MATH 415: HONORS ANALYSIS I

- Textbook is "Principles of Mathematical Analysis" by Walter Rudin (3rd ed).

- Content: we aim to cover Chapters $1 \to 5$ and 7 (time permitting).

- Goal: Make precise the concepts of Calculus (limits, derivatives etc.)

//

We start with some notation: (some people omit '0')

- The natural/counting numbers are denoted $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.
- The integers/whole numbers are denoted $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.
- The rationals/fractions are denoted $\mathbb{Q} = \{\frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0\}$.

[Here we assume familiarity with set notation, all symbolic statements can be written in words however!]

Familiarity with the rationals is assumed, we will not give an axiomatic treatment here (but this can be done, e.g. by Peano axioms).

//

The real and complex numbers

For many purposes the rationals are not 'big' enough:

Example: There is no $x \in \mathbb{Q}$ such that $x^2 = 2$.

Proof: We argue by contradiction and assume there was such an $x \in \mathbb{Q}$. We then write

$$x = \frac{p}{q} \text{ where } p, q \in \mathbb{Z},\ q \neq 0,$$ and at most one is even (ie. $\gcd(p, q) = 1$).

We now show this gives a false statement.

As $x^2=2$ we see that

$$\left(\frac{p}{q}\right)^2=2 \underset{\text{(if and only if)}}{\iff} \frac{p^2}{q^2}=2 \iff p^2=2q^2 \underset{\text{(implies)}}{\implies} p^2 \text{ is even.}$$

Now, if p were odd then $p=2k+1$ for some $k\in\mathbb{Z}$, but then

$$p^2=(2k+1)^2=4k^2+4k+1=2\underbrace{(2k^2+2k)}_{(\in\mathbb{Z})}+1,$$

which would imply p^2 is odd!

Thus, p is also even. Hence $p=2l$ for some $l\in\mathbb{Z}$ and thus as

$$p^2=2q^2 \implies 2q^2=(2l)^2=4l^2 \implies q^2=2l^2$$
$$\implies q^2 \text{ is even.}$$

The same argument as above (p^2 even $\Rightarrow$ p even) shows that q must also be even!

However, then both p and q are even, contradicting our assumption that at most one was even; thus there is no rational x such that $x^2=2$. ∎

In fact, there is not even a 'best' approximation in $\mathbb{Q}$ for the equation $x^2=2$. By considering the sets

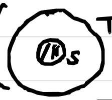
$$A=\{a\in\mathbb{Q} \mid a^2<2\}, \quad B=\{b\in\mathbb{Q} \mid b^2>2\},$$

we can see that there is no largest rational in A or smallest rational in B! (e.g. set $q=\frac{2p+2}{p+2}$ if $p>0$)

We will introduce the real (and complex numbers) to 'fill' these gaps in. To do this we will first discuss the notion of ordered sets and fields, as well as set theory that is of general use.

Set notation: Let S be a set:

- $x \in S$ means x is an element/member of S.
- $x \notin S$ means x is not an element/member of S.
- $S = \emptyset$, the empty set if S has no elements/members.
- S is nonempty if $S \neq \emptyset$; i.e. S has an element/member.
- $S \subset T$ means every element of S is an element of T. We then say that S is a subset of T.
- If $S \subset T$ and $S \neq T$ then S is a proper subset.

( $\leftarrow$ $S \subset T$ with $S \neq T$.)

Orders, Supremum + infimum

Definition: An order on a set S is a relation, $<$, with the properties that:

1. For $x, y \in S$ exactly one of the statements
$$x < y, \quad x = y, \quad y < x$$
is true.

2. For $x, y, z \in S$ then if (transitivity)
$$x < y \text{ and } y < z \text{ then } x < z.$$

We call a set with an order an ordered set, and for $x, y \in S$ write $x \leq y$ if $x < y$ or $x = y$ (i.e. $x \not> y$).

Examples: $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ are all ordered by $x < y \iff y - x$ is positive.

