exists a subsequence of (x_{n_k}) which is convergent in K. Since f is continuous, $(f(x_{n_k})$ is convergent in $f(K)$, which proves that $f(K)$ is compact. □

Corollary 1.3. *Let $f\colon X \to \mathbb{R}^k$ be a continuous function. If K is compact subset of X, then $f(K)$ is bounded (and closed).*

Corollary 1.4. *Let $f\colon X \to \mathbb{R}$ be a continuous function. If X is compact, and*

$$M = \sup_{x\in X} f(x), \quad m = \inf_{x\in X} f(x).$$

Then there exists x and x' in X such that $f(x) = M$ and $f(x') = m$.

Proof. Exercise. □

Email address: l.s.guetta@uu.nl

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht

ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 10

LÉONARD GUETTA

1. Pointwise convergence

Definition 1.1. Let X be a set and (Y, d_Y) a metric space. Let (f_n) be a sequence of functions $f_n : X \to Y$, we say that (f_n) *converges pointwise (on X) to a function* $f : X \to Y$, if for every $x \in X$, we have

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

Suppose that (f_n) converges pointwise on X to f. We can ask the following questions:

(1) If each (f_n) is continuous and the sequence converges pointwise to a function f, is f also continuous?

More generally, we can ask similar questions for any other kinds of properties involving limits.

(2) If $X = [a, b]$, do we have

$$\lim_{n \to +\infty} \int_a^b f(x)dx = \int_a^b f(x)dx?$$

(3) If X is a real interval and each f_n is derivable, do we have

$$\lim_{n \to +\infty} f_n'(x) = f'(x)?$$

In general, the answers to these questions are all no. In this lecture, we will focus on the first one of these questions, for which the following example proves that the answer is negative.

Example 1.2. Let $f_n(x) = x^n$. On $[0, 1]$ this sequence converges pointwise to the function f given by

$$f(x) = \begin{cases} 0 & \text{for } 0 \leq x < 1, \\ 1 & \text{for } x = 1. \end{cases}$$

While each f_n is continuous on $[0, 1]$, the limit function f is not.

2. Uniform convergence

Spelled out, the definition of pointwise convergence means that for every $x \in X$, for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that if $n \geq N$, then

$$d_Y(f_n(x), f(x)) < \epsilon.$$

In particular, the integer N can depend on the point x. This allows for a stronger notion of convergence as follows.

Definition 2.1. Let X be a set and (Y, d_Y) be a metric space. Let (f_n) be a sequence of functions $f_n\colon X \to Y$, we say that (f_n) *converges uniformely to a function* $f\colon X \to Y$, if for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$, such that if $n \geq N$ then for all $x \in X$ we have

$$d(f_n(x), f(x)) < \epsilon.$$

Note that uniform convergence implies pointwise convergence.

Uniform convergence can be reformulated as follows.

Proposition 2.2. *Let (f_n) be a sequence of functions that converges pointwise to f. Then (f_n) converges uniformly to f on X if and only if the sequence*

$$\sup_{x \in X} d_Y(f_n(x), f(x))$$

converges to 0 when $n \to +\infty$.

Note also that in the case of real valued functions, we can apply the Cauchy criterion.

Theorem 2.3 (Cauchy criterium)**.** *Let (f_n) be a sequence of function $f_n\colon X \to \mathbb{R}$. This sequence is uniformely convergent if and only if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $x \in X$, if $n, m \geq N$ then*

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

Our main result concerning uniform convergence is the following.

Theorem 2.4. *Let (X, d_X) and (Y, d_Y) be metric spaces and (f_n) a sequence of functions $f_n\colon X \to Y$ converging uniformely to $f\colon X \to Y$. If for every n, f_n is continuous then f is also continuous.*

Proof. Let $x_0 \in X$. Let $\epsilon > 0$, we know that there exists $N \in \mathbb{N}$ such that if $n \geq N$, then

$$d_Y(f_n(x), f(x)) < \frac{\epsilon}{3}$$

for all $x \in X$. Since f_N is continuous in x_0, there exists $\delta > 0$ such that if $d_X(x, x_0) < \delta$, then $d_Y(f_N(x_0), f_N(x)) < \frac{\epsilon}{3}$. Hence, if $d_X(x, x_0) < \delta$, we have

$$d_Y(f(x), f(x_0)) \leq d_Y(f(x), f_N(x)) + d_Y(f_N(x), f_N(x_0)) + d_Y(f_N(x_0), f(x_0)) < \epsilon.$$

□

3. Series of functions

Uniform convergence also applies to series of functions.

