UMA204: Introduction to Basic Analysis

Naman Mishra

January 2024

Contents

The Course

Lecture 01.
Mon 01 Jan '24

Instructor: Prof. Purvi Gupta
Office: L-25
Office hours: Wed 17:00–18:00
Lecture hours: MW 12:00–12:50, Thu 9:00–9:50
Tutorial hours: Fri 12:00–12:50

We assume the following.

- Basics of set theory
- Existence of $\mathbb{N} = \{0, 1, 2, \ldots\}$ with the usual operations $+$ and $\cdot$

For a recap, refer lectures 1 to 3 of UMA101.

Chapter I

Number Systems

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$$

I.1 The Naturals

(Recall from UM101) $\mathbb{N}$ is the unique minimal inductive set granted by the ZFC axioms. Addition and multiplication are defined by the recursion principle and we showed that they

- are associative and commutative,
- admit identity elements 0 and 1 respectively,
- satisfy the distributive law,
- satisfy cancellation laws,
- but do not admit inverses.

I.2 Relations

(Recall) A relation on a set A is a subset $R \subseteq A \times A$. We write $a\ R\ b$ to denote $(a, b) \in R$.

Definition I.2.1 (Partial order)**.** A relation R on A is called a *partial order* if it is

- reflexive – $a \ R \ a$ for all $a \in A$;
- antisymmetric – if $a \ R \ b$ and $b \ R \ a$ then $a = b$ for all $a, b \in A$;
- transitive – if $a \ R \ b$ and $b \ R \ c$ then $a \ R \ c$ for all $a, b, c \in A$.

Additionally, if for all $x, y \in A$, $x \ R \ y$ or $y \ R \ x$, then R is called a *total order*.

A set A equipped with a partial order $\leq$ is called a *partially ordered set* (or *poset*).
A set A equipped with a total order $\leq$ is called a *totally ordered set* or simply an *ordered set*.

Examples.

- $(\mathbb{N}, \leq)$ where we say that $a \leq b$ if $\exists \, c \in \mathbb{N}$ such that $a + c = b$.
- $(\mathbb{N}, |)$ where we say that $a \mid b$ if $\exists \, c \in \mathbb{N}$ such that $a \cdot c = b$.

In UMA101, we defined order slightly differently, where we said that either $a < b$ or $b < a$ but never both. This is a "strict order". We will denote a weak partial order by $\leq$ and a strict partial order by $<$. (the notation is suggestive of how to every order there is a corresponding strict order and vice versa).

Definition I.2.2 (Equivalence)**.** An *equivalence relation* on a set A is a relation R satisfying

- reflexivity;
- symmetry – if $a \ R \ b$ then $b \ R \ a$ for all $a, b \in A$;
- transitivity.

Notation. We write $[x]_R$ to denote the set $\{y \in A \mid x \ R \ y\}$.

Proposition I.2.3. *The collection $\mathscr{A} = \{[x]_R \mid x \in A\}$ partitions A under any equivalence relation R on A.*

Proof. For every $x \in A$, $x \in [x]_R$ and so $\bigcup \mathscr{A} = A$.

Let $[x]_R \cap [y]_R \neq \varnothing$, where $x, y \in A$. Then there exists $z \in A$ such that $x \ R \ z$ and $y \ R \ z$, from which it follows that $x \ R \ y$ and $[x]_R = [y]_R$. $\square$

I.3 The Integers

We cannot solve $3 + x = 2$ in $\mathbb{N}$. We introduce $\mathbb{Z}$ to solve this problem.

Consider the relation R on $\mathbb{N} \times \mathbb{N}$ given by

$$(a, b)\ R\ (c, d) \iff a + d = b + c.$$

(check that this is an equivalence relation trivial).

Definition I.3.1. We define $\mathbb{Z}$ to be the set of equivalence classes of R, denoted $(\mathbb{N} \times \mathbb{N})/R$.

Further, define

Definition I.3.2.

- $[(a, b)] +_{\mathbb{Z}} [(c, d)] := [(a + c, b + d)]$;
- $[(a, b)] \cdot_{\mathbb{Z}} [(c, d)] := [(ac + bd, ad + bc)]$.
- $z_1 \leq_{\mathbb{Z}} z_2$ iff there exists $n \in \mathbb{N}$ such that $z_1 +_{\mathbb{Z}} [(n, 0)] = z_2$ (alternatively, $[(a, b)] \leq_{\mathbb{Z}} [(c, d)]$ iff $a + d \leq b + c$).

We need to check that these are well-defined. What does this mean? Consider

$$[(1, 2)] +_{\mathbb{Z}} [(3, 4)] = [(4, 6)]$$
$$[(3, 4)] +_{\mathbb{Z}} [(3, 4)] = [(6, 8)]$$

Our definition must ensure that $[(4, 6)] = [(6, 8)]$.
In general, the definitions are well-defined if they are independent of the choice of representatives. Throughout this section, we will omit the parentheses in $[(a, b)]$ and write it as $[a, b]$.

Lecture 02.
Tue 02 Jan '24

Proposition I.3.3. *The operations $+_{\mathbb{Z}}$, $\cdot_{\mathbb{Z}}$ and the relation $\leq_{\mathbb{Z}}$ are well-defined.*

Proof. Suppose $x = [a, b] = [a', b']$ and $y = [c, d] = [c', d']$. Then

$$\begin{aligned} a + b' &= a' + b \\ c + d' &= c' + d \\ (a + c) + (b' + d') &= (a' + c') + (b + d) \\ (a + c, b + d) &\; R \; (a' + c', b' + d') \\ [a + c, b + d] &= [a' + c', b' + d'] \end{aligned}$$

Since $\leq_{\mathbb{Z}}$ is defined in terms of $+_{\mathbb{Z}}$, it is also well-defined. For multiplication,

$$\begin{aligned} (a + b')c + (a' + b)d &= (a' + b)c + (a + b')d \\ (ac + bd) + (a'd + b'c) &= (a'c + b'd) + (ad + bc) \\ [ac + bd, ad + bc] &= [a'c + b'd, a'd + b'c] \end{aligned}$$

and symmetrically

$$[a'c + b'd, a'd + b'c] = [a'c' + b'd', a'c' + b'd']$$

so by transitivity

$$[ac + bd, ad + bc] = [a'c' + b'd', a'c' + b'd']$$

□

Proposition I.3.4. *The relation $\leq_{\mathbb{Z}}$ is a total order on $\mathbb{Z}$.*

Proof. Let $x = [a, b], y = [c, d] \in \mathbb{Z}$. Since $x +_{\mathbb{Z}} [0, 0] = [a + 0, b + 0] = x$, $x \leq_{\mathbb{Z}} x$.

Suppose $x \leq_{\mathbb{Z}} y$ and $y \leq_{\mathbb{Z}} x$. Then there exist $m, n \in \mathbb{N}$ such that $x + [m, 0] = y$ and $y +_{\mathbb{Z}} [n, 0] = x$. Thus $x +_{\mathbb{Z}} [m, 0] +_{\mathbb{Z}} [n, 0] = [a + m + n, b] = [a, b]$. This gives $a + m + n + b = a + b$ and so $m + n = 0$. This can only be when $m = n = 0$ and so $x = y$.

Now suppose $x \leq_{\mathbb{Z}} y$ and $y \leq_{\mathbb{Z}} z$. Then there exist $m, n \in \mathbb{N}$ such that $x + [m, 0] = y$ and $y +_{\mathbb{Z}} [n, 0] = z$. This immediately gives $x + [m + n, 0] = z$ and so $x \leq_{\mathbb{Z}} z$.

For trichotomy, note that either $a + d \leq b + c$ or $b + c \leq a + d$ by trichotomy of $(\mathbb{N}, \leq)$. In the first case, $a + d + n = b + c$ for some $n \in \mathbb{N}$, so $[a, b] +_{\mathbb{Z}} [n, 0] = [c, d]$. Thus $x \leq_{\mathbb{Z}} y$. Similarly, in the second case, $y \leq x$. □

Definition I.3.5 (Ring)**.** A *ring* is a set S with two binary operations $+$ and $\cdot$ such that for all $a, b, c \in S$,

(R1) addition is associative,

(R2) addition is commutative,

(R3) there exists an additive identity 0,

(R4) there exists an additive inverse $-a$,

(R5) multiplication is associative,

(R6) there exists a multiplicative identity 1,

(R7) multiplication is distributive over addition (on both sides).

For a *commutative ring*, we require additionally that

(CR1) multiplication is commutative.

Note that inverses are unique, since if $a + b = 0$ and $a + b' = 0$, then $b = (b' + a) + b = b' + (a + b) = b'$.

Definition I.3.6 (Ordered Ring)**.** An *ordered ring* is a ring S with a total order $\leq$ such that for all $a, b, c \in S$,

(OR1) $a \leq b$ implies $a + c \leq b + c$,

(OR2) $0 \leq a$ and $0 \leq b$ implies $0 \leq ab$.

Theorem I.3.7.

- $(\mathbb{Z}, +_{\mathbb{Z}}, \cdot_{\mathbb{Z}}, \leq_{\mathbb{Z}})$ *is an ordered (commutative) ring.*
- *The map* $f = n \mapsto [n, 0]$ *from* $\mathbb{N}$ *to* $\mathbb{Z}$ *is an injective map that respects* $+$, $\cdot$ *and* $\leq$. *That is, for all* $n, m \in \mathbb{N}$,

  (i) $f(n + m) = f(n) +_{\mathbb{Z}} f(m)$,

  (ii) $f(nm) = f(n) \cdot_{\mathbb{Z}} f(m)$,

  (iii) $n \leq m \iff f(n) \leq_{\mathbb{Z}} f(m)$.

  In other words, f *is an isomorphism onto a subset of* $\mathbb{Z}$.

Proof. For the first part of the theorem, we check all commutative ring axioms. We omit the subscripts on $+$ and $\cdot$ for brevity.

(R1) Addition is associative:

$$\begin{aligned}\big([a,b]+[c,d]\big)+[e,f] &= [a+c,b+d]+[e,f]\\ &= [a+c+e,b+d+f]\\ &= [a,b]+[c+e,d+f]\\ &= [a,b]+\big([c,d]+[e,f]\big)\end{aligned}$$

(R2) Addition is commutative: immediate from commutativity of $+$ on $\mathbb{N}$.

(R3) Additive identity: $[a,b]+[0,0]=[a+0,b+0]=[a,b]$.

(R4) Additive inverse: $[a,b]+[b,a]=[a+b,b+a]=[0,0]$ since $a+b+0=b+a+0$.

(R5) Multiplication is associative:

$$\begin{aligned}\big([a,b]\cdot[c,d]\big)\cdot[e,f] &= [ac+bd,ad+bc]\cdot[e,f]\\ &= [ace+bde+adf+bcf,ade+bce+acf+bdf]\\ &= [a(ce+df)+b(cf+de),a(cf+de)+b(ce+df]\\ &= [a,b]\cdot[ce+df,cf+de]\\ &= [a,b]\cdot\big([c,d]\cdot[e,f]\big)\end{aligned}$$

(R6) Multiplicative identity: $[a,b]\cdot[1,0]=[a,b]$.

(R7) Multiplication distributes over addition:

$$\begin{aligned}[a,b]\cdot\big([c,d]+[e,f]\big) &= [a,b]\cdot[c+e,d+f]\\ &= [ac+ae+bd+bf,ad+af+bc+be]\\ &= [ac+bd,ad+bc]+[ae+bf,af+be]\\ &= [a,b]\cdot[c,d]+[a,b]\cdot[e,f]\end{aligned}$$

Distributivity on the other side follows from commutativity proved below.

For commutativity of multiplication,

$$\begin{aligned}[a,b]\cdot[c,d] &= [ac+bd,ad+bc]\\ &= [ca+db,cb+da]\\ &= [c,d]\cdot[a,b]\end{aligned}$$

(OR1) follows immediately from the definition. For (OR2), suppose $0 \leq x, y \in \mathbb{Z}$. Then $x = [n, 0]$ and $y = [m, 0]$ for some $n, m \in \mathbb{N}$. Thus $xy = [nm, 0]$ and so $0 \leq xy$.

The second part is again yawningly brute force.

(i) $f(n + m) = [n + m, 0] = [n, 0] + [m, 0] = f(n) +_{\mathbb{Z}} f(m)$.

(ii) $f(nm) = [nm, 0] = [n, 0] \cdot [m, 0] = f(n) \cdot_{\mathbb{Z}} f(m)$.

(iii) $n \leq m \iff \exists\, k \in \mathbb{N}(n + k = m) \iff \exists\, k \in \mathbb{N}\big([n, 0] + [k, 0] = [m, 0]\big) \iff f(n) \leq_{\mathbb{Z}} f(m)$. □

Thus, we may view $(\mathbb{N}, +, \cdot, \leq)$ as a subset of $(\mathbb{Z}, +_{\mathbb{Z}}, \cdot_{\mathbb{Z}}, \leq_{\mathbb{Z}})$, denote $[n, 0]$ as n and drop $\mathbb{Z}$ in the subscript. We further define $-[a, b] \coloneqq [b, a]$ and $z_1 - z_2 \coloneqq z_1 + (-z_2)$.

Moreover, we have the following properties.

Proposition I.3.8.

- *There are no zero divisors in $\mathbb{Z}$. That is, for all $x, y \in \mathbb{Z}$, $xy = 0$ implies $x = 0$ or $y = 0$.*
- *The cancellation laws hold: for all $x, y, z \in \mathbb{Z}$, $x + y = x + z$ implies $y = z$, and $xy = xz$ implies $x = 0$ or $y = z$.*
- *(trichotomy) For all $z \in \mathbb{Z}$, $z = n$ or $z = -n$ for some $n \in \mathbb{N}$.*

Proof.

- From trichotomy proven below, we have $x = n$ or $x = -n$ and $y = m$ or $y = -m$ for some $n, m \in \mathbb{N}$. In any case $xy = nm$ or $xy = -nm$. Since there are no zero divisors in $\mathbb{N}$, $xy = 0$ implies $n = 0$ or $m = 0$, which in turn implies $x = 0$ or $y = 0$.
- The first cancellation law follows from the fact that additive inverses exist. For the second, note that $xy = xz \iff x(y - z) = 0$ and invoke the fact that there are no zero divisors.

  Here we have also used that $-xz = x(-z)$, since $-\tilde{z} = -1 \cdot \tilde{z}$ for all $\tilde{z} \in \mathbb{Z}$, and multiplication is associative and commutative.
- Let $z = [a, b]$. From trichotomy of $\leq$ on $\mathbb{N}$ we know that either $a + n = b$ or $a = b + n$ for some $n \in \mathbb{N}$. That is, either $z = [0, n] = -n$, or $z = [n, 0] = n$. □

I.4 The Rationals

We cannot solve $3x = 2$ in $\mathbb{Z}$.

Proof. For $x \leq 0$, $3x \leq 0 < 2$. For $x \geq 1$, $3x \geq 3 > 2$. $\square$

We define $\mathbb{Z}^*$ to be $\mathbb{Z} \setminus \{0\}$ and define the relation R on $\mathbb{Z} \times \mathbb{Z}^*$ by $(a, b)R(c, d)$ if $ad = bc$. Then R is an equivalence relation on $\mathbb{Z} \times \mathbb{Z}^*$.

Definition I.4.1. We define $\mathbb{Q}$ to be the set of equivalence classes of R, denoted $(\mathbb{Z} \times \mathbb{Z}^*)/R$.

We define operations $+_{\mathbb{Q}}$ and $\cdot_{\mathbb{Q}}$ on $\mathbb{Q}$ by

$$\begin{aligned}
[(a, b)] +_{\mathbb{Q}} [(c, d)] &\coloneqq [(ad + bc, bd)] \\
[(a, b)] \cdot_{\mathbb{Q}} [(c, d)] &\coloneqq [(ac, bd)]
\end{aligned}$$

Since there are no zero divisors in $\mathbb{Z}$, $bd \neq 0$.

We define an order $\leq_{\mathbb{Q}}$ on $\mathbb{Q}$ by

$$[(a, b)] \leq_{\mathbb{Q}} [(c, d)] \iff (ad - bc)bd \leq 0.$$

We will again omit the parentheses in this section.

Proposition I.4.2. *The operations $+_{\mathbb{Q}}$, $\cdot_{\mathbb{Q}}$ and the relation $\leq_{\mathbb{Q}}$ are well-defined.*

Proof. Suppose $[a, b] = [a', b']$ and $[c, d] = [c', d']$. Then

$$\begin{aligned}
ab' &= a'b \\
cd' &= c'd \\
(ad + bc)(b'd') &= (a'd' + b'c')(bd) \\
[ad + bc, bd] &= [a'd' + b'c', b'd']
\end{aligned}$$

For multiplication,

$$\begin{aligned}
(ac)(b'd') &= (a'c')(bd) \\
[ac, bd] &= [a'c', b'd']
\end{aligned}$$

For order,

$$\begin{aligned}
&(ad-bc)bd \le 0\\
\iff\ &(b'd')(ad-bc)bd(b'd') \le 0\\
\iff\ &(ab'dd'-bb'cd')bdb'd' \le 0\\
\iff\ &(a'bdd'-bb'c'd)bdb'd' \le 0\\
\iff\ &(bd)^2(a'd'-b'c')b'd' \le 0\\
\iff\ &(a'd'-b'c')b'd' \le 0
\end{aligned}$$

since $bd \neq 0 \neq b'd'$. Thus $+_{\mathbb{Q}}$, $\cdot_{\mathbb{Q}}$ and $\le_{\mathbb{Q}}$ are well-defined. □

Proposition I.4.3. *The relation $\le_{\mathbb{Q}}$ is a total order on $\mathbb{Q}$.*

Proof. **Transitivity:** Suppose $(ad-bc)bd \le 0$ and $(cf-de)df \le 0$. Then $(adf-bcf)bdf \le 0$ and $(bcf-bde)bdf \le 0$. Adding these gives $(adf-bde)bdf \le 0$ and so $(af-be)bf \le 0$.

Antisymmetry: Suppose $(ad-bc)bd \le 0$ and $(cb-da)db \le 0$. Then $(ad-bc)bd = 0$ which gives $ad = bc$ so $x = y$. □

Theorem I.4.4.

- $(\mathbb{Q}, +_{\mathbb{Q}}, \cdot_{\mathbb{Q}}, \le_{\mathbb{Q}})$ *is an ordered field.*
- *The map $f = z \mapsto [z,1]$ from $\mathbb{Z}$ to $\mathbb{Q}$ is an injective map that respects $+$, $\cdot$ and $\le$. That is, for all $z_1, z_2 \in \mathbb{Z}$,*

  (i) $f(z_1+z_2) = f(z_1) +_{\mathbb{Q}} f(z_2)$,

  (ii) $f(z_1 z_2) = f(z_1) \cdot_{\mathbb{Q}} f(z_2)$,

  (iii) $z_1 \le z_2 \iff f(z_1) \le_{\mathbb{Q}} f(z_2)$.

  In other words, f is a commutative ring isomorphism into $\mathbb{Q}$.

Proof. For the first part, we check all ordered field axioms. We again omit the subscripts on $+$ and $\cdot$ for brevity. Numbering is from UMA101.

(F1) $+$ and $\cdot$ are commutative: immediate from commutativity of $+$ and $\cdot$ on $\mathbb{Z}$.

(F2) $+$ and $\cdot$ are associative:

$$\begin{aligned}\big([a,b]+[c,d]\big)+[e,f] &= [ad+bc,bd]+[e,f]\\ &= [(ad+bc)f+bde,bdf]\\ &= [adf+b(cf+de),bdf]\\ &= [a,b]+[cf+de,df]\\ &= [a,b]+\big([c,d]+[e,f]\big)\end{aligned}$$

Associativity of $\cdot$ is immediate from associativity on $\mathbb{Z}$.