Definition: Let S be an ordered set and $E \subset S$, we say that E is bounded above if there is some $\beta \in S$ such that
$$x \leq \beta \text{ for all } x \in E,$$
and call β an upper bound for E. Lower bounds are defined in the same way (swapping $\leq$ for $\geq$).

[Here we conviniently write $y > x$ for $x < y$].

Example: $A = \{a \in \mathbb{Q} \mid a^2 < 2\}$ is bounded above by 2 $\left(\text{as } 2^2 = 4 > 2\right)$.

Definition: Let S be an ordered set and $E \subset S$ be bounded above. If there exists $\alpha \in S$ such that

1. α is an upper bound for E.

2. If $\gamma < \alpha$, then γ is not an upper bound for E.
   ($\Leftrightarrow$ If γ is an upper bound then $\alpha \leq \gamma$)

then α is the supremum / least upper bound for E and we write $\alpha = \sup E$.
The notion of infimum / greatest lower bound is defined similarly.

Examples:
- The sets A/B as before are bounded above/below but, as noted, have no supremum / infimum respectively.

- $\mathbb{N} \subset \mathbb{Z}$ is bounded below by 0, a lower bound for $\mathbb{N}$, and for all $n \in \mathbb{Z}$ with $0 < n$, n is not a lower bound for $\mathbb{N} \Rightarrow \inf \mathbb{N} = 0$.

- Let $E = \{ \frac{1}{n} \mid n \geq 1 \}$, then $\sup E = 1$, $\inf E = 0$ but $0 \notin E$!

Remark: As the last example shows, the supremum/infimum of a set, when it exists, may or may not be an element.

Definition: An ordered set S has the least upper bound property if whenever $E \subset S$ is non-empty and bounded above, then $\sup E$ exists. The greatest upper bound property is defined similarly.

Examples:
- $\mathbb{Q}$ does not have either property. (A, B sets as before)

- Both $\mathbb{N}$ and $\mathbb{Z}$ have these properties. (check!)

In fact, $\mathbb{N}$ has the well ordering principle: every $E \subset \mathbb{N}$ non-empty has a smallest element.

<u>Theorem:</u> Let S be an ordered set with the least upper bound property, $B \subset S$ non-empty and bounded below. Let L be the set of all lower bounds for B, then $\sup L$ exists and is equal to $\inf B$; which in particular exists.

(i.e. least upper bound property $\Leftrightarrow$ greatest lower bound property)

<u>Proof:</u> $L \neq \emptyset$ since B is bounded below, and since $y \leq x$ for each $y \in L$ and $x \in B$, <u>every</u> $x \in B$ is an upper bound for L. Thus, L is bounded above ($B \neq \emptyset$) and hence by the least upper bound property, $\sup L$ exists.

As each $x \in B$ is an upper bound for L we have

$$\sup L \leq x \text{ for any } x \in B,$$

so that $\sup L$ is a lower bound for B. For any lower bound $y \in L$ of B we have

$$y \leq \sup L$$

and thus $\sup L$ must be the greatest lower bound for B; i.e. $\sup L = \inf B$. ∎

End L1

<u>Fields and $\mathbb{R}$</u>

Given rational numbers we can add, multiply and more generally manipulate them in the usual manner; the concept of a field generalises this:

<u>Definition:</u> A <u>field</u> is a set F with two operations, called <u>addition</u> (denoted $+$) and <u>multiplication</u> (denoted $\cdot$), which satisfy the '<u>field axioms</u>':

- $x, y \in F \Rightarrow x+y \in F$, $x \cdot y \in F$ (<u>closure</u>)
- $x+y = y+x$, $xy = yx$ for $x, y \in F$ (<u>commutativity</u>)
- $(x+y)+z = x+(y+z)$, $(xy)z = x(yz)$ for $x, y, z \in F$ (<u>associativity</u>)
- exists $0 \neq 1 \in F$ s.t. $0+x = x$, $1 \cdot x = x$ for $x \in F$ (<u>identities</u>)
- $\begin{cases} \text{if } x \in F \text{ then there exists } -x \in F \text{ s.t. } x+(-x) = 0 \\ \text{if } x \in F \text{ and } x \neq 0 \text{ then there exists } \frac{1}{x} \in F \text{ s.t. } x \cdot \frac{1}{x} = 1 \end{cases}$ (<u>inverses</u>)
- $(x+y)z = xz + yz$ for $x, y, z \in F$ (<u>distributivity</u>)