Definition 3.1. Let X be a set and (f_n) a sequence of function $f_n\colon X \to \mathbb{R}$. We say that the series $\sum f_n$ *converges uniformely* if the sequence of partial sums

$$\sum_{k=0}^{n} f_k$$

converges uniformely.

The following criterion is useful to determine the uniform convergence of a series of functions.

Theorem 3.2 (Weierstrass)**.** *Let* (f_n) *be a sequence of functions* $f_n\colon X \to \mathbb{R}$ *and suppose that there exists a real sequence* (M_n) *such that*

$$|f_n(x)| \leq M_n$$

for every $n \in \mathbb{N}$ *and every* $x \in X$*.*

If $\sum M_n$ *converges then* $\sum f_n$ *converges uniformely.*

Note that the important hypothesis in the previous theorem is that M_n does *not* depend on x.

Proof. Exercise. □

4. Uniform convergence and continuity

Email address: `l.s.guetta@uu.nl`

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht

ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 11

LÉONARD GUETTA

1. Geometric interpretation of the derivative

1.1. The idea of derivation is to approximate a function $f\colon I \to \mathbb{R}$ at a given $x_0 \in I$ by a line. Recall that the equation of a line is given by $y = ax + b$, hence we would like to say something like

$$f(x) \simeq ax + b$$

when "x is close to x_0". In particular, for $x = x_0$, we must have $f(x_0) = ax_0 + b$ and thus, we can rewrite the previous equation as

$$f(x) \simeq f(x_0) + a(x - x_0),$$

when x is close to x_0. Or, if we set $\delta x = x - x_0$, we can rewrite it as

$$f(x_0 + \delta x) \simeq f(x_0) + a\delta x,$$

when δx is close to 0. But how do we formalize the symbol $\simeq$? The idea is to say that this equality is true up to a remainder that is "neglectable in front of δx". Hence, we can write

$$f(x_0 + \delta x) = f(x_0) + a \cdot \delta x + r(\delta x),$$

where r is a function such that $\lim_{\delta x \to 0} \frac{r(\delta x)}{\delta x} = 0$, which means intuitively that $r(\delta x)$ tends to 0 much faster than δx. If we divide the previous equality by δx, we obtain

$$\frac{f(x_0 + \delta x) - f(x_0)}{\delta x} = a + \frac{r(\delta x)}{\delta x},$$

hence

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

Hence, a is nothing but the derivative of f at x_0.

Definition 1.2. Let I be an interval of $\mathbb{R}$, $f\colon I \to \mathbb{R}$ a function and $x_0 \in \mathbb{R}$. We say that f *is derivable at* x_0 if the following limit exists

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

In which case, the value of this limit is denoted by $f'(x_0)$ and referred to as the *derivative of* f *at* x_0.

If f is derivable at every point of I, we say that f *is derivable on* I, or simply *derivable*.

Proposition 1.3. *Let $f\colon I \to \mathbb{R}$ be a function and $x_0 \in I$. If f is derivable at x_0, then we have*

$$f(x_0 + h) = f(x_0) + f'(x_0) \cdot h + r(h),$$

where r is a function such that $lim_{h\to 0}\frac{r(h)}{h} = 0$.

Proof. It suffices to take $r(h) = f'(x_0) \cdot h - (f(x + h) - f(x))$. □

Remark 1.4. Sometimes, the equality of the previous proposition is written as

$$f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + r(x - x_0),$$

with r a function such that $\lim_{x\to x_0} \frac{r(x-x_0)}{x-x_0} = 0$.

Definition 1.5. If f is derivable at x_0, the line of equation

$$y = f(x_0) + f'(x_0) \cdot (x - x_0)$$

is called the *tangent of f at x_0*.

Let us end this section with an easy but essential result.

Proposition 1.6. *Let $f\colon I \to \mathbb{R}$ be a function and $x_0 \in I$. If f is derivable at x_0 then it is continuous at x_0.*

Proof. Because we have $f(x_0 + h) = f(x_0) + f'(x_0) \cdot h + r(h)$, we have

$$\lim_{h\to 0} f(x_0 + f) = f(x_0)$$

(because $\lim_{h\to 0} r(h) = 0$). □

2. Local study

Definition 2.1. Let $f\colon I \to \mathbb{R}$ a function, and x_0 an interior point of I. We say that x_0 is a *local maximum* (resp. *local minimum*) if there exists $\delta > 0$ such that for $x \in (x_0 - \delta, x_0 + \delta)$, we have

$$f(x) \geq f(x_0)$$

(resp.

$$f(x) \leq f(x_0)).$$

We say that x_0 is a *local extremum* if it is either a local maximum or a local minimum.