(F3) Distributivity:

$$\begin{aligned}[a,b]\cdot\big([c,d]+[e,f]\big) &= [a,b]\cdot[cf+de,df]\\ &= [acf+ade,bdf]\\ &= [abcf+abde,b^2df] \qquad (b \text{ is nonzero})\\ &= [(ac)(bf)+(bd)(ae),(bd)(bf)]\\ &= [ac,bd]+[ae,bf]\end{aligned}$$

(F4) Identities: $[0,1]\neq[1,1]$, $[a,b]+[0,1]=[a,b]$ and $[a,b]\cdot[1,1]=[a,b]$.

(F5) Additive inverse: $[a,b]+[-a,b]=[0,1]$.

(F6) Multiplicative inverse: $[a,b]\cdot[b,a]=[1,1]$ for $a\neq 0 \iff [a,b]\neq[0,1]$.

For the second part,

(i) $f(z_1+z_2)=[z_1+z_2,1]=[z_1,1]+[z_2,1]$.

(ii) $f(z_1z_2)=[z_1z_2,1]=[z_1,1]\cdot[z_2,1]$.

(iii) $f(z_1)\leq f(z_2) \iff (z_1-z_2)\leq 0 \iff z_1\leq z_2$. □

We now introduce the division operation $/:\mathbb{Q}\times\mathbb{Q}^*\to\mathbb{Q}$ by $a/b=\frac{a}{b}=ab^{-1}$.

Notation. Note that every rational number $x=[a,b]$ can be written as $x=a/b$. We thus largely drop the notation $[a,b]$ and write a/b instead.

We will now accept basic algebraic manipulations of rational numbers without justification.

Lecture 03.
Wed 03 Jan '24

Definition I.4.5 (Exponentiation)**.** The recursion principle guarantees the existence of pow : $\mathbb{Z} \times \mathbb{N} \to \mathbb{N}$ such that for all $n, m \in \mathbb{N}$,

$$\begin{aligned} \mathrm{pow}(m, 0) &= 1 \\ \mathrm{pow}(m, n+1) &= m \cdot \mathrm{pow}(m, n) \end{aligned}$$

We extend this to pow : $\mathbb{Q}^* \times \mathbb{Z} \to \mathbb{Q}$ as follows.

$$\mathrm{pow}\left(\frac{a}{b}, m\right) := \begin{cases} a^m/b^m & \text{if } m \in \mathbb{N} \\ b^m/a^m & \text{if } -m \in \mathbb{N} \end{cases}$$

We write z^n to denote $\mathrm{pow}(z, n)$.

Remark. Note that we have defined 0^0 to be 1, but we don't really care.

Proposition I.4.6. *Expoentiation is well-defined.*

Proof. Let $a/b = \tilde{a}/\tilde{b} \in \mathbb{Q}$. That is, $a\tilde{b} = b\tilde{a} \in \mathbb{Z}$. For $m \in \mathbb{N}$, thus $a^m\tilde{b}^m = b^m\tilde{a}^m$ (easily proved by induction).

Similarly if $-m \in \mathbb{N}$. □

Theorem I.4.7. *There is no rational whose square is* 2*.*

We first make note of the following lemma.

Lemma I.4.8. *Let $x \in \mathbb{Q}$. Then there exists $p \in \mathbb{Z}$, $q \in \mathbb{N}^*$ such that $x = p/q$.*

In particular, if $x > 0$, then $x = p/q$ for some $p \in \mathbb{N}$, $q \in \mathbb{N}^*$.

Proof. Let $x = a/b$. If $b \in \mathbb{N}$, we are done. Otherwise, $x = -a/-b$ and $-b \in \mathbb{N}$. □

We will make use of the well-ordered property of $(\mathbb{N}, \leq)$ proved below in theorem I.4.9.

Proof of theorem I.4.7. Suppose there exists such an x. By the field properties, $(-x)^2 = x^2$. Thus we may assume $x \geq 0$. Let $x = p/q$ for some $q \in \mathbb{N}^*$. Since $x \geq 0$, we have $p \geq 0 \iff p \in \mathbb{N}$.

Let $A = \{q \in \mathbb{N}^* \mid x = p/q \text{ for some } p \in \mathbb{N}\}$. A is non-empty.

By the well-ordering principle, A has a least element q_0. Let $p_0 \in \mathbb{N}$ such that $x = p_0/q_0$.

We know that $1 < x < 2$ [why? because $(\cdot)^2$ is an increasing function on positive rationals (why? difference of squares)] and so $0 < p_0 - q_0 < q_0$. Now

$$\begin{aligned}\frac{2q_0 - p_0}{p_0 - q_0} &= \frac{2-x}{x-1} \\ &= \frac{(2-x)(x+1)}{x^2-1} \\ &= 2x + 2 - x^2 - x \\ &= x,\end{aligned}$$

in contradiction to the minimality of q_0. □

Theorem I.4.9 (Well-ordering principle)**.** *Every non-empty subset of* $\mathbb{N}$ *has a least element.*

Proof. Let $S \subseteq \mathbb{N}$ be non-empty. We define $P(n)$ to be "if $n \in S$, then S has a least element". Clearly $P(0)$ holds.

Suppose $P(k)$ holds for all $k \leq n \in \mathbb{N}$.
If $n+1 \notin S$, $P(n+1)$ holds vacuously.
If $\exists\, m \in S(m < n+1)$, then $P(n+1)$ holds by virtue of $P(m)$.
Otherwise $n+1 \in S$ and $\forall\, m \in S(n+1 \leq m)$, so that $n+1$ is the least element of S.

In any case, $P(n+1)$ holds. □

Theorem I.4.10. *Let*

$$\begin{aligned}A &= \left\{x \in \mathbb{Q} \mid x^2 < 2\right\} \\ B &= \left\{x \in \mathbb{Q} \mid x^2 > 2, x > 0\right\}\end{aligned}$$

Then A has no largest element and B has no smallest element.

Proof. Let $a \in A$. $a > -2$ since otherwise $a^2 \geq 4$. Let $c = a + \frac{2-a^2}{2+a}$. Clearly

$c > a$. Now

$$c = \frac{2a+2}{2+a}$$
$$c^2 = \frac{4a^2+8a+4}{4+4a+a^2}$$
$$c^2 - 2 = \frac{2a^2-4}{(2+a)^2} < 0$$

Thus $c \in A$.

For B, let $c = b + \frac{2-b^2}{2+b} = \frac{2b+2}{2+b}$. Clearly $0 < c < b$ and $c^2 - 2 = \frac{2b^2-4}{(2+b)^2} > 0$. Thus $c \in B$. $\square$

Corollary I.4.11. $(\mathbb{Q}, \leq)$ *does not have the least upper bound property.*

Proof. Let b be an upper bound of A. Clearly $b > 0$. b cannot be in A since A has no largest element. b cannot have square 2 by theorem I.4.7. Thus $b \in B$. But since B has no smallest element, there is a $b' \in B$ which is less than b.

For any $a \in A$, if $a < 0$ then $a < b'$. Otherwise, $0 < (b')^2 - a^2 = (b'-a)(b'+a)$ and so $a < b'$. Thus b' is an upper bound of A which is less than b.

Since b was arbitrary, A cannot have a least upper bound. $\square$

I.5 Ordered Fields with LUB

(Recall from UMA101 Lecture 6) Given an ordered set $(X, \leq)$, a subset $S \subseteq X$ is said to be *bounded above* (resp. below) if there exists $x \in X$ such that for all $s \in S$, $s \leq x$ (resp. $x \leq s$), and any such x is called an *upper* (resp. lower) *bound* of S.

A (The) *supremum* or least upper bound of S is an element $x \in X$ such that x is an upper bound of S and for all upper bounds y of S, $x \leq y$. Similarly, infimum or greatest lower bound.

$(X, \leq)$ is said to have the least upper bound property if every non-empty bounded above subset of X admits a supremum.

Theorem I.5.1. *Every ordered field F "contains" $\mathbb{Q}$, i.e., there exists an injective map $f : \mathbb{Q} \to F$ that respects $+$, $\cdot$ and $\leq$.*

We will notate this statement as $\mathbb{Q} \subseteq F$.

Proof. Let $f : \mathbb{Z} \to F$ be defined as

$$f(n) = \begin{cases} 0_F & \text{if } n = 0 \\ \underbrace{1_F + \cdots + 1_F}_{n \text{ times}} & \text{if } n > 0 \\ \underbrace{(-1_F) + \cdots + (-1_F)}_{m \text{ times}} & \text{if } n = -m, m > 0 \end{cases}$$

Note that $f(-n) = -f(n)$ for all $n \in \mathbb{N}$. Let us show that $f(n+m) = f(n) + f(m)$ for all $n, m \in \mathbb{Z}$.

Case 1: $n = 0$ or $m = 0$. Immediate.

Case 2: $n > 0$ and $m > 0$. Then

$$\begin{aligned} f(n+m) &= \underbrace{1_F + \cdots + 1_F}_{n+m \text{ times}} \\ &= \underbrace{1_F + \cdots + 1_F}_{n \text{ times}} + \underbrace{1_F + \cdots + 1_F}_{m \text{ times}} \\ &= f(n) + f(m) \end{aligned}$$

Case 3: $n < 0$ and $m < 0$. Then $f(n+m) = -f((-n)+(-m)) = -(f(-n)+f(-m)) = f(n) + f(m)$.

Case 4: $nm < 0$. WLOG, let $m < 0 < n$. Suppose $0 < n + m$. Then $f(n+m) + f(-m) = f(n+m-m) = f(n)$ from case 2. Now suppose $n + m < 0$. Then $f(n) + f(-n-m) = f(n-n-m) = -f(m)$ from case 3. In either case, $f(n+m) = f(n) + f(m)$.

Now consider $f(nm)$. If $nm = 0$, then $f(nm) = 0_F = f(n)f(m)$. If $0 < n, m$, then

$$
\begin{aligned}
f(nm) &= \overbrace{1_F + \cdots + 1_F}^{nm \text{ times}} \\
&= \underbrace{\overbrace{(1_F + \cdots + 1_F)}^{n \text{ times}} + \cdots + \overbrace{(1_F + \cdots + 1_F)}^{n \text{ times}}}_{m \text{ times}} \\
&= \underbrace{(1_F + \cdots + 1_F)}_{n \text{ times}} \cdot \underbrace{(1_F + \cdots + 1_F)}_{m \text{ times}} \\
&= f(n)f(m)
\end{aligned}
$$

If either of n, m is negative, then we take the negative sign out and use the above case.

Thus f respects $+$ and $\cdot$.

Suppose that $m < n$. Then $f(n) - f(m) = f(n) + f(-m) = f(n - m) = (n - m)1_F$ (where $z1_F$ is notation for 1_F added z times). $n - m$ is positive, but 1_F added to itself a positive number of times must be positive. This is because $0_F < 1_F$ (UMA101) and so $k1_F < (k + 1)1_F$ for all $k \in \mathbb{N}^+$. Thus $f(m) < f(n)$ and so f respects $<$ (and hence $\leq$).

Finally, injectivity of f follows from order preservation.

We extend f to $\mathbb{Q}$ by defining $f(a/b) = f(a)f(b)^{-1}$. This continues to be an isomorphism.

Why? First note that

$$f(ka/kb) = f(ka)f(kb)^{-1} = f(k)f(a)(f(k)f(b))^{-1} = f(a)f(b)^{-1} = f(a/b),$$

so that f is well-defined. Now

$$
\begin{aligned}
f(a/b + c/d) &= f\Big(\frac{ad + bc}{bd}\Big) \\
&= f(ad + bc)f(bd)^{-1} \\
&= (f(a)f(d) + f(b)f(c))(f(b)f(d))^{-1} \\
&= f(a)f(b)^{-1} + f(c)f(d)^{-1} \\
&= f(a/b) + f(c/d)
\end{aligned}
$$

and

$$\begin{aligned} f(a/b \cdot c/d) &= f(ac/bd) \\ &= f(ac)f(bd)^{-1} \\ &= f(a)f(b)^{-1}f(c)f(d)^{-1} \\ &= f(a/b)f(c/d). \end{aligned}$$

Finally

$$\begin{aligned} f(a/b) \leq f(c/d) &\iff f(a)f(b)^{-1} \leq f(c)f(d)^{-1} \\ &\iff f(a)f(b)f(d)^2 \leq f(c)f(d)f(b)^2 \\ &\iff (f(ad) - f(bc))f(b)f(d) \leq 0 \\ &\iff (ad - bc)bd \leq 0 \\ &\iff a/b \leq c/d. \end{aligned}$$ $\square$

Assignment 1

Quiz 12 Jan 2024

Problem 1.1. *Let* $(\mathbb{Z}, +_{\mathbb{Z}}, \cdot_{\mathbb{Z}}, \leq_{\mathbb{Z}})$ *be defined as in class. Recall that we identify* $n \in \mathbb{N}$ *with* $[(n, 0)] \in \mathbb{Z}$*. Show that any element of* $\mathbb{Z}$ *is either* m *or* $-m$ *for some* $m \in \mathbb{N}$*.*

Proof. Proved in proposition I.3.8. ☐

Problem 1.2. *Recall the construction of* $\mathbb{Q}$ *as the set of equivalence classes of the relation* R *on* $\mathbb{Z} \times \mathbb{Z} \setminus \{0\}$ *given by* $(a, b)R(c, d) \iff ad = bc$*. We say that* $[(a, b)] \leq [(c, d)]$ *if* $(bc - ad)(bd) \geq 0$*. Using only the arithmetic and order properties of integers, show that the relation* $\leq$ *is well-defined. Remember you are no allowed to divide yet!*

Proof. Proposition I.4.2. ☐

Problem 1.3. *Without assuming the existence of irrational numbers, show that*

(a) $(\mathbb{Z}, \leq)$ *has the least upper bound property.*

(b) $(\mathbb{Q}, \leq)$ *does not have the least upper bound property.*

You may directly cite any theorem(s) proved in class.

Proof.

(a) Let S be a non-empty bounded above subset of $\mathbb{Z}$. Let b be an upper bounded of S and let $f : \mathbb{Z} \to \mathbb{N}$ be as $f(x) = b - x$. By the well-ordering principle, $f(S)$ has a least element m. Then $b - m$ is the maximum of S.

(b) Corollary I.4.11. ☐

Problem 1.4. *Let* F *be an ordered field. Recall that* $\mathbb{Q} \subseteq F$*. Show that the following two statements are equivalent.*

(i) For every $a, b > 0$ *in* F*, there is an* $n \in \mathbb{N}$ *such that* $na > b$*.*

(ii) For every $a < b$ *in* F*, there is an* $r \in \mathbb{Q}$ *such that* $a < r < b$*.*

Proof. Suppose item (i) holds. Let $a < b$ in F. Then $1/(b-a) > 0$. Let $n \in \mathbb{N}$ be such that $n > 1/(b-a)$, *i.e.*, $1/n < b-a$. We first show that there is a rational at most a. If $a \geq 0$, this is trivial. Otherwise, $-a > 0$ and so by item (i), there is an $m \in \mathbb{N}$ such that $m > 1/(-a) \iff -1/m < a$. Thus the set $S = \left\{k \in \mathbb{Z} \mid k \cdot \frac{1}{n} \leq a\right\}$ is non-empty. By item (i), it is bounded above. By problem 1.3(a), it has a maximum M. Then $\frac{M}{n} \leq a < \frac{M+1}{n} \leq a + \frac{1}{n} < b$. Thus $\frac{M+1}{n}$ is the required rational.

Suppose item (ii) holds. Let $0 < a, b$. Then there exist $p \in \mathbb{N}$ and $q \in \mathbb{N}^*$ such that $0 < b/a < p/q < b/a + 1$. Since $1 \leq q$, $p/q \leq p$. Then $b < pa$ as required. $\square$

Problem 1.5. *Let F be a field. An absolute value of F is a function $A\colon F \to \mathbb{R}$ satisfying*

(1) $A(x) \geq 0$ for all $x \in F$,

(2) $A(x) = 0$ if and only if $x = 0$,

(3) $A(xy) = A(x)A(y)$ for all $x, y \in F$,

(4) $A(x+y) \leq A(x) + A(y)$ for all $x, y \in F$.

A subset $S \subseteq F$ is said to be A-bounded if there exists an $M > 0$ such that $A(s) \leq M$ for all $s \in S$. This is a way to define boundedness of sets in the absence of an order relation.

Let $p \in \mathbb{N}$ be a prime number. Define $\nu_p\colon \mathbb{Z} \to \mathbb{Z} \cup \{\infty\}$ by

$$\nu_p(n) = \begin{cases} \max\{k \in \mathbb{N} : p^k \mid n\}, & \text{if } n \neq 0, \\ \infty, & \text{if } n = 0. \end{cases}$$

Extend ν_p to $\mathbb{Q}$ by

$$\nu_p(a/b) = \nu_p(a) - \nu_p(b), \quad a, b \in \mathbb{Z}, b \neq 0.$$

Now, define $A_p\colon \mathbb{Q} \to \mathbb{R}$ by $A_p(x) = e^{-\nu_p(x)}$ if $x \neq 0$, and $A_p(0) = 0$.

(a) Show that A_p is an absolute value on $\mathbb{Q}$.

(b) Show that

$$A_p(x+y) \leq \max\{A_p(x), A_p(y)\}, \quad x, y \in \mathbb{Q}.$$

(c) Show that $\mathbb{Z}$ is A_p-bounded.

You may use basic facts about factorization without proof, but clearly state what you are using.

Proof. A_p satisfies (1) and (2) by definition.

Let $x = a/b$, $y = c/d$ in $\mathbb{Q}$. If either is zero, (3) holds trivially. Otherwise $xy = ac/bd$ with $a, b, c, d \in \mathbb{Z}^*$. Let $a = p^{\nu_p(a)}a'$, $c = p^{\nu_p(c)}c'$, where a', c' are coprime to p. Then $ac = p^{\nu_p(a)+\nu_p(c)}(a'c')$. Thus $\nu_p(ac) = \nu_p(a)+\nu_p(c)$. Similarly, $\nu_p(bd) = \nu_p(b) + \nu_p(d)$. Thus $\nu_p(xy) = \nu_p(x) + \nu_p(y)$ and so $A_p(xy) = A_p(x)A_p(y)$.

(4) follows from (b), which we prove now. If either x or y is zero, (b) holds trivially. Let

$$x = \frac{p^\alpha a}{p^\beta b}, \quad y = \frac{p^\gamma c}{p^\delta d},$$

where $a, b, c, d \in \mathbb{Z}^*$ are coprime to p. Thus $\nu_p(x) = \alpha - \beta$ and $\nu_p(y) = \gamma - \delta$. WLOG suppose that $A_p(x) \geq A_p(y) \iff \nu_p(x) \leq \nu_p(y)$ which gives $\alpha - \beta \leq \gamma - \delta$.

$$\begin{aligned} x + y &= \frac{p^{\alpha+\delta}ad + p^{\beta+\gamma}bc}{p^{\beta+\delta}bd} \\ &= \frac{p^{\alpha+\delta}(ad + p^{\beta+\gamma-\alpha-\delta}bc)}{p^{\beta+\delta}bd} \end{aligned}$$

Thus $\nu_p(x + y) \geq \alpha + \delta - \beta - \delta = \alpha - \beta$ and so $A_p(x + y) \leq A_p(x) = \max\{A_p(x), A_p(y)\}$.