Remark: Common to write
$$\begin{cases} x-y \text{ for } x+(-y) \\ x^{-1} \text{ for } 1/x \\ x/y \text{ for } xy^{-1} \\ x^n = \underbrace{x\cdot\ldots\cdot x}_{(n\text{-times})} \\ nx = \overbrace{x+\ldots+x} \end{cases}$$

Examples:
- $\mathbb{Q}$ is a field with usual $+,\cdot$.
- $\{0,1\}$ with $1+1=0$ is a field (<u>trivial</u>)
  (indeed $\{0,\ldots,p-1\}$ where p is prime is a field under $+,\cdot$ modulo p, $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$)
- $\mathbb{Z}$ is <u>not</u> a field with usual $+,\cdot$ ($\frac{1}{2} \notin \mathbb{Z}$!).
  (this is a ring however)

Any general fact proven for a field must then hold in $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ etc., this is why abstraction can be useful!
(For instance Propositions 1.14/5/6/8 all follow easily from definitions).
(we will assume these from now on!)

We also have ordered fields:

Definition: An <u>ordered field</u> is a field, F, which is also an ordered set such that

① $x+y < x+z$ if $x,y,z \in F$ and $y<z$.

② $xy > 0$ if $x,y \in F$ and both $x,y>0$.

We call $x \in F$ <u>positive</u> if $x>0$, <u>negative</u> if $x<0$.

Example: $\mathbb{Q}$ is an ordered field, $\mathbb{Z}_p$ is not (nor is $\mathbb{C}$!).

Inequalities in ordered fields behave as expected.
(See Proposition 1.18).

The main reason we discuss fields is the following:

Theorem: There exists an ordered field, $\mathbb{R}$, which has the least upper bound property. Moreover, $\mathbb{Q} \subset \mathbb{R}$.

Definition: We call elements of $\mathbb{R}$ real numbers.

Proof: We will do it at the end of term if time allows; it is long and not too insightful, involving an explicit construction from $\mathbb{Q}$ itself. $\square$

The fact that $\mathbb{R}$ has the least upper bound property allows us to prove some useful, and familiar, facts:

Theorems: ① (Archimedean Property) If $x, y \in \mathbb{R}$ and $x > 0$, then there exists $n \in \mathbb{N}$ such that $nx > y$.

② (Density of $\mathbb{Q}$) If $x, y \in \mathbb{R}$, then there is a $p \in \mathbb{Q}$ such that $x < p < y$.

③ (Existence of roots) If $x > 0$ and $n \in \mathbb{N}$, there is a unique $y \in \mathbb{R}$ with $y > 0$ such that $y^n = x$. We call y the n-th root of x and write $y = \sqrt[n]{x} = x^{\frac{1}{n}}$.

Proof: ① Let $A = \{nx \mid n \in \mathbb{N}\}$ and assume that $nx \leq y$ for all $n \in \mathbb{N}$; thus y is an upper bound for A. Since $\mathbb{R}$ has the least upper bound property, $\sup A$ exists. Now, as $x > 0$ we have

$$\sup A - x < \sup A,$$

thus $\sup A - x$ is not an upper bound for A (by def$^{\underline{n}}$ of $\sup A$), so there must exist some $n \in \mathbb{N}$ with

$$\sup A - x < nx$$

but then $\sup A < (n+1)x$ which is a contradiction since $n+1 \in \mathbb{N}$ and $\sup A$ is an upper bound for A; thus $nx > y$ for some $n \in \mathbb{N}$. $\square$

(assume $x, y > 0$, if $x < 0 < y$ done, $x, y < 0$ same)