Proposition 2.2. *Let $f\colon I \to \mathbb{R}$ and x_0 an* interior point to I. *If x_0 is a local extremum of f and f is derivable at x_0, then $f'(x_0) = 0$.*

Proof. Suppose that $f'(x_0) \neq 0$, for example $f'(x_0) > 0$. In particular, there exists $\delta > 0$ such that for $x \in (x_0 - \delta, x_0 + \delta)$, we have

$$\frac{f(x) - f(x_0)}{x - x_0} > 0.$$

This means that $f(x) - f(x_0)$ is of the sign of $x - x_0$, which contradicts the fact that $f(x_0)$ is a local extremum.

A similar argument applies in the case $f'(x_0) < 0$. □

Even though the derivative allows us to find local extrema, it is not sufficient to determine whether this local extremum is a local minimum or a local maximum. In order to do that, we need higher order derivatives.

3. Higher order derivatives

3.1. Proposition 1.3 can be restated by saying that when f is derivable at x_0, then f can be approximated by a polynomial of degre 1 around x_0. In fact, we can go further and wonder if it's possible to approximate f by a polynomial of degree 2, that is, we wonder if there are $a, b, c \in \mathbb{R}$ such that

$$f(x_0 + h) = a + b \cdot h + c \cdot h^2 + r(h),$$

where r is a function such that $\lim_{h\to 0} \frac{r(h)}{h^2} = 0$. More generally, we can wonder if we can approximate f by a polynomial of degree n around x_0, that is if there exists $a_0, a_1, \cdots, a_n \in \mathbb{R}$

$$f(x_0 + h) = a_0 + a_1 \cdot h + \cdots + a_n \cdot h^n + r(h),$$

where $\lim\limits_{h\to 0} \dfrac{r(h)}{h^n} = 0$.

Definition 3.2. Let $f\colon I \to \mathbb{R}$ be a function, with I interval of $\mathbb{R}$. If f is derivable on I, and the derived function $f'\colon I \to \mathbb{R}$ is also derivable, then we denote by f'' or $f^{(2)}$ the derivative of f'. In this case, we say that the function f is *two-times derivable* and $f^{(2)}$ is called the *second derivative of* f.

More generally, we say that f is *n-times derivable* if $f^{(n-1)}$ is derivable, and we call $f^{(n)}$ the *n-th derivative of* f.

Definition 3.3. Let I be an interval of $\mathbb{R}$ and $n \geq 1$. We denote by $\mathcal{C}^n(I)$ the set of n-times derivable functions $f\colon I \to \mathbb{R}$ such that all $f^{(n)}$ is continuous.

We also denote by $\mathcal{C}^0(I)$ the set of continuous functions on I.

Finally, we denote by $\mathcal{C}^\infty(I)$ the set of infinitely derivable functions on I, that is

$$\mathcal{C}^\infty(I) = \bigcap_{n\geq 0} \mathcal{C}^n(I).$$

Remark 3.4. We have

$$\mathcal{C}^0(I) \supset \mathcal{C}^1(I) \supset \mathcal{C}^2(I) \supset \cdots \supset \mathcal{C}^\infty(I).$$

Proposition 3.5 (Taylor-Young)**.** *Let I be an interval of $\mathbb{R}$ and $x_0 \in I$. For any $f \in \mathcal{C}^n(I)$, we have*

$$f(x_0 + h) = f(x_0) + f'(x_0) \cdot h + \frac{f^{(2)}(x_0)}{2} \cdot h^2 + \cdots \frac{f^{(n)}(x_0)}{n!} \cdot h^n + r(h),$$

where r is a function such that $\lim\limits_{h\to 0} \dfrac{r(h)}{h^n} = 0$.

Definition 3.6. Let $f \in \mathcal{C}^n(I)$ and $x_0 \in I$. The polynomial in h

$$f(x_0) + f'(x_0) \cdot h + \frac{f^{(2)}(x_0)}{2} \cdot h^2 + \cdots \frac{f^{(n)}(x_0)}{n!}$$

is called the *Taylor polynomial of f of degree n*, and the formula

$$f(x_0 + h) = f(x_0) + f'(x_0) \cdot h + \frac{f^{(2)}(x_0)}{2} \cdot h^2 + \cdots \frac{f^{(n)}(x_0)}{n!} \cdot h^n + r(h)$$

is called the *Taylor expansion of f of degree n*.