(c) follows from $\nu_p(x) \geq 0$, so that $A_p(x) \leq 1$ for all $x \in \mathbb{Z}$. $\square$

Lecture 04.
Wed 10 Jan '24

Definition I.5.2 (Archimedean property)**.** An ordered field F is said to have the *Archimedean property* if for every $x, y > 0$, there exists an $n \in \mathbb{N} \subseteq F$ such that $nx > y$.

Theorem I.5.3. *$\mathbb{Q}$ has the Archimedean property.*

Proof. Let $x, y > 0$ be rationals. If $x > y$, $n = 1$ works. Suppose $x \leq y$. It suffices to show that $\exists\, n \in \mathbb{N}(nr > 1)$, where $r = x/y$. Since r is positive, we have $p, q \in \mathbb{N}^*$ such that $r = p/q$. Let $n = 2q$. This gives $nr > 1$. $\square$

Remark. Not all ordered fields have the Archimedean property.

Theorem I.5.4. *Let F be an ordered field with the LUB property. Then F has the Archimedean property.*

Proof. Let $x, y > 0$. Suppose $\forall\, n \in \mathbb{N}(nx \leq y)$. Let $A = \{nx \mid n \in \mathbb{N}\}$. Clearly A is non-empty and bounded above. Then $\sup A$ exists and so there exists an $m \in \mathbb{N}$ such that $\sup A - x < mx$. Thus $\sup A < (m+1)x \in A$, a contradiction. $\square$

Theorem I.5.5. *Let F be an ordered field with the LUB property. Then $\mathbb{Q}$ is dense in F, i.e., given $x < y \in F$, there exists a rational $r \in \mathbb{Q}$ such that $x < r < y$.*

Proof. Follows from theorem I.5.3 and problem 4 on assignment 1. $\square$

I.6 The Reals

Theorem I.6.1 (Dedekind/Cauchy)**.** *There exists a unique (up to isomorphism) ordered field with the LUB property.*

We first recover some properties of supremums.

Lemma I.6.2. *Let F be an ordered field with the LUB property. Let A and B be non-empty bounded above subsets of F. Then $\sup A + \sup B = \sup(A+B)$. Further, if all elements of A and B are non-negative, then $\sup A \sup B = \sup(AB)$.*

Here $A + B \coloneqq \{a + b \mid a \in A, b \in B\}$ and $AB \coloneqq \{ab \mid a \in A, b \in B\}$.

Proof. Let $\alpha = \sup A$ and $\beta = \sup B$. For all $a \in A$ and $b \in B$, $a+b \leq \alpha+\beta$. Thus $\alpha + \beta$ is an upper bound of $A + B$.

Let $c < \alpha + \beta$. Since $c - \beta < \alpha$, there exists an $a \in A$ larger than $c - \beta$. Then $c - a < \beta$ and so there exists a $b \in B$ larger than $c - a$. Thus $c < a + b \in A + B$ and so $\alpha + \beta = \sup(A + B)$.

Now suppose all elements of A and B are non-negative. If $\alpha = 0$ or $\beta = 0$, then $\alpha\beta = 0$ and so $\alpha\beta = \sup(AB)$.

For all $a \in A$ and $b \in B$, $ab \leq \alpha\beta$. Let $c < \alpha\beta$. Since $c/\beta < \alpha$, there exists an $a \in A$ larger than c/β. Then $c/a < \beta$ and so there exists a $b \in B$ larger than c/a. Thus $c < ab \in AB$ and so $\alpha\beta = \sup(AB)$. ☐

Proof of uniqueness. Let F and G be OFWLUB. Let h be identity on $\mathbb{Q} \subseteq F, G$. For $z \in F$ let

$$A_z = \{w \in \mathbb{Q} \mid w <_F z\}.$$

Claim: A_z is non-empty and bounded above when viewed as a subset of G, and therefore has a supremum in G.

First, A_z is non-empty by density applied to $(z - 1_F, z)$ or Archimedean applied to $-z$. Secondly, by Archimedean (or density) there exists a *rational* upper bound q of A_z in F. This q is also an upper bound of A_z in G.

By LUB, A_z has a supremum in G.

We define $h(z) \coloneqq \sup_G A_z$. For this we need to show that $h(r) = r$ for all $r \in \mathbb{Q}$, so that the definitions coincide. Let $r \in \mathbb{Q}$ so that $A_r = \{w \in \mathbb{Q} \mid w <_F r\}$. Clearly r is an upper bound of A_r in G. For any $g \in G$ less than r, there is some $q \in \mathbb{Q}$ such that $g <_G q <_G r$ (by density of $\mathbb{Q}$ in G). Thus g cannot be an upper bound of $A_r \subseteq G$. Thus $r = \sup_G A_r = h(r)$.

Claim: h preserves order.

Let $z < w \in F$. By density of $\mathbb{Q}$ in F, there exist rationals r, s, t such that $z < r < s < w$. Then $A_z \subsetneq A_w$ as subsets of F and hence of G. Thus

$$h(z) = \sup_G A_z \leq_G r < s \leq_G \sup_G A_w = h(w).$$

Claim: h preserves addition.

It is sufficient to show that $A_{x+y} = A_x + A_y$, where set addition is defined pairwise. If a rational $q \in A_x + A_y$, then clearly $q <_F x + y$ and so $q \in A_{x+y}$. Let $q \in A_{x+y} \iff q <_F x + y$. Then $q - x \in A_y$. Since A_y has no largest element (by density), there exists an $r \in A_y$ with $q - x < r < y$. Then $q - r < x$ and so $q - r \in A_x$. Thus $q = (q - r) + r \in A_x + A_y$ which gives equality of the sets.

From the previous lemma, $\sup A_x + \sup A_y = \sup(A_x + A_y) = \sup A_{x+y}$ and so h preserves addition.

Claim: h preserves multiplication.

Let $0 < x, y \in F$. Let $A_z^+ = \{w \in \mathbb{Q} \mid 0 < w <_F z\}$. We will show that $A_{xy}^+ = A_x^+ A_y^+$, where set product is defined pairwise. If a rational $q \in A_x^+ A_y^+$, then clearly $0 < q <_F xy$ and so $q \in A_{xy}^+$. Let $q \in A_{xy}^+ \iff 0 < q <_F xy$. Then $q/x \in A_y^+$. Since A_y^+ has no largest element, there exists an $r \in A_y^+$ with $q/x < r < y$. Then $q/r < x$ and so $q/r \in A_x^+$. Thus $q = (q/r) \cdot r \in A_x^+ A_y^+$ which gives equality of the sets.

From the previous lemma, $\sup A_x^+ \sup A_y^+ = \sup(A_x^+ A_y^+) = \sup A_{xy}^+$ and so h preserves multiplication of positive elements.

Since h preserves addition, h preserves additive inverses. So h preserves multiplication of all elements.

Thus h is an isomorphism between F and G. $\square$

I.6.1 Dedekind's Construction

Definition I.6.3 (Dedekind cut)**.** A *Dedekind cut* is a non-empty proper subset $A \subsetneq \mathbb{Q}$ such that

(i) if $a \in A$, then $b \in A$ for all $b \in \mathbb{Q}$ with $b < a$.

(ii) if $a \in A$, then there exists a $c \in A$ such that $a < c$.

Definition I.6.4 ($\mathbb{R}$)**.** We define

$$\mathbb{R} := \left\{A \in 2^{\mathbb{Q}} \mid A \text{ is a Dedekind cut}\right\}.$$

Further,

(i) $A \leq B \iff A \subseteq B$;

(ii) $A + B = \{a + b \mid a \in A, b \in B\}$.

(iii) for $A, B > 0$,

$$A \cdot B = \{q \in \mathbb{Q} \mid q \leq rs \text{ for some } r \in A, s \in B\}.$$

If $A < 0$ but $B > 0$, then $A \cdot B = -((-A) \cdot B)$. If $B < 0$ but $A > 0$, then $A \cdot B = -(A \cdot (-B))$. If $A < 0$ and $B < 0$, then $A \cdot B = (-A) \cdot (-B)$.

Proposition I.6.5. *$O = \{z \in \mathbb{Q} \mid z < 0\}$ is the additive identity of $\mathbb{R}$. For any $A \in \mathbb{R}$,*

$$B = \{x \in \mathbb{Q} \mid \exists\, r \in O(r - x \notin A)\}$$

is an additive inverse of A.

Proof. Let $A \in \mathbb{R}$. For all $a \in A$, there exists $a' \in A$ larger than a. So $a - a' \in O$ and thus $a' + (a - a') = a \in A + O$.

For all $a \in A + O$, there exists $a' \in A$ and $o \in O$ such that $a = a' + o$. But then $a' > a$, so $a \in A$. Thus $A + O = A$.

Let B be as defined. Let $a + b \in A + B$ where $a \in A$ and $b \in B$. Then there exists $r \in O$ such that $r - b \notin A$. So $r - b > a$ and thus $a + b < r < 0$.

Now let $o \in O$. Since O has no largest element, there exists an $o' \in O$ such that $o' > o$. Let $a \in A$. Consider the set $\alpha = \{n \in \mathbb{Z} \mid a + n(o' - o) \in A\}$. By archimedean property of $\mathbb{Q}$, α is bounded. It is obviously non-empty fucker. Thus it has a supremum n. Let $a' = a + n(o' - o)$. $a' + (o' - o) = o' - (o - a') \notin A$ because n was supremum. This gives $o - a' \in B$. Thus $o \in A + B$. $\square$

Lecture 05. Thu 11 Jan '24

Theorem I.6.6. $\mathbb{R}$ *has the least upper bound property.*

Proof. Let α be a non-empty subset of $\mathbb{R}$ that is bounded above. We claim that $S = \bigcup_{A \in \alpha} A$ is the supremum of α.

s **is a cut:** Since S is a union of a non-empty set of non-empty sets, it is non-empty. Since S is bounded above, say by some cut C, we have $S \subseteq C \subsetneq \mathbb{Q}$ and so $S \neq \mathbb{Q}$. If $a \in S$, then $a \in A$ for some $A \in \alpha$. Since A is a cut, every rational smaller than a is contained in A and thereby in S. Moreover, there exists an $a' \in A$ which is larger than a. Thus $a' \in S$ is larger than a.

upper bound: $A \subseteq S$ for all $A \in \alpha$.

least upper bound: For any $D \subsetneq S$, let $b \in S \setminus D$. But since $b \in A$ for some $A \in \alpha$, D is not an upper bound of α. ☐

Dedekind's construction is an "order completion". Thus all the order properties (LUB, density) are nice, but arithmetic is ugly.

I.6.2 Cauchy's Construction

There seem to be sequences in $\mathbb{Q}$ that "should" have a limit (*e.g.*, a monotone and bounded sequence) but do not (within $\mathbb{Q}$). We construct equivalence classes of sequences which "converge" to the same number, and define reals by those classes.

Definition I.6.7 (Sequence)**.** A sequence of rational numbers is a $f : \mathbb{N} \to \mathbb{Q}$. We usually denote $f(k)$ by a_k and call it the k-th term of the sequence. The function f is usually written as $(a_k)_{k \in \mathbb{N}}$.

Definition I.6.8. A sequence $(a_k)_{k \in \mathbb{N}} \subseteq \mathbb{Q}$ is said to be

(i) $\mathbb{Q}$-bounded if there exists an $M \in \mathbb{Q}$ such that $|a_k| \leq M$ for all $k \in \mathbb{N}$.

(ii) $\mathbb{Q}$-Cauchy if for every rational $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $|a_m - a_n| < \epsilon$ for all $m, n \geq N$.

(iii) convergent in $\mathbb{Q}$ if there exists an $L \in \mathbb{Q}$ such that for all (rational) $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ for all $n \geq N$.

Exercise I.6.9. *Show that if a sequence is convergent in $\mathbb{Q}$, then it is $\mathbb{Q}$-Cauchy, and if it is $\mathbb{Q}$-Cauchy, then it is $\mathbb{Q}$-bounded.*

Remark. From UMA101, we know that if a sequence is convergent in $\mathbb{Q}$, the limit is unique. We also know arithmetic laws of limits (which we proved over $\mathbb{R}$, but they hold over $\mathbb{Q}$as well).

Definition I.6.10. Two sequences $a = (a_n)_{n\in\mathbb{N}}$ and $b = (b_n)_{n\in\mathbb{N}}$ are said to be *equivalent* if their difference converges to 0.

Proposition I.6.11. *Let $\mathcal{C}$ denote the space of $\mathbb{Q}$-cauchy sequences. Then $\sim$ given by $a \sim b$ if a and b are equivalent (as per the previous definition) is an equivalence relation.*

Proof. Reflixivity and symmetry are immediate. Transitivity follows from the triangle inequality. □

Definition I.6.12 ($\mathbb{R}$)**.** We define

$$\mathbb{R} \coloneqq \mathcal{C}/\sim.$$

Further,

(i) $[(a_n)_n] +_{\mathbb{R}} [(b_n)_n] \coloneqq [(a_n + b_n)_n]$.

(ii) The additive identity $0 = [(0)_{n\in\mathbb{N}}]$.

(iii) $[(a_n)_n] \cdot_{\mathbb{R}} [(a_n)_n] \coloneqq [(a_n \cdot b_n)_n]$.

(iv) $[a]$ is positive iff there exists a rational $c > 0$ and an $N \in \mathbb{N}$ such that $a_n > c$ for all $n \geq N$.

Recall how we define order from a set of positive elements. Assuming the existence of additive inverses, we define $a <_R b$ iff $b - a$ is positive. The set of positive elements $\mathbb{R}^+$ must satisty (Apostol order axioms):

- $0 \notin \mathbb{R}^+$.
- For $x \neq 0$, exactly one of $x \in \mathbb{R}^+$ or $-x \in \mathbb{R}^+$ holds.
- For $x, y \in \mathbb{R}^+$, $x + y \in \mathbb{R}^+$ and $x \cdot y \in \mathbb{R}^+$.

The first two give trichotomy, and the third is the field order properties.

Proposition I.6.13. *The operations $+\mathbb{R}$ and $\cdot_{\mathbb{R}}$ and the relation $>_{\mathbb{R}}$ are well-defined.*

Proof. Let $a \sim a'$ and $b \sim b'$ be $\mathbb{Q}$-Cauchy sequences. Then $(a+b)-(a'+b') = (a-a')+(b-b') \to 0$. So $a+b \sim a'+b'$.

For multiplication,

$$\begin{aligned} ab - a'b' &= ab - ab' + ab' - a'b' \\ &= a(b-b') + b'(a-a') \\ &\to 0. \end{aligned}$$

So $ab \sim a'b'$.

Finally, for any positive c, $a_n - a'_n$ is eventually smaller than $c/2$, so $a_n > c$ implies $a'_n > c/2$ for sufficiently large n. □

Proposition I.6.14. *The relation $<_{\mathbb{R}}$ makes $\mathbb{R}$ an ordered field.*

Proof. First note that each $[x] \in \mathbb{R}$ has additive inverse $[-x]$ where each term is the additive inverse of the corresponding term in x.

$[0] \notin \mathbb{R}^+$ is obvious.

Let $[x] \neq 0$. Suppose $[x] \in \mathbb{R}^+$ and $-[x] \in \mathbb{R}^+$. Then there exists $c, d > 0$ and $N \in \mathbb{N}$ such that $x_n > c$ and $-x_n > d$ for all $n \geq N$. This is a contradiction.

Now suppose neither is positive. Then for any $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that $x_n < \varepsilon$ and $-x_n < \varepsilon$ for all $n \geq N$. But this means that $x_n \to 0$ so $[x] = 0$, a contradiction.

Finally, let $[x], [y] \in \mathbb{R}^+$. Then for c, d bounding x and y, we have $x_n + y_n > c + d$ for sufficiently large n. Similarly, $x_n y_n > cd$ for the same n. □

Assignment 2

Quiz 19 Jan 2024

Problem 2.1. *Let F and G be ordered fields with the LUB property. In Lecture 04, we defined $h: F \to G$ as*

$$h(z) = \sup_G \{w \in \mathbb{Q} : w \leq z\}.$$

Show that h is a field isomorphism, i.e.,

(1) h is a bijection between F and G,

(2) $h(x + y) = h(x) + h(y)$ for all $x, y \in F$,

(3) $h(x \cdot y) = h(x) \cdot h(y)$ for all $x, y \in F$.

Proof. Theorem I.6.1. □

Problem 2.2. *In this problem, you may assume the well-definedness, commutativity and associativty of addition of Dedekind cuts (as defined in Lecture 04). Let $O = \{z \in \mathbb{Q} : z < 0\}$. Verify that O is a Dedekind cut, and $A + O = A$ for all Dedekind cuts A. Let A be a Dedekind cut. Define a Dedekind cut B such that $A + B = O$. You must justify your answer.*

Proof. Proposition I.6.5. □

Problem 2.3. *Let $a = (a_n)_{n\in\mathbb{N}}$ and $b = (b_n)_{n\in\mathbb{N}}$ be sequences of rational numbers such that $b_n \neq 0$ for all $n \in \mathbb{N}$. Suppose*

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1.$$

(i) Are a and b equivalent?

(ii) Are a and b equivalent if a is a $\mathbb{Q}$-bounded sequence?

Solution.

(i) No. $a_n = n + 1$ and $b_n = n$ gives a counterexample.

(ii) Yes.

Let a be bounded by M. Let n_0 be such that for all $n \geq n_0$, $\frac{1}{2} < \frac{a_n}{b_n}$. Then, for all $n \geq n_0$, $|b_n| < 2|a_n| \leq 2M$. Thus b is bounded.

Let $\varepsilon > 0$. Let N be such that for all $n \geq N$,

$$\left|\frac{a_n}{b_n} - 1\right| < \frac{\varepsilon}{2M}.$$

Then for all $n \geq N$,

$$\begin{aligned} |a_n - b_n| &= |b_n|\left|\frac{a_n}{b_n} - 1\right| \\ &< 2M\frac{\varepsilon}{2M} \\ &= \varepsilon. \end{aligned}$$

■

Problem 2.4. *You cannot use the existence (or the LUB property) of the ordered field of real numbers in this problem, so you must work "within" $\mathbb{Q}$.*

(i) Show that every monotone $\mathbb{Q}$-bounded sequence of rational numbers is $\mathbb{Q}$-Cauchy.

(ii) Consider the following sequence:

$$x_n = \begin{cases} 2, & \textit{if } n = 0, \\ x_{n-1} - \frac{x_{n-1}^2 - 2}{2x_{n-1}} & \textit{if } n \geq 1 \end{cases}$$

We define an isomorphism from $\mathbb{Q}$ into $\mathbb{R}$ as **Lecture 06.** Mon 15 Jan '24

$$r \in \mathbb{Q} \mapsto [(r, r, \dots)] \in \mathbb{R}.$$

The proof is direct.

Theorem I.6.15. $(\mathbb{R}, +, \cdot, \leq)$ *satisfies the Archimedean property.*

Proof. Let $[a], [b] > 0$ be in $\mathbb{R}$. Since $[b]$ is $\mathbb{Q}$-Cauchy, there exists a positive $M \in \mathbb{Q}$ such that $b_n < M$ for all $n \in \mathbb{N}$.

Since $[a] > 0$, let $c \in \mathbb{Q}^+$ and $N \in \mathbb{N}$ be such that $a_n > c$ for all $n \geq N$. By the Archimedean property of $\mathbb{Q}$, there exists an $m \in \mathbb{N}$ such that $mc > M$. Thus $b_n < M < mc < ma_n$ for all $n \geq N$. Thus $(m+1)a_n - b_n > ma_n - b_n + c > c$ for all $n \geq N$ and so $[m+1][a] > [b]$. $\square$

Theorem I.6.16. $(\mathbb{R}, +, \cdot, \leq)$ *satisfies the LUB property.*

Proof. Let $A \subseteq \mathbb{R}$ be a non-empty bounded above set.