② Since $x < y$ we have $y - x > 0$, thus by applying the Archimedean property (① with $x \leftrightarrow y-x$, $y \leftrightarrow 1$) there is some $n \in \mathbb{N}$ such that $n(y-x) > 1$. Again by the Archimedean property there is some $k \in \mathbb{N}$ with $k > nx$ and hence $E = \{k \in \mathbb{N} \mid k > nx\} \neq \emptyset$. By the well ordering principle, E has a smallest element, let us call this smallest element $m \in E$; then

$$m - 1 \leq nx < m \quad (\text{Since } m-1 \notin E!).$$

Combining $\begin{cases} n(y-x) > 1 \iff nx < ny - 1, \\ m - 1 \leq nx < m \end{cases}$

we have $nx < m \leq nx + 1 < ny \implies x < \frac{m}{n} < y$, and so setting $p = \frac{m}{n} \in \mathbb{Q}$ concludes the proof. ▨

③ We first show uniqueness (assuming existence); if $y, \tilde{y} \in \mathbb{R}$ were both such that $y^n = \tilde{y}^n = x$ but $y \neq \tilde{y}$, then since either $y < \tilde{y}$ or $y > \tilde{y}$ we must have $y^n < \tilde{y}^n$ or $y^n > \tilde{y}^n$, a contradiction.
(since $y, \tilde{y} > 0$)

For existence we let $E = \{t > 0 \mid t^n < x\}$; this is non-empty since if $t = \frac{x}{x+1}$ then $0 < t < 1$ and so $t^n < t < x$. We also see that $t = 1 + x$ is an upper bound for E since then $t > 1$ and so $t^n > t > x$. Since $\mathbb{R}$ has the least upper bound property, $y = \sup E$ exists. We will prove that $y^n = x$ by showing that $y^n > x$ or $y^n < x$ give a contradiction. Note that $y > 0$ in the following!

<u>If $y^n < x$:</u> Whenever $0 < a < b$ we note that

$$(+) \qquad b^n - a^n = (b-a)(b^{n-1} + b^{n-2}a + \dots + a^{n-1}) \rightarrow \left(\begin{array}{c}\text{can prove by}\\ \text{induction}\\ (n \geq 2)\end{array}\right)$$
$$< (b-a) \cdot n \cdot b^{n-1}.$$

Choosing $0 < h < 1$ so that $h < \frac{x - y^n}{n(y+1)^{n+1}}$ we see that →

P.T.O

Setting $a = y$ and $b = y+h$ we have by (+) that

$$(y+h)^n - y^n < h \cdot n \cdot (y+h)^{n-1} \underset{(h<1)}{<} h \cdot n \cdot (y+1)^{n-1} < x - y^n$$

and so we see that

$$(y+h)^n < x \Rightarrow y+h \in E,$$

but $y = \sup E$ contradicting y being an upper bound for E!

<u>If $y^n > x$</u>: we set $k = \dfrac{y^n - x}{n y^{n-1}}$ and note that whenever $t \geq y-k$ we have that by (+) with $a = y-k$, $b = y$ $\rightarrow (k<y!)$

$$y^n - t^n \leq y^n - (y-k)^n < k \cdot n \cdot y^{n-1} = y^n - x,$$

and so $x < t^n \Rightarrow t \notin E$, and so $y-k$ is an upper bound for E. However, $y-k < y$ which contradicts the fact that $y = \sup E$, the <u>least</u> upper bound for E.

Hence, both $y^n > x$, $y^n < x$ give a contradiction and thus

$y^n = x$ as desired. ∎

$\boxed{\text{End L2}}$

//

<u>Extended real numbers, $\mathbb{C}$, and $\mathbb{R}^n$</u>

In order to ensure every subset of $\mathbb{R}$ has an inf/sup we sometimes make use of:

<u>Definition</u>: The <u>extended real numbers</u> consists of $\mathbb{R}$ and two symbols, $+\infty$ and $-\infty$, called <u>positive and negative infinity</u> respectively. The extended real numbers forms an ordered set by preserving the order on $\mathbb{R}$ and such that

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

<u>Remark</u>: Every $E \subset \mathbb{R}$ is then bounded above/below by $\pm\infty$ respectively. Thus, if E is not bounded above in $\mathbb{R} \Rightarrow \sup E = +\infty$.