Remark 3.7. As before, we can also write the Taylor expansion at x_0 as

$$f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + \cdots \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + r(x - x_0),$$

with r a function such that $\lim_{x \to x_0} \frac{r(x-x_0)}{(x-x_0)^n} = 0$.

4. Local analysis

The use of the Taylor expension is very useful to study the local behaviour of a function.

Proposition 4.1. *Let $f: I \to \mathbb{R}$ be a function in $\mathcal{C}^\infty(I)$ and let x_0 be an interior point of I such that $f'(x_0) \neq 0$. Suppose that there exists $n > 1$ such that $f^{(n)} \neq 0$ and let k be the first of such integers n. Then:*

(1) if k is even, then x_0 is a local extremum and more precisely:
  (a) a local minimum if $f^{(k)}(x_0) > 0$,
  (b) a local maximum if $f^{(k)}(x_0) < 0$,

(2) if k is odd, then x_0 is an inflexion point and more precisely
  (a) f is (locally) increasing at x_0 if $f^{(k)}(x_0) > 0$,
  (b) f is (locally) decreasing at x_0 if $f^{(k)}(x_0) < 0$.

Proof. By the Taylor formula, we have

$$f(x) - f(x_0) = \frac{f^{(k)}(x_0)}{k!} \cdot (x - x_0)^k + r(x - x_0).$$

In particular, since $f^{(k)}(x_0) \neq 0$, $f(x) - f(x_0)$ is of the sign of $f^{(k)}(x_0) \cdot (x - x_0)^k$ in a neighborhood of x. From this, all the cases follow easily. □

Email address: l.s.guetta@uu.nl

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht

ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 12

LÉONARD GUETTA

1. Derivatives

1.1. If $A\colon \mathbb{R}^m \to \mathbb{R}^n$ is a *linear* map and $h \in \mathbb{R}^m$, we will use the notation

$$A \cdot h \text{ or even } Ah \text{ instead of } A(h).$$

The notation Ah is consistent with the fact that A can be represented by a matrix and h by a column vector (i.e. a $m \times 1$ matrix), and Ah is obtained by matrix multiplication.

Definition 1.2. Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ be a function, with U an open subset of $\mathbb{R}^n$. We say that f *is differentiable (or derivable) at* $x \in U$ if there exists a linear map $A\colon \mathbb{R}^m \to \mathbb{R}^n$ such that

$$f(x+h) = f(x) + Ah + r(\|h\|),$$

for $h \in \mathbb{R}^m$ in a neighborhood of 0 (so that $x+h$ is in a neighborhood of x) and r is a function such that

$$\lim_{h \to 0} \frac{\|r(h)\|}{\|h\|} = 0.$$

We say that f is *differentiable (or derivable)* on U if it is differentiable at every point $x \in U$.

Remark 1.3. In other words, f is differentiable at x if and only if there exists a linear map $A\colon \mathbb{R}^n \to \mathbb{R}^m$ such that

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

Lemma 1.4. *If f is differentiable at x, then the linear map A as in the previous definition is unique.*

1.5. Consequently, if $f\colon U \subset \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at x_0 then we write $D_x f$ or $f'(x)$, and call refer to the *derivative of f at x* for the unique linear map $\mathbb{R}^n \to \mathbb{R}^m$ such that

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

As you know, a linear map can be represented by a matrix. We are now going to see what the matrix of $D_x f$ looks like. Let's start by the components.

1.6. Recall that a function $f\colon U \subset \mathbb{R}^m \to \mathbb{R}^n$ is of the form

$$f(x) = (f_1(x), f_2(x), \cdots, f_n(x)).$$

The f_j, for each $1 \leq j \leq m$ are called the *components of f*.