For $n \in \mathbb{N}^*$, let

$$U_n = \left\{ m \in \mathbb{Z} : \frac{m}{n} \text{ is an upper bound of } A \right\}.$$

From the Archimedean property of $\mathbb{R}$, U_n is non-empty and bounded below. By well-ordering, U_n has a minimum $m(n)$. Let $a_n = \frac{m(n)}{n}$ for each $n \in \mathbb{N}^*$.

Claim: $(a_n)_{n \in \mathbb{N}^*}$ is $\mathbb{Q}$-Cauchy.

Let ε be a positive rational number. By Archimedean, $\frac{1}{n} < \varepsilon$ for all n above some N in $\mathbb{N}$. Note that for any $n \in \mathbb{N}^*$, a_n is an upper bound of A, and $a_n - \frac{1}{n}$ is not an upper bound of A.

Thus for any $n, n' \geq N^*$, we have

$$\begin{aligned} \frac{m(n)}{n} &> \frac{m(n')}{n'} - \frac{1}{n'} \\ a_n - a_{n'} &> -\frac{1}{n'} \end{aligned} \qquad\qquad \begin{aligned} \frac{m(n')}{n'} &> \frac{m(n)}{n} - \frac{1}{n} \\ a_n - a_{n'} &< \frac{1}{n} \end{aligned}$$

and so $|a_n - a_{n'}| < \max\left\{\frac{1}{n}, \frac{1}{n'}\right\} < \varepsilon$.

Claim: $[(a_n)]$ is an upper bound of A.

Suppose there exists some $[x] > [a]$. That is, there is some positive rational

c such that $c < x_n - a_n$ for all n larger than some $N_1 \in \mathbb{N}^*$. Since (x_n) is $\mathbb{Q}$-Cauchy, $-c/2 < x_n - x_m < c/2$ for all n, m larger than some $N_2 \in \mathbb{N}^*$. $\square$

I.7 The Complex Numbers

Lecture 07. Wed 17 Jan '24

Definition I.7.1. A *complex number* is an ordered pair of real numbers. We define operations on the set $\mathbb{C}$ of complex numbers as follows.

$$\begin{aligned} (a,b)+(c,d) &= (a+c, b+d) \\ (a,b)\cdot(c,d) &= (ac-bd, ad+bc) \\ |(a,b)| &= \sqrt{a^2+b^2} \end{aligned}$$

We further define i to be $(0,1)$.

Remark. These operations make $\mathbb{C}$ a *normed field*.

Theorem I.7.2. *The map $f\colon \mathbb{R} \to \mathbb{C}$ given by $f(x) = (x,0)$ is an isomorphism into $\mathbb{C}$.*

This allows us to identify $x \in \mathbb{R}$ with $(x,0) \in \mathbb{C}$.

Remark. $(a,b) = a + ib$ for any $a, b \in \mathbb{R}$. $i^2 = -1$.

0 is the additive identity and $(-a) + i(-b)$ is the additive inverse of $a + ib$.

1 is the multiplicative identity and for $a + ib \neq 0$, $\frac{a}{a^2+b^2} + i\frac{-b}{a^2+b^2}$ is the multiplicative inverse of (a,b).

Theorem I.7.3 (Cauchy-Schwarz inequality)**.** *Let $a_1, \dots, a_n$ and $b_1, \dots, b_n$ be complex numbers. Then*

$$\Big|\sum_{j=1}^n a_j\overline{b_j}\Big|^2 \leq \Big(\sum_{j=1}^n |a_j|^2\Big)\Big(\sum_{j=1}^n |b_j|^2\Big).$$

Proof. Let $\lambda = u + iv \in \mathbb{C}$.

$$\begin{aligned} 0 &\leq \sum_{j=1}^n (a_j + \lambda b_j)\overline{(a_j+\lambda b_j)} \\ &= \sum_{j=1}^n (a_j\overline{a_j} + \overline{\lambda} a_j \overline{b_j} + \lambda b_j \overline{a_j} + |\lambda|^2 b_j \overline{b_j}) \\ &= \sum_{j=1}^n |a_j|^2 + 2[u\Re(A) + v\Im(A)] + (u^2+v^2)B \end{aligned}$$

where $A = \sum_{j=1}^n a_j\overline{b_j}$ and $B = \sum_{j=1}^n |b_j|^2$.

Let the right hand expression be $F(u, v)$. Then $F_u(u, v) = 2\Re(A) + 2uB$ and $F_v(u, v) = 2\Im(A) + 2vB$. Setting both to be 0 gives $u = -\frac{\Re(A)}{B}$ and $v = -\frac{\Im(A)}{B}$. These values of u and v give $\lambda = -A/B$. Thus

$$F(u, v) = \sum_{j=1}^n |a_j|^2 - \frac{2|A|^2}{B} + \frac{|A|^2}{B}$$

and so

$$|A|^2 \leq \sum_{j=1}^n |a_j|^2 \sum_{j=1}^n |b_j|^2. \qquad \square$$

Chapter II

Metric Spaces

II.1 Definitions & examples

Definition II.1.1. A *metric space* is a pair (X, d) consisting of a set X and a "distance function" $d\colon X \times X \to [0, \infty)$ such that

(M1) $d(x, y) = 0$ iff $x = y$,

(M2) $d(x, y) = d(y, x)$,

(M3) $d(x, z) \le d(x, y) + d(y, z)$ (triangle inequality).

Examples.

- $X = \mathbb{R}$, $d(x, y) = |x - y|$.
- (Real Euclidean space) $X = \mathbb{R}^n$. The inner product $\langle x,\, y\rangle = \sum_{j=1}^n x_j y_j$ gives the *Euclidean* distance $d(x, y) = \sqrt{\langle x - y,\, x - y\rangle}$.
- (Discrete metric) Let X be any set. Then $[x \ne y]$ is a distance function on X.
- $X = \mathbb{R}^n$, $p \in [1, \infty]$. For $p \ne \infty$,

$$d_p(x, y) = \left(\sum_{j=1}^n |x_j - y_j|^p \right)^{1/p}$$

and

$$d_\infty(x, y) = \max_{1\le j\le n} |x_j - y_j|.$$

If $p \neq 2$, then d_p is not induced by an inner product.

- For any metric space (X, d) and a subset $Y \subseteq X$, the restriction of d to $Y \times Y$ is a distance on Y.

Proposition II.1.2. *Given $a, b \in \mathbb{R}^n$,*

$$\big|\|a\| - \|b\|\big| \le \|a + b\| \le \|a\| + \|b\|.$$

Proof. From Cauchy-Schwarz,

$$\begin{aligned} \|a+b\|^2 &= \langle a+b,\, a+b\rangle \\ &= \|a\|^2 + 2\langle a,\, b\rangle + \|b\|^2 \\ &\le \|a\|^2 + 2\|a\|\|b\| + \|b\|^2 \\ &= (\|a\| + \|b\|)^2. \end{aligned}$$

□

II.2 Metric Topology

Lecture 08.
Thu 18 Jan '24

Definition II.2.1. Let (X, d) be a metric space.

(i) The *open ball* centered at p or radius $\varepsilon > 0$ is the set

$$B_d(p; \varepsilon) \coloneqq \{x \in X : d(p, x) < \varepsilon\}$$

This set is also called the ε-neighborhood of p. Similarly, the *closed ball* centered at p or radius $\varepsilon > 0$ is the set

$$\{x \in X : d(p, x) \leq \varepsilon\}$$

(ii) Given a set $E \subseteq X$ and $p \in X$, p is an *interior point* of E if there exists some $\varepsilon > 0$ such that the ε-neighborhood $B(p; e)$ is contained in E. The collection of all interior points of E, denoted E°, is called the *interior* of E.

(iii) A set $E \subseteq X$ is said to be *open* if it is equal to its interior.

(iv) The collection of all open sets of (X, d) is called the d-topology on X.

Remark. The empty set is always open.

Examples.

- The open ball on $\mathbb{R}$ is an interval $(p - \varepsilon, p + \varepsilon)$.
-
- For the discrete metric,

$$B_d(p; \varepsilon) = \begin{cases} \{p\} & \varepsilon < 1 \\ X & \varepsilon \geq 1 \end{cases}$$

  Every set is open, by taking any $\varepsilon = 1$.

Proposition II.2.2. *Every open ball is an open set.*

Proof. Let (X, d) be the metric. Let $p \in X$, $\varepsilon > 0$, and $q \in B(p; \varepsilon)$. Choose $\delta = \varepsilon - d(p, q) > 0$ works. We show that $B(q; \delta) \subseteq B(p; \varepsilon)$. Let $r \in B(q; \delta)$. Then from the triangle inequality,

$$\begin{aligned} d(p, r) &\leq d(p, q) + d(q, r) \\ &< d(p, q) + \delta \\ &= \varepsilon \end{aligned}$$

□

Proposition II.2.3. *The union of any collection of open sets is open, and the intersection of any finite collection of open sets is open.*

Proof. Let $\mathscr{U}$ be a collection of open sets. Let $E = \bigcup_{U \in \mathscr{U}} U$. For any $p \in E$, p is contained in some $U \in \mathscr{U}$. Then there exists some $\varepsilon > 0$ such that $B(p;\varepsilon) \subseteq U \subseteq E$.

Let $U_1, \ldots, U_n$ be open sets and let $E = \bigcap_{i=1}^{n} U_i$. For any $p \in E$, $p \in U_i$ for all i. Then there exist $\varepsilon_1, \ldots, \varepsilon_n > 0$ such that $B(p;\varepsilon_i) \subseteq U_i$ for all i. Letting ε be the minimum of the ε_i's, we have $B(p;\varepsilon) \subseteq U_i$ for all i. So $B(p;\varepsilon) \subseteq E$. $\square$

Definition II.2.4. Let (X, d) be a metric space and $E \subseteq X$.

(i) Given $p \in X$, we say that p is an *accumulation point* of E if for every $\varepsilon > 0$, $B(p;\varepsilon)$ contains a point $q \in E$ such that $q \neq p$.

(ii) A point $p \in E$ is said to be *isolated* in E if it is not an accumulation point of E.

(iii) E is said to be *closed* if it contains all its accumulation points.

(iv) The *closure* of E, denoted $\overline{E}$, is the union of E with all its accumulation points.

(v) The boundary of E is the set $\partial E = \overline{E} \setminus E^{\circ}$.

Examples.

- In the discrete metric, every point is isolated in every subset.
- Finite subsets have no accumulation points.

Remarks.

- p need not lie in E to be an accumulation point.
- If p is an accumulation point of E, then every neighborhood of p contains infinitely many points of E.

Examples.

Lecture 09.
Mon 21 Jan '24

- Closed intervals are closed sets.
- $E = [0,1)$ is neither open nor closed. $\overline{E} = [0,1]$, $E^\circ = (0,1)$ and $\partial E = \{0,1\}$.
- Finite sets are always closed.
- In the discrete metric, every set is both open and closed.

Proposition II.2.5. *A set is closed iff its complement is open.*

Proof. Let $E \subseteq X$ be closed. Let $x \in E^c$. Since x is not a limit point of E, there exists an open ball around it that contains no points of E. Thus x is an interior point of E^c. This gives that E^c is open.

Now let $E \subseteq X$ be open. Let x be a limit point of E^c. Then there is no open ball around x that lies entirely in E. Thus x cannot lie in E. This gives that E^c is closed. □

Corollary II.2.6. *The intersection of any collection of closed sets is closed, and the union of any finite collection of closed sets is closed.*

Proof. Propositions II.2.3 and II.2.5 and De Morgan's laws. □

Theorem II.2.7.

(i) The closure of a set is closed.

(ii) If $A \subseteq B$, then $\overline{A} \subseteq \overline{B}$.

(iii) $\overline{E}$ is the smallest closed set containing E. That is,

$$\overline{E} = \bigcap_{\substack{F \supseteq E \\ F \text{ closed}}} F.$$

Proof.

(i) Let x be a limit point of $\overline{E}$. We will show that x is a limit point of E. Let $\varepsilon > 0$. If $B(x;\varepsilon)$ contains no point from E, it must contain a limit point of E. But then there are points of E arbitrarily close to it, within the ball. Thus $B(x;\varepsilon)$ contains points of E for any $\varepsilon > 0$ and so x is a limit point of E.

(ii) Every point in A is a point in B. Every limit point of A is a limit point of B.

(iii) Every closed set F containing E must contain

- every point of E, and
- every limit point of E.

Thus $\overline{E} \subseteq F$ for every such F. so $\overline{E}$ is the smallest closed set containing E. $\square$

Definition II.2.8 (Relative topology)**.** Given a metric space (X, d) and subsets $E \subseteq Y$, we say that E is open (resp. closed) *relative to* Y if E is an open (resp. closed) set in the metric space $(Y, d|_Y)$.

Proposition II.2.9. *Let (X, d) be metric space and $E \subseteq Y \subseteq X$. Then E is open relative to Y iff there exists an open set $F \subseteq X$ such that $E = F \cap Y$.*

Proof. Let E be open relative to Y. For each $x \in E$, there exists an $\varepsilon_x > 0$ such that $B_Y(x, \varepsilon_x) \subseteq E$. Let $F = \bigcup_{x \in E} B_X(x, \varepsilon_x)$. Then F is open relative to X and $F \cap Y = E$.

Let $E = F \cap Y$ where F is open in X. For every $x \in F$ there exists an $\varepsilon > 0$ such that $B_X(x, \varepsilon) \subseteq F$. Then for any $y \in E$, $B_Y(y, \varepsilon) = B_X(y, \varepsilon) \cap Y$ is contained in $F \cap Y = E$. So E is open in Y. $\square$

Corollary II.2.10. *Let (X, d) be metric space and $E \subseteq Y \subseteq X$. Then E is closed relative to Y iff there exists a closed set $F \subseteq X$ such that $E = F \cap Y$.*

Proof. Propositions II.2.5 and II.2.9. $\square$

II.3 Compactness

Lecture 10.
Wed 24 Jan '24

Definition II.3.1. A subset $E \subseteq (X, d)$ is said to be *bounded* if there exists a $p \in X$ and $M > 0$ such that $E \subseteq B(p; M)$.

Equivalently, E is bounded if for all $x \in X$, there exists an $M > 0$ such that $E \subseteq B(x; M)$. Or, E is bounded if there exists an $M > 0$ such that $d(x, y) < M$ for all $x, y \in E$.

Consider $E = \{p \in \mathbb{Q} : 2 < p^2 < 3\}$. Then E is both closed and bounded in $(\mathbb{Q}, |\cdot|)$. However, continuous functions on E are neither uniformly continuous nor bounded.

Definition II.3.2 (Open cover)**.** Let $E \subseteq (X, d)$. An open cover $\{\mathcal{U}_\alpha\}_{\alpha\in\Lambda}$ of E in X is a collection of open sets $\mathcal{U}_\alpha$ such that $E \subseteq \bigcup_{\alpha\in\Lambda} \mathcal{U}_\alpha$.

Definition II.3.3 (Compact set)**.** A subset $E \subseteq (X, d)$ is said to be compact if any open cover $\mathcal{U} = \{\mathcal{U}_\alpha\}_{\alpha\in\Lambda}$ of E in X admits a finite subcover of E, *i.e.*, there exist $\alpha_1, \ldots, \alpha_k \in \Lambda$ such that $E \subseteq \bigcup_{i=1}^{k} \mathcal{U}_{\alpha_i}$.

Examples.

- $E \subseteq (X, d)$ is finite. Let $\mathcal{U}$ be an open cover of $E = \{p_1, \ldots, p_n\}$. Then for each $p_j \in E$, there exists $\alpha_j \in \Lambda$ such that $p_j \in \mathcal{U}_{\alpha_j}$. Then $E \subseteq \bigcup_{j=1}^{n} \mathcal{U}_{\alpha_j}$.

- $E = (0, 1)$ is not compact in $(\mathbb{R}, |\cdot|)$.

  Proof. Let $\mathcal{U}_n = (\frac{1}{n+2}, \frac{1}{n})$ for $n \in \mathbb{N}^*$. Then $\mathcal{U} = \{\mathcal{U}_n\}_{n\in\mathbb{N}^*}$ is an open cover of E. However, $\mathcal{U}$ does not admit a finite subcover of E. For any finite $\{\mathcal{U}_{n_1}, \ldots, \mathcal{U}_{n_k}\}$, let $n_0 = \max\{n_j : 1 \le j \le k\}$. Then $\bigcup \mathcal{U}_{n_j} \subseteq (\frac{1}{n_0+2}, 1)$ and thus is not a cover of E. $\square$

- $E = [0, 1]$ is compact in $(\mathbb{R}, |\cdot|)$. In fact, all rectangles (sets of the form $[a_1, b_1] \times \cdots \times [a_n, b_n]$) are compact in $(\mathbb{R}^n, \|\cdot\|)$.

Theorem II.3.4. *Let $E \subseteq (\mathbb{R}^n, \|\cdot\|)$. Then the following are equivalent:*

(1) E is compact.

(2) E is closed and bounded.

(3) Every infinite subset of E admits a limit point in E.

Remark. The equivalence between (1) and (2) is known as the Heine-Borel theorem.

We first show that (1) $\implies$ (2) in a general metric space. This is not necessary for the theorem, since we will later show (1) $\implies$ (3) $\implies$ (2) $\implies$ (1) anyway. But (2) $\implies$ (1) is only valid in $\mathbb{R}^n$, so that won't yield (1) $\implies$ (2) in a general space.

Proof that (1) $\implies$ (2) in general. Let $E \subseteq (X, d)$ be compact. Let $z \in E^c$. For any $y \in E$, let $\delta_y = d(y, z)/2$. Note that $B(z, \delta_y) \cap B(y, \delta_y) = \varnothing$.

Then $\mathcal{U} = \{B(y; \delta_y) : y \in E\}$ is an open cover of E. Since E is compact, $\mathcal{U}$ admits a finite subcover of E. That is, there exist $y_1, \ldots, y_k \in E$ such that $E \subseteq \bigcup_{i=1}^k B(y_i; \delta_{y_i})$. Let $\delta = \min\{\delta_{y_i}\}$. Then $B(z; \delta) \cap E = \varnothing$, so $B(z; \delta) \subseteq E^c$. This shows that E^c is open, so E is closed.

For boundedness, take the largest ball from finite subcover of E in $\bigcup_{R>0} B(p; R)$ for some $p \in E$. $\square$

To show that (2) $\implies$ (1) in $(\mathbb{R}^n, \|\cdot\|)$, we first show that for any $R \in \mathbb{R}$, the set $[-R, R]^n$ is compact.

Theorem II.3.5. *All rectangles of the form $[-R, R]^n$ are compact in $(\mathbb{R}^n, \|\cdot\|)$.*

Proof. Fix an $R > 0$. Let $I_0 = [-R, R]^n$. Note that

$$\operatorname{diam}(I_0) = \max\{\|x - y\| \mid x, y \in I_0\} = 2R\sqrt{n}.$$

Let $\mathcal{U}$ be an open cover of I_0. Suppose this has no finite subcover of I_0. We partition I_0 into 2^n equal rectangles of the form $J_1 \times J_2 \times \ldots J_n$, where each J_i is either $[-R, 0]$ or $[0, R]$. The diameter of each of these rectangles is $R\sqrt{n}$. By the pigeonhole principle, there exists I_1 among these rectangles that does not admit a finite subcover in $\mathcal{U}$.

We continue this process to obtain a sequence of rectangles $I_0 \supseteq I_1 \supseteq I_2 \supseteq \ldots$, none of which admit a finite subcover in $\mathcal{U}$, with $\operatorname{diam}(I_k) = \frac{2R}{2^k}\sqrt{n}$.