We also make the conventions that if $x \in \mathbb{R}$ then:

$$x \pm \infty = \pm\infty, \quad \frac{x}{\pm\infty} = 0, \quad \begin{cases} x > 0 \Rightarrow x \cdot (\pm\infty) = \pm\infty, \\ x < 0 \Rightarrow x \cdot (\pm\infty) = \mp\infty. \end{cases}$$

Even with these, the extended real numbers are <u>not</u> a field! ($\pm\infty$ have no additive or multiplicative inverses.)

We also introduce the complex numbers, which help solve the equation $x^2 = -1$ (just as $\mathbb{R}$ helped for $x^2 = 2$):

<u>Definition</u>: The <u>complex numbers</u>, $\mathbb{C}$, is the set of ordered pairs, (a, b), where $a, b \in \mathbb{R}$. For $x = (a, b), y = (c, d) \in \mathbb{C}$ we write $x = y \Leftrightarrow a = c$ and $b = d$, and define

$$x + y = (a + c, b + d), \quad x \cdot y = (ac - bd, ad + bc).$$

<u>Remarks</u>:
- By an <u>ordered pair</u> we mean that $(a, b) \neq (b, a)$ if $b \neq a$. We view $\mathbb{R} \subset \mathbb{C}$ by identifying $a \in \mathbb{R}$ with $(a, 0)$.
- One can show (Theorem 1.25) that the operations $+, \cdot$ turn $\mathbb{C}$ into a field with $(0,0), (1,0)$ the additive and multiplicative identity respectively.
- By defining $i = (0, 1)$ we see that $i^2 = (-1, 0)$ from which we can identify $(a, b) \in \mathbb{C}$ with $a + bi$. We will <u>only</u> use this latter form from now on.

<u>Definitions</u>:
- For $z = a + bi \in \mathbb{C}$ we call $\bar{z} = a - bi \in \mathbb{C}$ the <u>conjugate</u> of z.
- For $z = a + bi \in \mathbb{C}$ we call $a = Re(z)$ the <u>real part of z</u> and $b = Im(z)$ the <u>imaginary part of z</u>.
- If $x_1, \ldots, x_n \in \mathbb{C}$ we write $\sum_{i=1}^{n} x_i = x_1 + \ldots + x_n$.

- If $z \in \mathbb{C}$ then the ~~modulus~~ absolute value of z is $|z| = (z\bar{z})^{\frac{1}{2}}$
(One can show that $z\bar{z} \in \mathbb{R}$ and ≥ 0 so well defined)
When $x \in \mathbb{R}$, $\bar{x} = x$ so $|x| = (x^2)^{\frac{1}{2}}$ and so

$$|x| = \begin{cases} x, & x \geq 0 \\ -x, & x < 0 \end{cases}$$

//

Theorems: Let $z, w \in \mathbb{C}$. Then:

1. $\overline{z+w} = \bar{z} + \bar{w}$, $\overline{zw} = \bar{z}\bar{w}$.
2. $z + \bar{z} = 2\mathrm{Re}(z)$, $z - \bar{z} = 2i\,\mathrm{Im}(z)$.
3. $z\bar{z} \in \mathbb{R}$, $z\bar{z} \geq 0$ and $z\bar{z} = 0 \iff z = 0$.
4. $|z| \geq 0$ and $|z| = 0 \iff z = 0$
5. $|\bar{z}| = |z|$, $|zw| = |z||w|$, $|\mathrm{Re}(z)| \leq |z|$
6. $|z+w| \leq |z| + |w|$ (triangle inequality)

Proof: • 1 → 5 are easier to verify from definitions.