Note that using the matrix notation, we are saying that

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix}$$

Proposition 1.7. *A function $f\colon U \subset \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at $x \in U$ if and only if each of the components functions f_j is differentiable at x. In this case, we have*

$$D_x f \cdot h = (D_x f_1 \cdot h, D_x f_2 \cdot h, \cdots, D_x f_m \cdot h).$$

Note that in matrix notation, it becomes

$$D_x f \cdot h = \begin{pmatrix} D_x f_1 \cdot h \\ D_x f_2 \cdot h \\ \vdots \\ D_x f_m \cdot h \end{pmatrix}.$$

Definition 1.8. Let $f = (f_1, \cdots, f_n)\colon U \to \mathbb{R}^n$ be a function, with U an open subset, and let $\mathbf{x} = (x_1, \cdots, x_m)$ a point of U. The partial derivative

$$\frac{\partial f_j}{\partial x_i}$$

for $1 \leq j \leq n$ and $1 \leq i \leq m$, is the derivative of the the real function

$$t \mapsto f_j(x_0, \cdots, x_{i-1}, t, x_{i+1}, \cdots, x_m).$$

Remark 1.9. In other words, we have

$$\frac{\partial f_j}{\partial x_i} = \lim_{t \to 0} \frac{f(\mathbf{x} + te_i) - f(\mathbf{x})}{t},$$

where e_i is the i-th vector of the standard base of $\mathbb{R}^m$.

Proposition 1.10. *Let $f\colon U \to \mathbb{R}^n$ be a function, with $U \subseteq \mathbb{R}^m$ an open subset. If f is differentiable at $x \in U$, then the matrix of $D_x f$, referred to as the* Jacobi matrix *of f at x, and denoted by $J_x f$ is given by*

$$J_x f = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix}.$$

Remark 1.11. More explicitly, the previous proposition says that

$$D_x f \cdot h = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

Email address: `l.s.guetta@uu.nl`

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht

ADVANCED MATHEMATICS 2023-2024
LECTURE – WEEK 13

LÉONARD GUETTA

1. Continuously differentiable functions

1.1. We denote by $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ the set of linear maps $A\colon \mathbb{R}^m \to \mathbb{R}^n$. Henceforth, we shall identify such maps with $n \times m$ matrices.

Definition 1.2. Let $A \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$, seen as a matrix $(a_{i,j})$. We define its norm $\|A\|$ as the following non-negative real number

$$\|A\| := \sqrt{\sum_{j=1}^{m} \sum_{i=1}^{n} |a_{i,j}|^2}$$

Lemma 1.3. *Let $A, B \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ and $X \in \mathbb{R}^m$, the following holds:*

(1) $\|A\| = 0$ *if and only if* $A = 0$,
(2) $\|A + B\| \leq \|A\| + \|B\|$,
(3) $\|AX\| \leq \|A\| \, \|X\|$.

1.4. In particular, it follows that we have a metric on $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ defined as $d(A, B) = \|A - B\|$. We shall now always equip the set $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ with this metrc. This allows us to speak of the continuity of $x \mapsto D_x f$ as is stated in the following definition.

Definition 1.5. Let $U \subset \mathbb{R}^m$ an open subset and $f\colon U \to \mathbb{R}^m$ a function. We say that f is $\mathcal{C}^1$ if f is differentiable in U and the map

$$\begin{aligned} Df\colon U &\to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n) \\ x &\mapsto D_x f \end{aligned}$$

is continuous.

Proposition 1.6. *A map $f\colon U \to \mathbb{R}^n$, where U is an open subset of $\mathbb{R}^m$, is $\mathcal{C}^1$ if and only if all the partial derivative*

$$\frac{\partial f_j}{\partial x_i}$$

for $1 \leq i \leq m$ and $1 \leq j \leq n$, exist on U and are continuous.

Remark 1.7. In particular, the previous theorem tells us that if all the partial derivatives exist and are continuous, then f is differentiable on U.

Note that if all partial derivatives of f are differentiable, we can take their derivatives again, leading to higher derivatives.

Definition 1.8. Let $U \subseteq \mathbb{R}^m$ an open subset. A function $f\colon U \to \mathbb{R}^k$ is $\mathcal{C}^k$ if all higher partial derivatives

$$\frac{\partial^k f_i}{\partial x_{j_1} \partial x_{j_2} \cdots \partial x_{j_k}}$$

exist and are continuous on U.