We use the nested interval property (lemma II.3.6) to obtain a point $x \in \bigcap_{k=1}^\infty I_k$. Since $\mathcal{U}$ covers I_0, there exists $U \in \mathcal{U}$ that contains x. Since U is open, it contains some ε-ball around x. But the diameters of the I_ks decrease to 0, so $I_k \subseteq U$ for some k. This contradicts that I_k does not admit a finite subcover in $\mathcal{U}$.

Thus the original rectangle I_0 admits a finite subcover in $\mathcal{U}$. $\square$

Lemma II.3.6 (Nested interval theorem)**.** *Let $I_0 \supseteq I_1 \supseteq \ldots$ be a sequence of closed rectangles in $\mathbb{R}^n$. Then $\bigcap_{k=1}^\infty I_k \neq \varnothing$.*

Proof. Let each $I_k = [a_{1k}, b_{1k}] \times \cdots \times [a_{nk}, b_{nk}]$. Then $(a_{ik})_{k\in\mathbb{N}}$ and $(b_{ik})_{kin\mathbb{N}}$ are bounded sequences in $\mathbb{R}$. Let $A_i = \sup\{a_{ik} \mid k \in \mathbb{N}\}$. Note that each b_{ik} is an upper bound for the set $\{a_{ik} \mid k \in \mathbb{N}\}$, so $A_i \le b_{ik}$ for all k.

Let $A = (A_1, \ldots, A_n)$. For each k, $A \in I_k$ since $A_i \in [a_{ik}, b_{ik}]$ for all i. Thus $\bigcap_{k=1}^{\infty} I_k \ni A$. $\square$

Lecture 11. Thu 25 Jan '24

How do we go from the compactness of rectangles to the compactness of arbitrary closed and bounded sets in $\mathbb{R}^n$? We need the following theorem.

Theorem II.3.7. *A closed subset of a compact set is compact.*

Proof. Let $E \subseteq X$ be compact and $F \subseteq E$ be closed. Let $\mathcal{U}$ be an open cover of F. Then $\mathcal{U} \cup \{F^c\}$ is an open cover of E. This contains a finite subcover $\mathcal{V}$. Then $\mathcal{V} \setminus \{F^c\} \subseteq \mathcal{U}$ is a finite subcover of F. $\square$

We are now ready to show that closed and bounded sets in $\mathbb{R}^n$ are compact.

Proof that (2) $\implies$ (1) in $\mathbb{R}^n$. Let $E \subseteq \mathbb{R}^n$ be closed and bounded. There is an $R > 0$ such that $E \subseteq B(0; R)$, but $B(0; R) \subseteq [-R, R]^n$. So E is a closed subset of the compact set $[-R, R]^n$. By theorem II.3.7, E is compact. $\square$

Theorem II.3.8. *Let $\{K_\alpha\}_{\alpha\in\Lambda}$ be a collection of compact sets in (X, d) such that any non-empty finite subcollection has non-empty intersection. Then $\bigcap_{\alpha\in\Lambda} K_\alpha \neq \varnothing$.*

Proof. Suppose $\bigcap_{\alpha\in\Lambda} K_\alpha = \varnothing$. Choose a $K \in \{K_\alpha\}_{\alpha\in\Lambda}$. No element in K is in every other K_α. Let $U_\alpha = K_\alpha^c$ for each α. Any point in K is in at least one U_α. Then $\{U_\alpha\}_{\alpha\in\Lambda}$ is an open cover of K. Since K is compact, there is a finite subcover $U_{\alpha_1}, \ldots, U_{\alpha_n}$. But then

$$\begin{aligned} K &\subseteq U_{\alpha_1} \cup \cdots \cup U_{\alpha_n} \\ &= K_{\alpha_1}^c \cup \cdots \cup K_{\alpha_n}^c \\ &= (K_{\alpha_1} \cap \cdots \cap K_{\alpha_n})^c \end{aligned}$$

so $K \cap K_{\alpha_1} \cap \cdots \cap K_{\alpha_n} = \varnothing$, a contradiction. $\square$

Corollary II.3.9. *Let $K_0 \supseteq K_1 \supseteq K_2 \supseteq \ldots$ be non-empty compact sets in (X, d). Then $\bigcap_{n\in\mathbb{N}} K_n \neq \varnothing$.*

Remark. This cannot be used to prove the nested interval property, since the proof of compactness of rectangles in $\mathbb{R}^n$ itself relies on the nested interval property.

Theorem II.3.10. *Every infinite subset of a compact set has a limit point in the set.*

Proof. Let $E \subseteq (X, d)$ be compact and $F \subseteq E$ be infinite. Suppose F has no limit point in E. Then for every $z \in E$, there exists an $\varepsilon_z > 0$ such that $B(z; \varepsilon_z)$ contains no point of F, except possibly z. Then $\{B(z; \varepsilon_z)\}_{z \in E}$ is an open cover of E.

Since E is compact, this contains a finite subcover. But each $B(z; \varepsilon_z)$ contains at most one point of F, so only finitely many points of F are covered. Contradiction. $\square$

Proof that (3) $\implies$ (2). Suppose (3) holds on some $E \subseteq (\mathbb{R}^n, \|\cdot\|)$ but E is not bounded. Let $x_0 \in E$. We can produce a sequence $(x_n)_{n \in \mathbb{N}} \subseteq E$ such that

$$\|x_{n+1}\| > \|x_n\| + 1 \text{ for all } n \in \mathbb{N}.$$

This cannot have a limit point in E (or anywhere) since for any $x \in E$, $B(x, 1)$ contains at most one point of the sequence.

Now suppose (3) holds on E but E is not closed. Then there exists a $z \in E^c$ such that z is a limit point of E. Then there exists a sequence $(x_n)_{n \in \mathbb{N}} \subseteq E$ such that $\|x_j - z\| < \frac{1}{j}$ for all $j \in \mathbb{N}$. The set $F = \{x_n\}_{n \in \mathbb{N}}$ is infinite (otherwise, the minimum distance is the infimum, which is zero, but $z \notin E$). Then F must have a limit point in E.

For any $y \in \mathbb{R}^n$,

$$\begin{aligned} \|x_j - y\| &\geq \|z - y\| - \|x_j - z\| \\ &\geq \|z - y\| - \frac{1}{j}. \end{aligned}$$

If $\|z - y\|$ is positive, then there are only finitely many x_j within a distance $\|z - y\|$ of y. Hence y can be a limit point of F only if $y = z$. $\square$

Theorem II.3.11. *Let $E \subseteq Y \subseteq (X, d)$ where Y is compact in X. Then E is compact relative to Y if and only if it is compact in X.*

Proof. We use proposition II.2.9.

Suppose E is compact in Y. Let $\mathcal{U}$ be an open cover of E in X. Then $\mathcal{V} = \{U \cap Y\}_{U \in \mathcal{U}}$ is an open cover of E in Y. This has a finite subcover $\{U_1 \cap Y, \dots, U_n \cap Y\}$. Then $\{U_1, \dots, U_n\}$ is a finite subcover of E in $\mathcal{U}$.

Suppose E is compact in X. Let $\mathcal{V}$ be an open cover of E in Y. Then for each $V \in \mathcal{V}$, there exists a set U_V open in X such that $V = U_V \cap Y$, so that $\mathcal{U} = \{U_V\}_{V \in \mathcal{V}}$ is an open cover of E in X. This has a finite subcover $\{U_{V_1}, \ldots, U_{V_n}\}$. Then $\{V_1, \ldots, V_n\}$ is a finite subcover of E in $\mathcal{V}$. $\square$

II.4 Connected Sets

Lecture 12.
Mon 29 Jan '24

Definition II.4.1.

(a) Let (X, d) be a metric space. A pair of sets $A, B \subseteq X$ are said to be *separated* in X if $\overline{A} \cap B = A \cap \overline{B} = \varnothing$.

(b) A set $E \subseteq X$ is said to be *disconnected* if it is the union of two separated sets in X.

(c) E is *connected* if it is not disconnected.

Examples.

- Sets $A = (-1, 0)$ and $B = (0, 1)$ are separated in $\mathbb{R}$. Note that sgn is continuous on $A \cup B$ but does not satisfy the intermediate value property.

  However, if $A = (-1, 0]$ instead, then all continuous functions on $A \cup B$ satisfy the intermediate value property.

- The empty set is connected.

- $\mathbb{Q}$ is disconnected in $\mathbb{R}$. The partition $\{\mathbb{Q} \cap (-\infty, \sqrt{2}), \mathbb{Q} \cap (\sqrt{2}, \infty)\}$ separates $\mathbb{Q}$.

- $\mathbb{Q}$ is disconnected even in $\mathbb{Q}$.

Exercise II.4.2. *Let $E \subseteq Y \subseteq (X, d)$. Then E is connected relative to Y iff E is connected in X.*

Theorem II.4.3. *Let $E \subseteq \mathbb{R}$. Then E is connected iff E is convex, i.e., for all $x < y \in E$, $[x, y] \subseteq E$.*

Proof. Suppose E is connected, but not convex, *i.e.*, there exist $x < y \in E$ and some $r \in (x, y)$ that is not in E. Then $A = (-\infty, r] \cap E$ and $B = [r, \infty) \cap E$ separate E.

Conversely, suppose E is convex but not connected. Then there exist $A, B \subseteq E$ that separate E. Let $x \in A$ and $y \in B$ and suppose WLOG that $x < y$. Note that $A \cap [x, y]$ is non-empty and bounded. Let $r = \sup(A \cap [x, y])$.

By the lemma below, $r \in \overline{A \cap [x, y]} \subseteq \overline{A} \cap [x, y]$ so $r \in \overline{A}$. Disconnectedness forces that $r \notin B \iff r \in A$ so $x \leq r < y$.

But since r is the supremum of $A \cap [x, y]$, $(r, y) \subseteq B$. This gives $r \in \overline{B}$, violating the separation of A and B. $\square$

II.5 The Cantor Set

Definition II.5.1 (Perfect set)**.** A set $E \subseteq (X, d)$ is said to be *perfect* if every point of E is a limit point of E.

Note that $E = [0, 1]$ is perfect in $\mathbb{R}$. Can we produce a "sparse" perfect set? Throwing away isolated points makes the set open. Throwing away a finite number of open sets preserves perfectness, but there are still *intervals of positive length*.

Can we produce a perfect set such that

(i) it contains no intervals of positive length?

(ii) E is *nowhere dense*, *i.e.*, $(\overline{E})^\circ = \varnothing$?

Note that the second condition implies the first.

Lecture 13.
Wed 31 Jan '24

Definition II.5.2 (Ternary expansion)**.** Let $x \in [0, 1]$. A *ternary expansion* of x is a sequence $(d_1, d_2, \dots) \subseteq \{0, 1, 2\}$ such that

$$x = \sup\left\{ D_k = \sum_{j=1}^{k-1} \frac{d_j}{3^j} : k \geq 1 \right\}$$

which is equivalent to

$$\sum_{j=1}^{\infty} \frac{d_j}{3^j} = x$$

We write $x = 0.d_1 d_2 d_3 \dots$ to denote this.

Example. For $x = \frac{1}{3}$, we have both $x = 0.1000\dots$ and $x = 0.0222\dots$, so ternary expansions are not unique.

	0	1	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{4}{9}$	$\frac{1}{2}$
A	$0.000\ldots$		$0.100\ldots$	$0.200\ldots$	$0.1100\ldots$	$0.111\ldots$
B		$0.222\ldots$	$0.022\ldots$	$0.122\ldots$	$0.1022\ldots$	$0.111\ldots$

Table II.1: Scheme A vs Scheme B

Let $I_0 = [0, \frac{1}{3}]$, $I_1 = [\frac{1}{3}, \frac{2}{3}]$ and $I_2 = [\frac{2}{3}, 1]$. Let $x \in [0,1]$. Choose $d_1 = j$ such that $x \in I_j$ (in ambiguous cases, pick any one). Then

$$\begin{aligned} x &\in \left[\frac{d_1}{3}, \frac{d_1+1}{3}\right] \\ \implies 0 &\leq x - \frac{d_1}{3} \leq \frac{1}{3} \\ \implies D_1 &\leq x \leq D_1 + \frac{1}{3} \end{aligned}$$

Let I_{j0}, I_{j1}, I_{j2} be the subdivisions of I_j. Choose $d_2 = l$, where $x \in I_{jl}$ iff

$$\begin{aligned} x &\in \left[\frac{d_1}{3} + \frac{d_2}{9}, \frac{d_1}{3} + \frac{d_2+1}{9}\right] \\ \implies D_2 &\leq x \leq D_2 + \frac{1}{9} \end{aligned}$$

How do we break ties?

Scheme A If at the k^{th} state, $x \in [0,1)$ is an endpoint of 2 intervals, pick the right interval. This gives a unique expansion. That is, pick d_k such that $D_k \leq x < D_k + \frac{1}{3}$.

Scheme B For $x \in (0,1]$, always pick the left interval. That is, pick d_k such that $D_k < x \leq D_k + \frac{1}{3}$.

We make the following observations:

- Ambiguity only occurs at endpoints of "middle thirds".
- Say x is an endpoint of a middle third. Let k be the first stage where ambiguity occurs. Then if x is the left endpoint, scheme A gives $x = 0.d_1 d_2 \ldots d_{k-1}1000\ldots$ and scheme B gives $x = 0.d_1 d_2 \ldots d_{k-1}0222\ldots$. If x is the right endpoint, scheme A gives $x = 0.d_1 d_2 \ldots d_{k-1}2000\ldots$ and scheme B gives $x = 0.d_1 d_2 \ldots d_{k-1}1222\ldots$.

Note that this ambiguity can be resolved by a scheme C, which picks the expansion which has no 1 starting from the point of ambiguity.

Theorem II.5.3. *There exists a non-empty $E \subseteq [0,1]$ such that*

(i) E *is compact.*

(ii) $E = \{$*limit points of* $E\}$.

(iii) $E^\circ = \overline{E}^\circ = \varnothing$.

(iv) E *is uncountable.*

Proof.

$$E = \{x \in [0,1] : x \text{ admits at least one ternary expansion with only 0's and 2's}\}$$

We can construct this set by removing the middle thirds.

$$\begin{aligned}
E_0 &= [0,1] \\
E_1 &= E_0 \setminus \left(\frac{1}{3}, \frac{2}{3}\right) \\
&= \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right] \\
E_2 &= E_1 \setminus \left[\left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{4}{9}, \frac{5}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right)\right] \\
E_m &= E_{m-1} \setminus \bigcup_{k=0}^{3^{m-1}-1} \left(\frac{3k+1}{3^m}, \frac{3k+2}{3^m}\right)
\end{aligned}$$

We claim that $E = \bigcap_{m=1}^{\infty} E_m$ satisfies the conditions of the theorem. First note that $E_1 \subseteq E_2 \subseteq \dots$, so for any $m_1 < m_2 < \cdots < m_k$, $\bigcap_{i=1}^{k} E_{m_i} = E_{m_k}$ is non-empty. By theorem II.3.8, E is non-empty

Since E is the intersection of closed sets, E is closed. Since E is bounded, E is compact.

We have that $E^\circ = \varnothing$ since E does not contain any open intervals. *Formally*, we will show that for any interval (a,b), there exist k and m such that $\left(\frac{3k+1}{3^m}, \frac{3k+2}{3^m}\right)$ is contained in (a,b).

Heuristically, we see that the length of the removed intervals is $\frac{1}{3}+\frac{1}{9}+\cdots = 1$, so that the remaining set cannot contain any interval of positive length.

Uncountability is by a diagonal argument. $\square$

Chapter III

Sequences & Series

Lecture 14.
Thu 01 Feb '24

III.1 Sequences & Subsequences

Definition III.1.1. Let (X, d) be a metric space. A squence in X is a function $f\colon \mathbb{N} \to X$, more commonly written as $(f(k))_{k\in\mathbb{N}} \subseteq X$.

We say that a sequence $(x_n)_{n\in\mathbb{N}}$ *converges* in X if there exists an $x \in X$ such that for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that for all $n \geq N$, $d(x_n, x) < \varepsilon$. In this case, we call x a limit of $(x_n)_{n\in\mathbb{N}}$ and write

$$\lim_{k\to\infty} x_k = x \quad \text{or} \quad x_k \to x \text{ as } k \to \infty.$$

If $(x_n)_{n\in\mathbb{N}}$ does not converge, we say that it *diverges*.

Examples.

- When $(X, d) = (\mathbb{R}, |\cdot|)$, this definition reduces to the definition in UMA101.
- Let $x_n = (\frac{1}{n}, \frac{2}{n^2}) \in (\mathbb{R}^2, \|\cdot\|)$ for each $n \geq 1$.
  We claim that $\lim_{n\to\infty} x_n = (0, 0)$.

  Proof. Let $\varepsilon > 0$. Choose an $N > \frac{\sqrt{5}}{\varepsilon}$. Then for all $n \geq N$,

  $$\begin{aligned}\left\|\left(\frac{1}{n}, \frac{2}{n^2}\right)\right\|^2 &= \frac{1}{n^2} + \frac{4}{n^4} \\ &\leq \frac{5}{n^2} \\ &< \varepsilon.\end{aligned}$$ ∎

- Let $x = \left(\frac{1}{n}, (-1)^n\right)_{n\in\mathbb{N}^*}$ with standard norm. Then $(x_n)_{n\in\mathbb{N}^*}$ diverges.

Theorem III.1.2. *Let (X, d) be a metric space.*

(i) Let $(x_n)_{n\in\mathbb{N}} \subseteq X$. Then, $\lim_{n\to\infty} x_n = x$ iff every ε-ball centred at x contains all but finitely many terms of (x_n).

(ii) Suppose $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} x_n = y$. Then $x = y$.

(iii) If $(x_n)_{n\in\mathbb{N}} \subseteq X$ converges, then $\{x_n : n \in \mathbb{N}\}$ is a bounded set in (X, d).

(iv) Let $E \subseteq X$. Then $x \in \overline{E}$ iff there exists a sequence $(x_n) \subseteq E$ such that $\lim_{n\to\infty} x_n = x$.

Proof.

(i) Let (x_n) be convergent to x. Then all terms except the first N lie inside the ε- neighborhood of x. The converse is similarly true.

(ii) Let x and y be distinct limits of (x_n). Choose $\varepsilon = \frac{d(x,y)}{2} > 0$. Then for large enough n,

$$\begin{aligned} d(x, y) &\leq d(x, x_n) + d(x_n, y) \\ &< \varepsilon + \varepsilon \\ &= d(x, y). \end{aligned}$$

□

(iii) Let (x_n) be convergent to x. Let N be such that for all $n \geq N$, $d(x_n, x) < 1$. Then $\rho = \sum_{k=0}^{N} d(x_k, x) + 1$ works as a radius for $B(x, \rho) \supseteq \{x_n : n \in \mathbb{N}\}$.

(iv) Let $x \in \overline{E}$. Then every ε-neighborhood of x intersects E. By the axiom of choice, we can choose a sequence $(x_n) \subseteq E$ such that $d(x_n, x) < \frac{1}{n}$. This converges to x.

Conversely if there exists a sequence $(x_n) \to x$ within E, then every ε-neighborhood of x intersects E.

Definition III.1.3. Let $(x_n)_{n\in\mathbb{N}} \subseteq X$. Let $(n_k)_{k\in\mathbb{N}}$ be a strictly increasing sequence in $\mathbb{N}$. Then $(x_{n_k})_{k\in\mathbb{N}}$ is called a *subsequence* of (x_n).

Any limit of a subsequence of (x_n) is called a *subsequential limit* of (x_n).

Lecture 15.
Mon 05 Feb '24

Example. Let $x_n = (\frac{1}{n}, (-1)^n) \subseteq \mathbb{R}^2$ for $n \geq 1$. Then (x_n) is not convergent, but has subsuential limits $(0, 1)$ and $(0, -1)$ corresponsing to the subsequences (x_{2n}) and (x_{2n-1}) respectively.

Theorem III.1.4. *Let $(x_n)_{n\in\mathbb{N}} \subseteq (X,d)$. Then $\lim_{n\to\infty} x_n = x$ iff every subsequence converges to x.*

Proof. Suppose (x_n) is convergent. Let $(y_k) = (x_{n_k})$ be a subsequence. Then for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \varepsilon$ for all $n \geq N$. But this same N works for the subsequence, since $n_k \geq k$. Thus each subsequence converges to x.

Now suppose every subsequence converges to x. Since the sequence itself is a subsequence, it converges to x. $\square$

Theorem III.1.5. *Let $E \subseteq (X,d)$. Then the following are equivalent.*

(1) E is compact.

(2) Every infinite subset of E has a limit point in E.

(3) Every sequence in E has a subsequential limit in E.

(1) $\iff$ (2) is by theorem II.3.4. We prove (2) $\iff$ (3).

Proof of (2) $\Rightarrow$ (3). Suppose every infinite subset of E has a limit point in E. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in E, and let S be the set of all x_n.

If S is finite, then by the Pigeonhole Principle, there exists some $x \in S$ such that $x_n = x$ for infinitely many n. Then the constant sequence (x) is a subsequence of (x_n), whose limit x is in E.

If not, then S is infinite, so it has a limit point $p \in E$. Thus for every $k \in \mathbb{N}$, there exists an $N_k \in \mathbb{N}$ such that $x_{N_k} \neq p \in B(p; \frac{1}{k})$.

Let n_1 be such that $d(x_{n_1}, p) < 1$. For n_{k+1}, consider $S \setminus \{x_0, \ldots, x_{n_k}\}$. p is also a limit point of this set (why?), so there exists an $n_{k+1} > n_k$ such that $d(x_{n_{k+1}}, p) < \frac{1}{k+1}$. Then $(x_{n_k})_k$ is a subsequence of $(x_n)_n$, and $\lim_{k\to\infty} x_{n_k} = p \in E$. $\square$

Corollary III.1.6. *Let $(x_n)_{n\in\mathbb{N}} \subseteq (\mathbb{R}^k, \|\cdot\|)$ be a bounded sequence. Then (x_n) has a convergent subsequence.*

Proof. Let $p \in \mathbb{R}^k$ and $R > 0$ be such that $(x_n) \subseteq B(p; R) \subseteq \overline{B(p; R)}$ which is compact (why?). Then by the previous theorem, (x_n) has a convergent subsequence. $\square$

Lecture 16. Wed 07 Feb '24

Proof of (3) $\Rightarrow$ (2). Let $S \subseteq E$ be an infinite set. Thus there exists a sequence $(x_n)_n \subseteq S$ of distinct elements.

By (3), there exists a subsequence $(x_{n_k})_k$ of $(x_n)_n$ such that $(x_{n_k})_k$ is convergent to some $x \in E$. By the sequential characterization of closures, $x \in \overline{S}$.

Thus for all $\varepsilon > 0$, there exists a $k_\varepsilon \in \mathbb{N}$ such that for all $k \geq k_\varepsilon$, we have that $d(x_{n_k}, x) < \varepsilon$. Thus x is a limit point of S in E. $\square$

III.2 Cauchy Sequences & Completeness

Recall the HW2 problem to show that the sequence $(x_n)_n$ given by

$$x_n = \begin{cases} 2 & \text{if } n = 0 \\ x_{n-1} - \frac{x_{n-1}^2 - 2}{2x_{n-1}} & \text{if } n \geq 1 \end{cases}$$

is $\mathbb{Q}$-Cauchy but not convergent in $\mathbb{Q}$. This is an application of the Newton-Raphson method.

III.2.1 Newton-Raphson Method (Informal)

Given a function $f \colon \mathbb{R} \to \mathbb{R}$, we want to find a root of f. We pick some initial guess $x_0 \in \mathbb{R}$, and iterate via

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}.$$

Under *some* assumptions on f and x_0, $(x_n)_n$ is Cauchy. Then

$$f(x_{n-1}) = f'(x_{n-1})(x_{n-1} - x_n) \to 0$$

If $\lim_{n\to\infty} x_n = l$, and f is continuous, then

$$f(l) = \lim_{n\to\infty} f(x_n) = 0.$$

Definition III.2.1 (Cauchy sequence)**.** Let $(x_n)_{n\in\mathbb{N}} \subseteq (X, d)$. We say that (x_n) is a *Cauchy sequence* if for every $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that whenever $n, m \geq N$, $d(x_n, x_m) < \varepsilon$.

Definition III.2.2 (Completeness)**.** (X, d) is said to be a *complete* metric space if every Cauchy sequence in (X, d) is convergent.

Theorem III.2.3.

(a) Every convergent sequence is Cauchy.

(b) Every Cauchy sequence is bounded.

Proof. Trivial. □

Theorem III.2.4. *Every compact metric space is complete.*

Proof. Let (X, d) be compact and let $(x_n)_n$ be a Cauchy sequence in X. Since X is compact, $(x_n)_n$ has a convergent subsequence $(x_{n_k})_k$ converging to some $x \in X$ (by theorem III.1.5).

Then $(x_n)_n$ also converges to x by the triangle inequality. For large enough n, $d(x_n, x) \leq d(x_n, x_{n_n}) + d(x_{n_n}, x) < 2\varepsilon$. □

Theorem III.2.5. $(\mathbb{R}^d, \|\cdot\|)$ *is complete.*

Proof. Let $(x_n)_n$ be a Cauchy sequence in $\mathbb{R}^d$. Then it must be bounded. Take a closed ball B centered at x_0 containing all elements of $(x_n)_n$. This is compact, and so the above theorem applies to give that $(x_n)_n$ has a limit in $B \subseteq \mathbb{R}^d$. □

Exercise III.2.6. *Every increasing and bounded above sequence in* $\mathbb{Q}$ *or* $\mathbb{R}$ *is Cauchy.*

Proof. Suppose not. Then there exists an $\varepsilon > 0$ such that for all $N \in \mathbb{N}$, there exist $n(N) > m(N) \geq N$ such that $\left|x_{n(N)} - x_{m(N)}\right| \geq \varepsilon$.

Let $m_0 = m(0)$ and $n_0 = n(0)$. For $k \geq 1$, let $m_k = m(n_{k-1})$ and $n_k = n(n_{k-1})$. Then

$$\begin{aligned} x_{n_k} &\geq x_{m_k} + \varepsilon \\ &\geq x_{n_{k-1}} + \varepsilon \end{aligned}$$

and so $(x_{n_k})_k$ is a subsequence with each term at least ε more than the last. Thus $x_{n_k} \geq x_0 + k\varepsilon$ for all $k \in \mathbb{N}$, which contradicts boundedness. □

Alternatively, we could see $\mathbb{Q}$ as a subset of $\mathbb{R}$, and use the fact that a bounded monotone sequence in $\mathbb{R}$ is convergent.

III.3 Sequences in $\mathbb{R}$

Lecture 17.
Thu 08 Feb '24

Definition III.3.1 (The Extended Reals)**.** The *extended real line* is the set of real numbers along with 2 formal sumbols $+\infty$ and $-\infty$, denoted by

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}.$$

$\overline{\mathbb{R}}$ will be endowed with the order

$$-\infty < x < \infty \text{ for all } x \in \mathbb{R},$$

along with the usual order on $\mathbb{R}$. We extend the algebraic operations on $\mathbb{R}$ to $\overline{\mathbb{R}}$.

- $x + \infty = +\infty$, $x - \infty = -\infty$ for all $x \in \mathbb{R}$.
- $x \cdot (+\infty) = +\infty$, $x \cdot (-\infty) = -\infty$ for all $x \in \mathbb{R}$, $x > 0$.
- $x \cdot (+\infty) = -\infty$, $x \cdot (-\infty) = +\infty$ for all $x \in \mathbb{R}$, $x < 0$.
- $\frac{x}{+\infty} = \frac{x}{-\infty} = 0$, for all $x \in \mathbb{R}$.

If $E \subseteq \mathbb{R}$ is not bounded above in $\mathbb{R}$, we say $\sup E = +\infty$.

When constructing $\mathbb{R}$ through Dedekind cuts, $\overline{\mathbb{R}}$ can be constructed by relaxing the condition that a cut must be neither empty nor the whole of $\mathbb{Q}$. Then $\varnothing$ is a Dedekind cut represented as $-\infty$, and $\mathbb{Q}$ is a Dedekind cut represented as $+\infty$.

Definition III.3.2. Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$. Suppose that for all $M \in \mathbb{R}$, there is an $N \in \mathbb{N}$ such that for all $n \geq N$, $x_n \geq M$. Then we say that $x_n \to +\infty$. If $-x_n \to +\infty$, we say that $x_n \to -\infty$.

Definition III.3.3. Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$. Let $E \subseteq \overline{\mathbb{R}}$ denote the set of subsequential limits of $(x_n)_n$ in the extended real line. The supremum of E is called the *upper limit* or *limit superior* of $(x_n)_n$, and is denoted by $\limsup_{n\to\infty} x_n$.

The infimum of E is called the *lower limit* or *limit inferior* of $(x_n)_n$, denoted $\liminf_{n\to\infty} x_n$.

Example. Let $(x_n = (-1)^n)_{n\in\mathbb{N}}$. Then $E = \{-1, +1\}$ so $\limsup_{n\to\infty} x_n = 1$ and $\liminf_{n\to\infty} x_n = -1$.

Theorem III.3.4. *Let $(x_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ be a sequence and E be the set of subsequential limits of $(x_n)_n$ in $\overline{\mathbb{R}}$.*

(1) E is non-empty.

(2) $\sup E$ and $\inf E$ are contained in E.

(3) If $x > \sup E$ (resp. $x < \inf E$), then there is an $N \in \mathbb{N}$ such that for all $n \geq N$, $x_n < x$ (resp. $x_n > x$).

(4) $\sup E$ (resp. $\inf E$) is the only element of $\overline{\mathbb{R}}$ statisfying both (2) and (3).

Proof. **(1)** If $(x_n)_n$ is bounded, then E is non-empty by theorem III.1.5.

Let $(x_n)_{n\in\mathbb{N}}$ be unbounded above. Let $n_0 = 0$, and for $k \geq 0$, let

$$n_{k+1} = \min\{n > n_k \mid x_n > x_{n_k}\}$$

This exists since $(x_n)_n$ is unbounded above.

Suppose $m \notin (n_k)_{k\in\mathbb{N}}$. Let k be such that $n_k < m < n_{k+1}$. $x_m > x_{n_k}$ would imply $n_{k+1} \leq m$, so $x_m \leq x_{n_k}$. This shows that each x_m not in the subsequence is bounded above by some element of the subsequence.

Thus $(x_{n_k})_k$ is unbounded above, for if it weren't, all of $(x_n)_n$ would be bounded above. So for every $M \in \mathbb{R}$, there is a K such that $x_{n_K} > M$, but since the subsequence is increasing, $x_{n_k} > M$ for all $k \geq K$. Thus $\lim x_{n_k} = +\infty$.

(2) If $\sup E = +\infty$, then for all $M \in \mathbb{R}$, there is an $e_M \in E$ larger than $M + 1$, so there is some x_n larger than M. Thus $(x_n)_n$ is unbounded above, so by the previous argument, $+\infty \in E$.

Now suppose $\sup E = x \in \mathbb{R}$. Let $\varepsilon_n = \frac{1}{2n}$. Let $n_0 = 0$. For $k > 0$, let e_k be an element of E larger than $x - \varepsilon_k$. Let $n_k > n_{k-1}$ be such that $x_{n_k} \in (e_k - \varepsilon_k, e_k + \varepsilon_k)$. Then $|x_{n_k} - x| < 2\varepsilon_k = \frac{1}{k}$. Thus $x_{n_k} \to x$, so $x \in E$. □

Example. Let $(x_n)_{n\in\mathbb{N}}$ be an enumeration of $\mathbb{Q}$. Then $E = \overline{\mathbb{R}}$.

Proof. Let $x \in \mathbb{R}$. Then for any $\varepsilon > 0$, there are infinitely many rationals that are ε-close to x. Thus $x \in E$.

For $x = \pm\infty$, replace "ε-close" with "larger/smaller than M". □

Theorem III.3.5.

(1) Suppose $x_n \le y_n$ for all $n \in \mathbb{N}$. Then

$$\liminf x_n \le \liminf y_n \quad \textit{and} \quad \limsup x_n \le \limsup y_n.$$

(2) $\lim x_n = x$ *iff* $\limsup x_n = \liminf x_n = x$.

Lecture 18.
Mon 12 Feb '24

Proof of theorem III.3.4 (continued). Let $\alpha^* = \sup E$.

(3) Suppose not. Let $x > \sup E$ such that for every $k \in \mathbb{N}$, there exists an $m(k) \ge k$ such that $x_{m(k)} \ge x$. Let $n_0 = m(0)$, and for $l \ge 1$, let $n_k = m(n_{k-1}+1)$. Then $n_0 < n_1 < n_2 < \cdots$ and $x_{n_k} \ge x$ for all k. Thus $\gamma = (x_{n_k})_k$ is a subsequence of $(x_n)_n$, but all subsequential limits of γ are at least $x > \sup E$. But a subsequential limit of γ is a subsequential limit of $(x_n)_n$, so $\sup E \ge x$, a contradiction.

(4) Suppose $y < z$ in $\overline{\mathbb{R}}$ satisfy both (2) and (3). That is, both y and z are sequential limits of $(x_n)_n$, and if $x > y$ (or $x > z$), then there exists an $N \in \mathbb{N}$ such that $x_n < x$ for all $n \ge N$.

Choose

$$x = \begin{cases} 0 & \text{if } y = -\infty, z = +\infty \\ z - 1 & \text{if } y = -\infty, z \in \mathbb{R} \\ y + 1 & \text{if } y \in \mathbb{R}, z = +\infty \\ \frac{y+z}{2} & \text{if } y, z \in \mathbb{R} \end{cases}$$

In each case, $y < x < z$. By (3) applied to x, all but finitely many x_n are less than x. By (2) applied to z, infinitely many x_n are greater than x. Contradiction.

$\square$

Theorem III.3.6.

(1) The following sequences admit limits in $\overline{\mathbb{R}}$*.*

$$y_n = \sup\{x_k : k \geq n\}$$
$$z_n = \inf\{x_k : k \geq n\}$$

(2) Moreover,

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} y_n$$
$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} z_n$$

where limits are taken in $\overline{\mathbb{R}}$*.*

Remark. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of subsets of X. Define

$$A^* = \limsup_{n\to\infty} A_n = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k$$
$$A_* = \liminf_{n\to\infty} A_n = \bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} A_k$$

Then $x \in A^*$ iff x is in infinitely many A_n, and $x \in A_*$ iff x is in all but finitely many A_n.

We say that $(A_n)_{n\in\mathbb{N}}$ *converges* if $A^* = A_*$.

We can characterize this using indicator functions.

$$\mathbf{1}_{A^*} = \limsup_{n\to\infty} \mathbf{1}_{A_n}$$
$$\mathbf{1}_{A_*} = \liminf_{n\to\infty} \mathbf{1}_{A_n}$$

which is to say that for each $x \in X$,

$$\mathbf{1}_{A^*}(x) = \limsup_{n\to\infty} \mathbf{1}_{A_n}(x)$$
$$\mathbf{1}_{A_*}(x) = \liminf_{n\to\infty} \mathbf{1}_{A_n}(x)$$

Proof. $(y_n)_n$ is a decreasing sequence, so it has a limit in $\overline{\mathbb{R}}$, since if it is not bounded, it converges to $-\infty$.

Let $y = \lim_{n\to\infty} y_n$. Since (y_n) is decreasing, given $k \in \mathbb{N}$, there exists an

$N(k) \in \mathbb{N}$ such that for all $n \geq N(k)$,

$$y \leq y_n < y + \frac{1}{k}.$$

But $y_n = \sup\{x_i : i \geq n\}$, so for all $n \geq N(k)$, there exists an $m(k,n)$ such that $y_n - \frac{1}{k} < x_{m(k,n)} \leq y_n$.

Let

$$\begin{aligned}
n_1 &= m(1, N(1)) \\
n_2 &= m(2, n_1 \vee N(2) + 1) > n_1 \vee N(2) \\
&\vdots \\
n_k &= m(k, n_{k-1} \vee N(k) + 1) > n_{k-1} \vee N(k)
\end{aligned}$$

$\square$

Assignment 5

Quiz 1 Mar 2024

> **Problem 3.1.** *Let $(x_n)_{n\in\mathbb{N}}$ be a convergent sequence in $\mathbb{R}$, with $x_n \geq 0$ for all $n \in \mathbb{N}$. Let $k \in \mathbb{N}^*$. Show that*
> $$\lim_{n\to\infty} (x_n)^{1/k} = \left(\lim_{n\to\infty} x_n\right)^{1/k}.$$

Solution. Call the limit of $(x_n)_n$ L. Let $\varepsilon > 0$ and let $\varepsilon' = \frac{\varepsilon}{L^{1/k}}$. Then for sufficiently small ε',

$$\begin{aligned}(L^{1/k} - \varepsilon)^k &\leq L(1 - k\varepsilon' + 2^k(\varepsilon')^2) < L \\ (L^{1/k} + \varepsilon)^k &\geq L(1 + k\varepsilon') > L\end{aligned}$$

But $x_n \to L$, so eventually $x_n \in (L(1 - k\varepsilon' + 2^k(\varepsilon')^2), L(1 + k\varepsilon'))$.
Then $x_n^{1/k} \in (L^{1/k} - \varepsilon, L^{1/k} + \varepsilon)$. ■

Solution. [Alternative] Let $L = \lim_{n\to\infty} x_n$. Then $\frac{x_n}{L} \to 1$. But notice that for any real $a > 0$,

$$\left|1 - a^{1/k}\right| \leq |1 - a|$$

because $x^k - 1 = (x-1)(x^{k-1} + x^{k-2} + \cdots + 1)$ where the second term is obviously larger than 1.

But then

$$\left(\frac{x_n}{x}\right)^{1/k} \to 1$$

which proves the result. ■

> **Problem 3.2.** *Let (X, d) be a complete metric space, and $Y \subseteq X$. Show that $(Y, d|_Y)$ is a complete metric space if and only if Y is closed in (X, d).*

Solution. Y is a complete metric space iff every Cauchy sequence in Y converges in Y. But X is complete, so every Cauchy sequence in Y converges in X. Thus, Y is complete iff every convergent sequence in Y (viewed as a sequence in X) converges in Y. This is true iff Y is closed in X. ■

> **Problem 3.3.** *Let (X, d) be a metric space and $A \subseteq X$ be a dense subset, i.e., $\overline{A} = X$. Show that if every Cauchy sequence in A converges to a limit in X, then X is a complete metric space.*

Solution. Let $(x_n)_n$ be a Cauchy sequence in X. For each $n \in \mathbb{N}$, there exists $a_n \in A$ such that $d(x_n, a_n) < \frac{1}{n}$. Then $(a_n)_n$ is a Cauchy sequence in A, so it converges to some $a \in X$. But $d(x_n, a_n) \to 0$, so $x_n \to a$. ■

Problem 3.4. *For any real sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ show that*

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

Solution. Let $X = \limsup_{n\to\infty} x_n$ and $Y = \limsup_{n\to\infty} y_n$. Then for any $z > X + Y$, rewrite z as $(X + \delta) + (Y + \delta) + \delta$. Then there is an N such that for all $n \geq N$,

$$x_n < X + \delta \quad \text{and} \quad y_n < Y + \delta$$

so that

$$x_n + y_n < z - \delta.$$

But then $z - \delta$ cannot be a subsequential limit of $(x_n + y_n)_n$. Thus

$$\limsup_{n\to\infty} x_n + y_n \leq X + Y.$$

■

Problem 3.5. *Compute* $\limsup_{n\to\infty} x_n$ *and* $\liminf_{n\to\infty} x_n$*, where the sequence* $(x_n)_{n\in\mathbb{N}^*} \subseteq \mathbb{R}$ *is given by*

$$x_1 = 0,$$
$$x_{2m} = \frac{x_{2m-1}}{2}, \quad m \geq 1,$$
$$x_{2m+1} = \frac{1}{2} + x_{2m}, \quad m \geq 1.$$

Solution. **Claim:** $x_{2m+1} = 1 - \frac{1}{2^m}$.