Theorem 1.9. *Let $f\colon U \to \mathbb{R}^n$ be a $\mathcal{C}^k$-function on an open set $U \subseteq \mathbb{R}^m$ for $k \geq 2$. Let $j_1, \dots, j_k \in \{1, \dots, n\}$. If σ is a permutation of $\{1, \dots, k\}$, then*

$$\frac{\partial^k f}{\partial x_{j_1} \cdots \partial x_{j_k}} = \frac{\partial^k f}{\partial x_{j_{\sigma 1}} \cdots \partial x_{j_{\sigma k}}}.$$

For $k = 2$, this theorem says that

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

2. The chain rule

Theorem 2.1. *Let $U \subseteq \mathbb{R}^m$ and $V \subseteq \mathbb{R}^n$ be open subsets. Let $f\colon U \to \mathbb{R}^n$ and $g\colon V \to \mathbb{R}^l$ be functions such that $f(U) \subseteq V$ so that $g \circ f\colon U \to \mathbb{R}^l$ is well-defined.*

If g and f are differentiable, then so is $g \circ f$ and we have

$$D_x(g \circ f) = D_{f(x)} g \circ D_x f.$$

2.2. In terms of Jacobi matrices, the previous theorem says that Jacobi matrix of $g \circ f$ is the multiplication of the Jacobi matrix of g with the Jacobi matrix of f

$$J_x(g \circ f) = J_{f(x)} g \circ J_x f.$$

3. The inverse function theorem

Definition 3.1. Let U and V be open subsets of $\mathbb{R}^n$. A function $\varphi : U \to V$ is a *diffeomorphism* if φ is bijective and both the function φ and its inverse φ^{-1} are differentiable. We say that φ is a $\mathcal{C}^k$*-diffeomorphism* if φ is bijective and both the function φ and its inverse φ^{-1} are $\mathcal{C}^k$-functions.

Proposition 3.2. *If $\varphi\colon U \to V$ is a diffeomorphism, then for every $x \in U$, the derivative $D_x f\colon \mathbb{R}^n \to \mathbb{R}^n$ is invertible and*

$$(D_x \varphi)^{-1} = D_{\varphi(x)}(\varphi^{-1}).$$

Proof. □

Diffeomorphisms are a very important class of functions and we would like to characterize them. In order to do that, we first introduce a notion close to that of diffeomorphism.

Definition 3.3. Let $\varphi\colon U \to \mathbb{R}^n$ a function, where U is an open subset of $\mathbb{R}^n$. We say that φ is a *local diffeomorphism* (resp. a *local $\mathcal{C}^k$ diffeomorphism*) if for every $x \in U$, there is an open neighborhood B of x such that $\varphi\colon B \to \varphi(B)$ is a diffeomorphism (resp. a $\mathcal{C}^k$-diffeomorphism). Note that in this case the set $\varphi(B)$ is open.

As we shall see soon, local diffeomorphisms have a nice characterization. First, let us see how they relate to diffeomorphism.

Lemma 3.4. *A function $f\colon U \to \mathbb{R}^n$, where U is an open subset of $\mathbb{R}^n$. If the following condition are satisfied*

- *f is injective,*
- *f is a local diffeomorphism (resp. a local $\mathcal{C}^k$-diffeomorphism),*

then f defines a diffeomorphism (resp. a $\mathcal{C}^k$-diffeormorphism) onto its image

$$f\colon U \to f(U).$$

Sketch of proof. This follows trivially from the fact that the derivability is a local notion. □

The following very important theorem gives a characterization of local diffeomorphisms.

Theorem 3.5 (Inverse function theorem)**.** *Let $E \subseteq \mathbb{R}^n$ be an open subset and let $f\colon E \to \mathbb{R}^n$ be a $\mathcal{C}^k$-function on U (with $k \geq 1$). If for $x \in U$, the derivative $D_x f$ is invertible, then there exists and open subset U containing x, such that $f(U)$ is open and the restriction of f to U*

$$f|_U\colon U \to f(U)$$

is a $\mathcal{C}^k$-diffeomorphism.

In other words, the inverse function theorem says that in order to show that f is a $\mathcal{C}^k$ local diffeomorphism, it suffices to proves that f is $\mathcal{C}^k$ with $k \geq 1$ and that the linear approximation of f (i.e. the derivative of f) is invertible.

Corollary 3.6. *A function $f\colon U \to \mathbb{R}^n$, with U open subset of $\mathbb{R}^n$ is a local $\mathcal{C}^k$ diffeomorphism (with $k \geq 1$) if and only if it is $\mathcal{C}^k$ on U and its derivative $D_x f$ is invertible for every $x \in U$.*

Email address: `l.s.guetta@uu.nl`

Mathematical Institute, Utrecht University, Hans Freudenthalgebouw, Budapeestlaan 6, 3584 CD Utrecht