Proof. Induction. □

Corollary: $x_{2m} = \frac{1}{2} - \frac{1}{2^m}$.
Thus $\inf_{n\geq 2m} x_n = x_{2m}$. Then

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty}\left(\frac{1}{2} - \frac{1}{2^n}\right) = \frac{1}{2}.$$

For limsup, note that each term is less than 1, but 1 is a subsequential limit via the odd terms. Thus

$$\limsup_{n\to\infty} x_n = 1. \qquad \blacksquare$$

III.4 Series

Lecture 19.
Mon 26 Feb '24

We have seen infinite sums already, in the form of decimal expansions, and also in computing the "length" of the Cantor set.

Definition III.4.1. Let $(a_n)_{n\in\mathbb{N}} \subseteq \mathbb{C}$. We say that the infinite sum

$$\sum_{n=0}^{\infty} a_n$$

converges to a iff the sequence of partial sums $(S_n)_{n\in\mathbb{N}}$ given by

$$\begin{aligned} S_0 &= a_0 \\ S_1 &= a_0 + a_1 \\ &\vdots \\ S_n &= a_0 + a_1 + \cdots + a_n \end{aligned}$$

converges to a. In that case, a is said to be the sum of the given series. If $(S_n)_n$ diverges, then the series is said to diverge.

Example. Let $z \in \mathbb{C}$. Let $a_n = z^n$ for $n \in \mathbb{N}$. We wish to compute $\sum_{n=0}^{\infty} z^n$. Note that

$$\begin{aligned} S_n &= z^0 + z^1 + \cdots + z^n \\ \implies (1-z)S_n &= 1 - z^{n+1} \end{aligned}$$

so if $z \neq 1$,

$$S_n = \frac{1-z^{n+1}}{1-z}$$

We have three cases:

- If $|z| < 1$, then $\lim_{n\to\infty} |z|^{n+1} = 0$. So by the algebra of limits, $\lim_{n\to\infty} S_n = \frac{1}{1-z}$ so that the series converges to $\frac{1}{1-z}$. Here, we have used that for any complex sequence $(z_n)_n$, $z_n \to 0$ iff $|z_n| \to 0$.
- If $|z| > 1$, then $\lim_{n\to\infty} z^{n+1}$ does not exist, so the series diverges. So by the algebra of limits, $\lim_{n\to\infty} S_n$ does not exist.
- For $|z| = 1$, we consider a few special cases first.
  - If $z = 1$, then $S_n = n + 1 \to \infty$.

– If $z = -1$, then $S_n = \mathbf{1}_{n \text{ is even}}$ which is divergent.

Even for the general case, we can use theorem III.4.2 to conclude that the series diverges, since $z^n \not\to 0$.

Theorem III.4.2. *Let $(a_n)_{n\in\mathbb{N}} \subseteq \mathbb{C}$. Then*

(1) $\sum_{n=0}^{\infty} a_n$ converges iff for every $\varepsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n > m \geq N$, we have that $\left|\sum_{j=m}^{n-1} a_j\right| < \varepsilon$.

(2) If $\sum_{n=0}^{\infty} a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.

(3) Suppose $a_n \geq 0$ (a_n is real). Then $\sum_{n=0}^{\infty} a_n$ converges iff the sequence of partial sums is bounded above.

Proof.

(1) The metric space $(\mathbb{C}, |\cdot|) \cong (\mathbb{R}^2, \|\cdot\|)$ is complete, so we can use the Cauchy criterion for convergence applied to the sequence of partial sums.

(2) From (1) using $n = m + 1$.

(3) Monotone convergence theorem applied to the SOPS. □

Theorem III.4.3 (Cauchy condensation test)**.** *Let $a_1 \geq a_2 \geq \cdots \geq 0$. Then $\sum_{n=1}^{\infty} a_n$ converges iff $\sum_{n=1}^{\infty} 2^n a_{2^n}$ converges.*

Lecture 20.
Wed 28 Feb '24

Proof. Let the SOPS of the two series be $(S_n)_{n\geq 1}$ and $(T_n)_{n\geq 1}$ respectively. Note that we only need to show that $(S_n)_n$ is bounded above iff $(T_n)_n$ is.

Let k and n be such that $n \leq 2^k$. Then

$$\begin{aligned}
S_n &= a_1 + a_2 + a_3 + \cdots + a_n \\
&\leq a_1 + (a_2 + a_3) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1}) \\
&\leq a_1 + 2a_2 + \cdots + 2^k a_{2^k} \\
&= T_k
\end{aligned}$$

Thus if $(T_n)_n$ is bounded then so is $(S_n)_n$.

Now let $2^k < n$. Then

$$\begin{aligned}
S_n &= a_1 + a_2 + a_3 + \cdots + a_n \\
&\geq a_1 + a_2 + a_3 + \ldots a_{2^k} \\
&= a_1 + a_2 + (a_3 + a_4) + \cdots + (a_{2^{k-1}+1} + a_{2^k}) \\
&\geq a_1/2 + a_2 + 2a_4 + \cdots + 2^{k-1}a_{2^k} \\
&= \frac{1}{2}(a_1 + 2a_2 + 4a_4 + \ldots 2^k a_{2^k}) \\
&= T_k
\end{aligned}$$

Thus if $(S_n)_n$ is bounded then so is $(T_n)_n$. $\square$

Corollary III.4.4. $\sum_{n=1}^{\infty} \frac{1}{n^p}$ *is convergent iff* $p > 1$.

Proof. If $p \leq 0$, $\frac{1}{n^p} \not\to 0$, so the series cannot converge. If $p > 0$, then $\frac{1}{n^p}$ is decreasing and non-negative, so by the Cauchy condensation test, this converges iff $\sum_{n=1}^{\infty} 2^n \frac{1}{(2^n)^p}$ converges. But this is a geometric series with ratio 2^{1-p}. This converges iff $2^{1-p} < 1 \iff p > 1$. $\square$

Corollary III.4.5. $\sum_{n=2}^{\infty} \frac{1}{n(\log n)^p}$ *converges iff* $p > 1$.

Proof. If $p \leq 0$, then the series is bounded below by $\frac{1}{n}$ for $n \geq 3$. So by the comparison test, the series diverges.

For $p > 0$, the terms are decreasing and non-negative. $\sum_{n=1}^{\infty} \frac{2^n}{2^n n^p (\log 2)^p}$ converges iff $p > 1$ by the previous corollary. $\square$

Recall that in UMA101 we defined e as $\sum_{n=0}^{\infty} \frac{1}{n!}$.

Theorem III.4.6. $e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$.

Proof. Let $s_n = \sum_{j=0}^{n} \frac{1}{j!}$ and $t_n = \left(1 + \frac{1}{n}\right)^n$.

$$\begin{aligned}
t_n &= \sum_{j=0}^{n} \binom{n}{j} \frac{1}{n^j} \\
&= \sum_{j=0}^{n} \frac{1}{j!} \frac{n(n-1)\ldots(n-j+1)}{n^j} \\
&= 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \frac{1}{3!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \cdots + \\
&\quad \frac{1}{n!}\left(1 - \frac{1}{n}\right)\ldots\left(1 - \frac{n-1}{n}\right)
\end{aligned}$$

which clearly shows that $t_n \leq s_n$ so $\limsup_{n\to\infty} t_n \leq e$.

Fix $m \in \mathbb{N}$. For all $n \geq m \geq 3$,

$$\begin{aligned}\liminf_{n\to\infty} &\geq 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \cdots + \\ &\qquad \frac{1}{m!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{m-1}{n}\right) \\ &\xrightarrow[m \text{ fixed}]{n\to\infty} 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{m!} \\ &= S_m\end{aligned}$$

As $m \to \infty$,

$$\liminf_{n\to\infty} t_n \geq \lim_{m\to\infty} S_m = e$$

□

Definition III.4.7 (Formal power series)**.** Let $(c_n)_{n\in\mathbb{N}} \subseteq \mathbb{C}$. Then the formal sum

$$\sum_{n=0}^{\infty} c_n z^n$$

is called a *power series*. Here, z is an arbitrary complex number, and the convergence of this series depends on z.

Examples.

Lecture 21.
Thu 29 Feb '24

- $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ converges everywhere.
- $\sum_{n=0}^{\infty} z^n$ converges nowhere on $|z| = 1$.
-

III.4.1 Combining Series

Theorem III.4.8. *For $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ convergent to a and b respectively,*

(i) (Addition) $\sum(a_n + b_n)$ converges to $a + b$,

(ii) (Scalar product) $\sum ca_n$ converges to ca for any $c \in \mathbb{C}$,

(iii) (Termwise product) $\sum a_n b_n$ *need not converge.*

Proof. The first two parts are trivial. For the third, consider the series $\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}$, which converges by the alternating series test. However, taking its termwise product with itself gives $\sum_{n=1}^{\infty} \frac{1}{n}$, which diverges. □

Definition III.4.9 (Absolute convergence)**.** A series $\sum_{n=0}^{\infty} a_n$ is said to be *absolutely convergent* if $\sum_{n=0}^{\infty} |a_n|$ converges. If it converges but not absolutely, it is said to be *conditionally convergent*.

The counterexample in the proof of the theorem is an example of a conditionally convergent series.

Theorem III.4.10. *Let* $(a_n)_{n\in\mathbb{N}} \subseteq \mathbb{C}$ *and* $(b_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$*. Suppose that the SOPS of* $\sum_{n=0}^{\infty} a_n$ *is bounded, and that* $b_1 \geq b_2 \geq \ldots$ *with* $b_n \to 0$*. Then* $\sum_{n=0}^{\infty} a_n b_n$ *converges.*

Proof. Let $(A_n)_n < M$ be the SOPS of $\sum_{n=0}^{\infty} a_n$. Then for any $p < q$,

$$\begin{aligned}
\sum_{n=p}^{q} a_n b_n &= \sum_{n=p}^{q} (A_n - A_{n-1}) b_n \\
&= \sum_{n=p}^{q} A_n b_n - \sum_{n=p}^{q} A_{n-1} b_n \\
&= \sum_{n=p}^{q} A_n b_n - \sum_{n=p-1}^{q-1} A_n b_{n+1} \\
&= A_q b_q - A_{p-1} b_p + \sum_{n=p}^{q-1} A_n (b_n - b_{n+1})
\end{aligned}$$

Taking absolute values,

$$\begin{aligned}
\left| \sum_{n=p}^{q} a_n b_n \right| &\leq |A_q| b_q + |A_{p-1}| b_p + \sum_{n=p}^{q-1} |A_n| (b_n - b_{n+1}) \\
&\leq M b_q + M b_p + \sum_{n=p}^{q-1} M (b_n - b_{n+1}) \\
&= M(b_p + b_q) + M(b_p - b_q) \\
&= 2M b_p
\end{aligned}$$

Since $b_n \to 0$, $2Mb_p < \varepsilon$ for sufficiently large p. Thus $\sum_{n=0}^{\infty} a_n b_n$ converges by the Cauchy criterion. $\square$

Remarks.

- If $\sum a_n$ converges, then the condition on the SOPS is satisfied. But the condition on the b_n is much more special.
- The alternating series test is a special case of this theorem, with $a_n = (-1)^n$ and b_n decreasing to 0.
- Suppose $(b_n)_n$ is as in the theorem, and $\sum b_n z^n$ has radius of convergence 1. Then it converges for all $|z| = 1$, except possibly at $z = 1$. This is because $\sum_{n=0}^{\infty} z^n$ has bounded sops for each $|z| = 1$, $z \neq 1$.

  Proof. Let $(Z_n)_n$ be the sops. Since $z \neq 1$, $Z_n = \frac{1-z^{n+1}}{1-z}$. The numerator is bounded by 2, and the denominator is constant. $\square$

Corollary III.4.11 (Alternating series test)**.** *Let $(a_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}$ be decreasing to 0. Then $\sum_{n=0}^{\infty}(-1)^n a_n$ converges.*

Definition III.4.12. Given two formal series $\sum_{n\in\mathbb{N}} a_n$ and $\sum_{n\in\mathbb{N}} b_n$, their *Cauchy product* is the series $\sum_{n\in\mathbb{N}} c_n$ where

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}$$

This is motivated by the termwise product of power series.

Lecture 22. Mon 04 Mar '24

If $\sum_{n\in\mathbb{N}} a_n$ and $\sum_{n\in\mathbb{N}} b_n$ converge, does $\sum_{n\in\mathbb{N}} c_n$ have to converge?

No. Take the conditionally convergent series $a_n = b_n = \frac{(-1)^n}{\sqrt{n}}$. Then

$$\begin{aligned}
c_n &= \sum_{k=1}^{n} a_k b_{n+1-k} \\
&= \sum_{k=1}^{n} \frac{(-1)^{n+1}}{\sqrt{k}\sqrt{n+1-k}} \\
|c_n| &= \sum_{k=1}^{n} \frac{1}{\sqrt{k}\sqrt{n+1-k}} \\
&\geq \sum_{k=1}^{n} \frac{1}{\sqrt{n}\sqrt{n}} \\
&= \frac{n}{n} = 1.
\end{aligned}$$

Thus $\sum_{n\in\mathbb{N}} c_n$ cannot converge.

Theorem III.4.13 (Mertens)**.** *Suppose that*

(i) $\sum_{n\in\mathbb{N}} a_n$ *converges absolutely to* A*, and*

(ii) $\sum_{n\in\mathbb{N}} b_n$ *converges to* B*.*

Then their Cauchy product $\sum_{n\in\mathbb{N}} c_n$ *converges to* AB*.*

Proof. Let $(A_n)_{n\in\mathbb{N}}$, $(B_n)_{n\in\mathbb{N}}$ and $(C_n)_{n\in\mathbb{N}}$ be the sops of $\sum a_n$, $\sum b_n$ and $\sum c_n$ respectively.

$$\begin{aligned}
C_n &= c_0 + c_1 + \cdots + c_n \\
&= a_0 b_0 + (a_0 b_1 + a_1 b_0) + \cdots + (a_0 b_n + \cdots + a_n b_0) \\
&= a_0 B_n + a_1 B_{n-1} + \cdots + a_n B_0 \\
&= A_n B + (a_0(B_n - B) + \cdots + a_n(B_0 - B)).
\end{aligned}$$

Introduce the notation β_k for $B_k - B$.

$$= A_n B + (a_0 \beta_n + \cdots + a_n \beta_0).$$

Define $\delta_n = a_0 \beta_n + \cdots + a_n \beta_0$. It suffices to show that

$$\lim_{n\to\infty} \delta_n = 0.$$

Let $\alpha = \sum_{n\in\mathbb{N}} |a_n|$. Since $\lim_{n\to\infty} \beta_n = 0$, we have that for every $\epsilon > 0$, there exists $m \in \mathbb{N}$ such that $|\beta_n| < \varepsilon$ for all $n \geq m$.

Fix an $\varepsilon > 0$ and choose an appropriate m. Then for $n \geq m$,

$$\begin{aligned}|\delta_n| &= |a_0\beta_n + \cdots + a_n\beta_0| \\ &= \left|a_0\beta_n + \cdots + a_{n-(m-1)}\beta_{m-1}\right| + \varepsilon \sum_{j=m}^{n} |a_j|\end{aligned}$$

Taking limit as $n \to \infty$,

$$\limsup_{n\to\infty} |\delta_n| \leq \varepsilon\alpha.$$

Since ε was arbitrary, we have that $\limsup_{n\to\infty}|\delta_n| = 0$ and so $|\delta_n| \to 0$. This gives $\delta_n \to 0$ so that $C_n = A_n B + \delta_n \to AB$. □

III.5 Rearrangements

Conditionally convergent real series have the remarkable feature that their terms can be rearranged to converge to any real number.

Example. Consider the series $\sum \frac{(-1)^n}{n}$. Let $(S_n)_{n\in\mathbb{N}^*}$ be its partial sums. Note that $S_3 > S_4, S_5, S_6, \ldots$. So the limit is less that $1 - \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$.

Now consider the rearranged series

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \ldots.$$

We can show that this converges, by rewriting it as

$$1 + \frac{1}{3} - \frac{2}{4} + \frac{1}{5} + \frac{1}{7} - \frac{2}{8} + \frac{1}{9} + \frac{1}{11} - \frac{2}{12} + \ldots$$

and applying theorem III.4.10 to

$$\begin{aligned}(a_n)_{n\in\mathbb{N}} &= (1, 1, -2, 1, 1, -2, 1, 1, -2, \ldots) \\ (b_n)_{n\in\mathbb{N}} &= \left(1, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{7}, \frac{1}{8}, \frac{1}{9}, \frac{1}{11}, \frac{1}{12}, \ldots\right)\end{aligned}$$

Grouping terms in groups of three, we notice that each group is positive. So the rearranged series converges to a number greater than $1 + \frac{1}{3} - \frac{1}{2} = \frac{5}{6}$. Thus rearranging the terms has created a series that has a different sum.

Chapter IV

Functional Limits & Continuity

IV.1 Definitions

Lecture 23.
Wed 06 Mar '24

Definition IV.1.1 (Limit)**.** Let $f\colon E \to Y$ be a function from a subset E of a metric space (X, d_X) to another metric space (Y, d_Y). Let $p \in X$ be a limit point of E. We say that

$$\lim_{x \to p} f(x) = q \in Y$$

if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that whenever $0 < d_X(x, p) < \delta$ and $x \in E$, then $d_Y(f(x), q) < \varepsilon$.

Remarks.

- p need not be in E, *i.e.*, f need not be defined at p.
- Even if f is defined at p, we are not requiring that $f(p) = q$. For example, consider the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by

  $$f(x) = 0^{0^{x^2}}.$$

  Then $\lim_{x\to 0} f(x) = 1$ even though $f(0) = 0$.

Theorem IV.1.2 (Sequential characterization of limits)**.** *Let X, Y, E, f and p be as in definition IV.1.1. Then*

$$\lim_{x \to p} f(x) = q$$

if and only if for every sequence $(p_n)_{n\in\mathbb{N}} \subseteq E \setminus \{p\}$ such that $p_n \to p$, we have $f(p_n) \to q$.

Proof. Suppose that $\lim_{x\to p} f(x) = q$. Let $(p_n)_{n\in\mathbb{N}} \subseteq E \setminus \{p\} \to p$. Let $\varepsilon > 0$. Then there exists a $\delta > 0$ such that whenever $d_X(x,p) < \delta$ and $x \in E \setminus \{p\}$, then $d_Y(f(x),q) < \varepsilon$.

But since $(p_n)_{n\in\mathbb{N}} \to p$, there exists an $N \in \mathbb{N}$ such that for all $n \geq N$, $d_X(p_n,p) < \delta$. So for all $n \geq N$, $d_Y(f(p_n),q) < \varepsilon$, which proves that $f(p_n) \to q$.

For the converse, we will prove the contrapositive. Now suppose that $\lim_{x\to p} f(x) \neq q$. That is, there is an $\varepsilon > 0$ such that for every $\delta > 0$, there exists a $p_\delta \in E$ such that $0 < d_X(p_\delta,p) < \delta$ but $d_Y(f(p_\delta),q) \geq \varepsilon$.

Consider the sequence $(p_{\frac{1}{n}})_{n\in\mathbb{N}}$. Then $p_{\frac{1}{n}} \to p$, but $f(p_{\frac{1}{n}}) \not\to q$. Thus we have constructed a sequence that does not satisfy the condition in the theorem. $\square$

Corollary IV.1.3.

(i) Functional limits are unique.

(ii) Let $f,g\colon E \to \mathbb{C}$ and p be a limit point of E. Assume that limits $\lim_{x\to p} f(x) = q$ and $\lim_{x\to p} g(x) = r$ exist. Then

$$\begin{aligned}
\lim_{x\to p}(f(x)+g(x)) &= q+r,\\
\lim_{x\to p}(f(x)-g(x)) &= q-r,\\
\lim_{x\to p}(f(x)\cdot g(x)) &= q\cdot r,\\
\lim_{x\to p}(f(x)/g(x)) &= q/r \quad \textit{if } r \neq 0.
\end{aligned}$$

(iii) Let $f,g\colon E \to \mathbb{R}^m$ and p, q, r be as above. Then

$$\begin{aligned}
\lim_{x\to p}(f(x)+g(x)) &= q+r,\\
\lim_{x\to p}\langle f(x),\, g(x)\rangle &= \langle q,\, r\rangle
\end{aligned}$$

Definition IV.1.4 (Continuity)**.** Let X and Y be metric spaces, $E \subseteq X$ and $f\colon E \to Y$. Let $p \in E$. Then we say that f is *continuous at p* if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in E$ with $d_X(x,p) < \delta$, we have $d_Y(f(x),f(p)) < \varepsilon$.

Remarks.

- If p is a limit point of E, then f is continuous at p iff $\lim_{x\to p} f(x)$ exists and equals $f(p)$.
- If p is an isolated point of E, then any f is continuous at p.

Theorem IV.1.5 (Sequential characterization of continuity)**.** *Let X, Y, E, f and p be as in definition IV.1.4. Then f is continuous at p if and only if for every sequence $(p_n)_{n\in\mathbb{N}} \subseteq E$ converging to p, we have $f(p_n) \to f(p)$.*

Corollary IV.1.6.

(i) Let $f, g\colon E \to \mathbb{C}$ and $p \in E$ such that f and g are continuous at p. Then $f+g$, $f-g$, $f\cdot g$ and f/g (if $g(p) \neq 0$) are continuous at p.

(ii) Let $f, g\colon E \to \mathbb{R}^m$ and p, q, r be as above. Then $f+g$ and $\langle f, g\rangle$ are continuous at p.

Exercise IV.1.7 (Composition of continuous functions)**.** *Let X, Y, Z be metric spaces and $E \subseteq X$. Let $f\colon E \to Y$ and $g\colon f(E) \to Z$. If f is continuous at $p \in E$ and g is continuous at $f(p)$, then $g \circ f$ is continuous at p.*

Proof. Let $p \in E$ be as in the statement and let $z = g(f(p))$. Let $\varepsilon > 0$. By continuity of g at $f(p)$, there exists a $\delta > 0$ such that for all $y \in B_{f(E)}(f(p);\delta)$, we have $g(y) \in B_Z(z;\varepsilon)$.

But by the continuity of f at p, there exists a $\delta' > 0$ such that for all $x \in B_E(p;\delta')$, we have $f(x) \in B_{f(E)}(f(p);\delta)$, so that $(g\circ f)(x) \in B_Z(z;\varepsilon)$.

Thus $g \circ f$ is continuous at p. □

Lecture 24.
Thu 07 Mar '24

Examples (continuous functions).

- All polynomials $p \in \mathbb{R}[x]$ are continuous.
- All rational functions p/q, where $p, q \in \mathbb{R}[x]$ are continuous.
- The exponential function $\exp\colon \mathbb{R} \to \mathbb{R}$ is continuous.
- The logarithm function $\log\colon (0,\infty) \to \mathbb{R}$ is continuous.
- $x \in \mathbb{R} \mapsto \log_x(a)$, or simply the reciprocal of log, is continuous.

IV.2 Topology & Continuity

Theorem IV.2.1. *Let (X, d_X) and (Y, d_Y) be metric spaces. Let $f\colon X \to Y$ be a function. Then the following are equivalent:*

(i) f is continuous on X.

(ii) $f^{-1}(U)$ is open in X for every open set U in Y.

(iii) $f^{-1}(C)$ is closed in X for every closed set C in Y.

Proof. For (ii) $\iff$ (iii), use that $f^{-1}(Y \setminus C) = X \setminus f^{-1}(C)$.

We now show (i) $\iff$ (ii), Suppose f is continuous on X. Let U be open in Y and $x \in f^{-1}(U)$. Then $f(x) \in U$. Since U is open, there exists an $\varepsilon > 0$ such that $B_Y(f(x); \varepsilon) \subseteq U$. Since f is continuous at x, there exists $\delta > 0$ such that for all $x' \in X$ with $d_X(x, x') < \delta$, $d_Y(f(x), f(x')) < \varepsilon$ so that $f(x') \in U$. Then $B_X(x; \delta) \subseteq f^{-1}(U)$.

Conversely, suppose $f^{-1}(U)$ is open for every open set U in Y. Let $x \in X$ and $\varepsilon > 0$. Choose $U = B_Y(f(x); \varepsilon)$. Then $f^{-1}(U)$ is open, so there exists $\delta > 0$ such that $B_X(x; \delta) \subseteq f^{-1}(U)$. Then for all $x' \in X$ with $d_X(x, x') < \delta$, $d_Y(f(x), f(x')) < \varepsilon$. $\square$

Remark. Continuous functions need not map open sets to open sets, nor closed sets to closed sets.

$x \mapsto 1$ on $\mathbb{R}$ is continuous, but maps open sets to the closed set $\{1\}$.

$x \mapsto \sin x$ on $\mathbb{R}$ is continuous, but the image of the closed set $\mathbb{N}^*$ doesn't contain 0 (π is irrational), even though it is a limit point of the image.

Theorem IV.2.2 (Compactness)**.** *If $f\colon X \to Y$ is continuous and $K \subseteq X$ is compact, then $f(K)$ is compact.*

Remarks.

- The pre-image of a compact set under a continuous function need not be compact.
- Even if the image of every compact set under a function f is compact, f need not be continuous.

Lecture 25.
Wed 13 Mar '24

Definition IV.2.3 (Uniform continuity)**.** A function $f\colon X \to Y$ is said to be *uniformly continuous* if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $x \in X$, $f(B_X(x; \delta)) \subseteq B_Y(f(x); \varepsilon)$.

Lecture 26.
Wed 13 Mar '24

Theorem IV.2.4. *Let $f\colon X \to Y$ be continuous on a compact set X. Then f is uniformly continuous.*

Proof. Let $\varepsilon > 0$ be given. For each $x \in X$, choose $\delta_x > 0$ such that $f(B_X(x;\delta_x)) \subseteq B_Y(f(x);\varepsilon)$. The collection $\left\{B_X(x;\frac{1}{2}\delta_x)\right\}_{x\in X}$ is an open cover of X. Since X is compact, only finitely many of these are needed to cover X. Label the centers $x_1, \dots, x_n$ and the corresponding radii $\frac{\delta_1}{2}, \dots, \frac{\delta_n}{2}$. Let $\delta > 0$ be the smallest of these radii.

Let $p, q \in X$ be such that $d_X(p,q) < \delta$. Then there exists an i such that $d_X(p,x_i) < \frac{\delta_i}{2}$. But $d_X(p,q) < \delta < \frac{\delta_i}{2}$, so $d_X(q,x_i) < \delta_i$ by the triangle inequality.

Thus $f(p)$ and $f(q)$ are both at most ε away from $f(x_i)$, so they are at most 2ε away from each other.

Thus for any $2\varepsilon > 0$, we have produced a $\delta > 0$ such that any points within δ of each other are mapped to points within 2ε of each other. $\square$

Theorem IV.2.5. *Continuous functions map connected sets to connected sets.*

Proof. Let X be connected and $f\colon X \to Y$ continuous. Suppose $f(X)$ is not connected. That is, there exist nonempty sets $A, B \subseteq f(X)$ such that $A \cup B = f(X)$ and $\overline{A} \cap B = A \cap \overline{B} = \varnothing$.

Let $A_* = f^{-1}(A)$ and $B_* = f^{-1}(B)$. They are nonempty since A and B are nonempty and in the range. Also, $X = f^{-1}(f(X)) = f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$. So A_* and B_* cover X.

Recall that f is continuous implies that the preimage of a closed set is closed. $A_* = f^{-1}(A) \subseteq f^{-1}(\overline{A})$, so $\overline{A_*} \subseteq f^{-1}(\overline{A})$.

Then

$$\begin{aligned}\overline{A_*} \cap B_* &\subseteq f^{-1}(\overline{A}) \cap f^{-1}(B)\\ &\subseteq f^{-1}(\overline{A} \cap B)\\ &= f^{-1}(\varnothing) = \varnothing.\end{aligned}$$

Similarly, $A_* \cap \overline{B_*} = \varnothing$.

Thus A_* and B_* are a separation of X, contradicting the connectedness of X. $\square$

IV.3 Discontinuities

Examples.

- The *Heaviside function* defined by

$$H(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$$

- The *sign function* defined by

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

- $$f(x) = \begin{cases} \frac{3x - x^2}{x(x^2+2)} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

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

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

- The *Dirichlet function*

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{otherwise} \end{cases} = \mathbf{1}_{\mathbb{Q}}.$$

  This is discontinuous *everywhere.*

Definition IV.3.1. Given $f\colon (a,b) \to \mathbb{R}$ and $c \in (a,b)$, we say that

(i) f has a *simple discontinuity* or a *discontinuity of the first kind* at c if the left-hand and right-hand limits $\lim_{x \to c^-} f(x)$ and $\lim_{x \to c^+} f(x)$ exist but are either unequal, or unequal to $f(c)$.

(ii) f has a *discontinuity of the second kind* at c if either the left-hand or right-hand limit does not exist.

The first three examples above have simple discontinuities at 0. The last three have discontinuities of the second kind. The third example has a *removable* discontinuity at 0, since both one-sided limits exist and are equal.

Theorem IV.3.2. *Monotone functions do not have discontinuities of the second kind.*

Corollary IV.3.3. *Monotone functions have at most countably many discontinuities.*

Example. Discontinuities of monotone functions can be dense.

Let $D \subseteq (0,1)$ be a dense countable set. Let $\{x_1, x_2, \ldots\}$ be an enumeration. Define

$$f(x) = \sum_{n: x_n < x} \frac{1}{n^2}.$$

Lecture 27.
Thu 14 Mar '24

IV.4 Mean Value Theorems & Applications

Lecture 28.
Mon 18 Mar '24

Example. Let

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

We show that f is differentiable on $\mathbb{R}$, but f' is discontinuous at 0.

Proof. When $x \neq 0$, we use the fact that polynomials and trignometric functions are differentiable, so that

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}.$$

At $x = 0$, we have

$$\lim_{h \to 0} \frac{f(h) - f(0)}{h} = \lim_{h \to 0} h \sin \frac{1}{h} = 0.$$

Thus the derivative is well-defined with

$$f'(x) = \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

□

Theorem IV.4.1. *Let $f \colon [a,b] \to \mathbb{R}$ be a function. Suppose f attains a local extremum at $c \in (a,b)$ and f is differentiable at c. Then $f'(c) = 0$.*

Proof. Suppose WLOG that f attains a local maximum at c. Then we have $f(c) \geq f(x)$ for all x close to c. But then for $h > 0$,

$$\frac{f(c+h) - f(c)}{h} \leq 0$$

and

$$\frac{f(c-h) - f(c)}{-h} \geq 0.$$

If the left hand and right hand limits of the difference quotient exist and are equal, then they must both be zero. $\square$

Exercise IV.4.2. *Let $f\colon (a,b) \to \mathbb{R}$ be a function. Suppose f is differentiable at $c \in (a,b)$ and $f'(c) > 0$. Then does there exist an interval around c such that f is increasing on that interval?*

Exercise IV.4.3. *Let $f\colon (a,b) \to \mathbb{R}$ be differentiable. Suppose $f'(x) \geq 0$ for all $x \in (a,b)$, and furthermore that*

- *f' is not identically zero on any interval. Is f strictly increasing on (a,b)?*
- *f' is zero on a discrete set of points. Is f strictly increasing on (a,b)?*

Theorem IV.4.4 (Intermediate value property)**.** *Suppose $f\colon (p,q) \to \mathbb{R}$ is differentiable and $[a,b] \subseteq (p,q)$. Suppose $f'(a) < \lambda < f'(b)$. Then there exists a $c \in (a,b)$ such that $f'(c) = \lambda$.*

Exercise IV.4.5. *If $g\colon (a,b) \to \mathbb{R}$ has a simple discontinuity at $c \in (a,b)$, then g does not satisfy the intermediate value property on some neighbourhood of c.*

Proof. We have two cases.

$\boldsymbol{f(c^-) = f(c^+) = L \neq f(c)}$**.** Choose $\lambda = \frac{1}{2}(L + f(c))$. In some δ neighbourhood of c (excluding c itself), we have $f(x) - L < \frac{1}{2}(f(c) - L)$ so that $f(x) < \lambda$. At c itself, we have $f(c) > \lambda$. Thus λ is never attained.

$\boldsymbol{f(c^-) \neq f(c^+)}$**.** Choose λ between the two limits, but unequal to $f(c)$. $\square$

Corollary IV.4.6. *Let f and $[a,b]$ be as in theorem IV.4.4. Then f' only has discontinuities of the second kind.*

Proof of theorem IV.4.4. Let $g(x) = f(x) - \lambda x$ for $x \in (p,q) \supseteq [a,b]$. Then $g'(x) = f'(x) - \lambda$ so that $g'(a) < 0 < g'(b)$. This means that g is strictly decreasing from a and strictly increasing to b.

Thus g attains a minimum at some $c \in (a,b)$. Then $g'(c) = 0$ so that $f'(c) = \lambda$. ∎

Theorem IV.4.7 (Generalised mean value theorem)**.** *Let $f, g\colon [a,b] \to \mathbb{R}$ be continuous functions that are differentiable on (a,b). Then there exists a $c \in (a,b)$ such that*

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$

Remark. If $g(x) = x$, we recover the mean value theorem. If furthermore $f(a) = f(b)$, we recover Rolle's theorem.

Lecture 29.
Wed 20 Mar '24

Proof. Let

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

Then

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

Thus $h(a) = h(b)$. By Rolle's theorem, there exists $c \in (a,b)$ such that $h'(c) = 0$. This proves the result. ∎

Theorem IV.4.8 (Taylor's theorem)**.** *Let $n \in \mathbb{N}^*$. Suppose that $f\colon (a,b) \to \mathbb{R}$ is n times differentiable on (a,b). Further assume that $f, f', \dots, f^{(n-1)}$ extend continuously to $[a,b]$. Then there exists a point $c \in (a,b)$ such that*

$$f(b) = f(a) + f'(a)(b-a) + \dots + \frac{f^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1} + \frac{f^{(n)}(c)}{n!}(b-a)^n.$$

Proof. Suppose the result holds in the case when

$$f(a) = f'(a) = \dots = f^{(n-1)}(a) = 0. \tag{$\star$}$$

Let f be given as per the theorem. Define

$$\begin{aligned} F(x) &= f(x) - \sum_{j=0}^{n-1} \frac{f^{(j)}(a)}{j!}(x-a)^j \\ \implies F^{(k)}(x) &= f^{(k)}(x) - f^{(k)}(a) - \sum_{j=k+1}^{n-1} a_j (x-a)^{j-k} \\ \implies F^{(k)}(a) &= 0 \end{aligned}$$

for every $k \in \{0, 1, \ldots, n-1\}$.

Then from the case ($\star$), there exists $c \in (a, b)$ such that

$$F(b) = \frac{F^{(n)}(c)}{n!}(b-a)^n.$$

But $F^{(n)} = f^{(n)}$, so this immediately gives

$$f(b) = f(a) + f'(a)(b-a) + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1} + \frac{f^{(n)}(c)}{n!}(b-a)^n.$$

But how do we prove the case ($\star$)?

Suppose f satisfies ($\star$). Define

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

This ensures that $g(a) = g(b) = 0$, and since derivatives of f extend continuously to $[a, b]$, so do the derivatives of g. Then $g'(a) = g''(a) = \cdots = g^{(n-1)}(a) = 0$. For the nth derivative,

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

Now apply Rolle's theorem iteratively to g, g', ..., $g^{(n-1)}$.

(1) From the first application, there exists $c_1 \in (a, b)$ such that $g'(c_1) = 0$.

(2) But $g'(a)$ is also zero, so by a second application there exists $c_2 \in (a, c_1)$ such that $g''(c_2) = 0$.

⋮

($n-1$) Continuing in this way, we find c_{n-1} such that $g^{(n-1)}(c_{n-1}) = 0$.

(n) In the final application, we find $c_n \in (a, c_{n-1})$ such that

$$g^{(n)}(c_n) = f^{(n)}(c_n) - \frac{f(b)n!}{(b-a)^n} = 0.$$

Thus finally, we have that there exists a point $c_n \in (a,b)$ such that $f(b) = \frac{(b-a)^n}{n!} f^{(n)}(c_n)$. This proves the case ($\star$), and hence the theorem. $\square$

Theorem IV.4.9 (L'Hôpital's rule)**.** *Let $a, b \in \overline{\mathbb{R}}$. Suppose $f, g\colon (a,b) \to \mathbb{R}$ are differentiable on (a,b), and that g and g' are never zero on (a,b). Suppose also that there exists an $A \in \overline{\mathbb{R}}$ such that*

$$\lim_{x \to a^+} \frac{f'(x)}{g'(x)} = A.$$

Then

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = A$$

whenever any of the following conditions hold.

(i) $\lim_{x \to a^+} f(x) = \lim_{x \to a^+} g(x) = 0$*, or*

(ii) $\lim_{x \to a^+} g(x) = +\infty$*, or*

(iii) $\lim_{x \to a^+} g(x) = -\infty$.

Proof. We will only consider the case when $a, A \in \mathbb{R}$.

Let $\varepsilon > 0$. Then there exists $\delta > 0$ such that for every $x \in (a, a+\delta)$,

$$A - \varepsilon < \frac{f'(x)}{g'(x)} < A + \varepsilon.$$

Fix a $y \in (a, a+\delta)$. Let $(z_n)_{n \in \mathbb{N}} \subseteq (a, y)$ be a sequence converging to a such that $g(z_n)$ is never $g(y)$.

Why does such a sequence exist? We can construct this iteratively. Since g' is never zero on (a,b), it cannot be constant on any interval. Thus every interval $(a, a + \frac{1}{n}) \cap (a,b)$ contains a point z_n such that $g(z_n) \neq g(y)$. This gives us the sequence $(z_n)_{n \in \mathbb{N}}$.

Then by the generalized mean value theorem, for every $n \in \mathbb{N}$, there exists a $w_n \in (z_n, y)$ such that

$$\frac{f(y) - f(z_n)}{g(y) - g(z_n)} = \frac{f'(w_n)}{g'(w_n)}.$$

But then

$$A - \varepsilon < \frac{f(y) - f(z_n)}{g(y) - g(z_n)} < A + \varepsilon.$$

Taking limits as $n \to \infty$, we find

$$A - \varepsilon < \frac{f(y)}{g(y)} < A + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this proves the result. $\square$