- For 6 we notice that as $\overline{(z\bar{w})} = \bar{z}w$ we have

$$\begin{aligned}
|z+w|^2 &= (z+w)(\bar{z}+\bar{w}) = z\bar{z} + z\bar{w} + \bar{z}w + w\bar{w} \\
&\overset{①}{=} |z|^2 + 2\mathrm{Re}(z\bar{w}) + |w|^2 \\
&\overset{②}{\leq} |z|^2 + 2|z\bar{w}| + |w|^2 \underset{⑤}{=} |z|^2 + 2|z||w| + |w|^2 \\
&\overset{⑤}{=} (|z| + |w|)^2,
\end{aligned}$$

taking square roots of both sides we deduce that

$$|z+w| \leq |z| + |w|$$ ☒

Remark: The triangle inequality is extremely useful; be sure to remember it!

Another useful inequality is:

<u>Theorem:</u> Let $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{C}$. Then
$$\left| \sum_{i=1}^{n} a_i \overline{b_i} \right|^2 \leq \sum_{i=1}^{n} |a_i|^2 \cdot \sum_{i=1}^{n} |b_i|^2,$$
this is called the <u>Cauchy-Schwartz inequality</u>.

<u>Proof:</u> Let $A = \sum_{i=1}^{n} |a_i|^2$, $B = \sum_{i=1}^{n} |b_i|^2$, and $C = \sum_{i=1}^{n} a_i \overline{b_i}$; we want to show that $|C|^2 \leq A \cdot B$. Note first that if $B = 0 \Rightarrow b_1 = \dots = b_n = 0$ and hence the inequality holds immediately. Thus we assume that $B > 0$ and see that
$$0 \leq \sum_{i=1}^{n} |Ba_i - Cb_i|^2 = \sum_{i=1}^{n} (Ba_i - Cb_i)(B\overline{a_i} - \overline{C}\overline{b_i})$$
$$= \sum_{i=1}^{n} B^2|a_i|^2 - B\overline{C} a_i \overline{b_i} - BC\overline{a_i} b_i + |C|^2 |b_i|^2$$
$$= B^2 A - B|C|^2 = B(AB - |C|^2),$$
now as $B > 0$ we see that
$$AB - |C|^2 \geq 0 \iff |C|^2 \leq AB$$
$\blacksquare$

We also introduce high dimensional spaces:

<u>Definition:</u> For each $n \in \mathbb{N}$ the <u>n-dimensional Euclidean Space</u>, $\mathbb{R}^n$, is the set of all ordered n-tuples, called <u>Points/Vectors</u>,
$$x = (x_1, \dots, x_n),$$
where $x_1, \dots, x_n \in \mathbb{R}$ are called the <u>coordinates</u> of x.
For $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ we define
$$x + y = (x_1 + y_1, \dots, x_n + y_n), \quad x \cdot y = \sum_{i=1}^{n} x_i y_i \in \mathbb{R}, \text{ (inner product)}$$
$$\alpha x = (\alpha x_1, \dots, \alpha x_n), \quad |x| = \left(\sum_{i=1}^{n} (x_i)^2 \right)^{\frac{1}{2}} = (x \cdot x)^{\frac{1}{2}} \text{ (norm)}$$
(Scalar multiplication)

Remarks: • The addition of vectors and scalar multiplication give $\mathbb{R}^n$ the structure of a vector space over the field $\mathbb{R}$ (see linear Algebra)

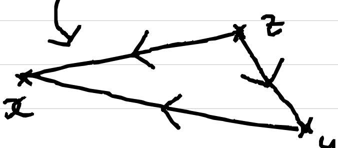
• The origin of $\mathbb{R}^n$ is the vector $0 = (0, \dots, 0)$.

• $\mathbb{R}^0$ is a point (0-tuple), $\mathbb{R}^1 = \mathbb{R}$ is the real line, $\mathbb{R}^2$ is called the plane ($\mathbb{R}^2 = \mathbb{C}$ with norm $\leftrightarrow$ absolute value).

We have the following properties:

Theorem: Let $x, y, z \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$. Then:

1. $|x| \geq 0$ and $|x| = 0 \iff x = 0$.
2. $|\alpha x| = |\alpha||x|$.
3. $|x \cdot y| \leq |x||y|$. (Cauchy-Schwarz inequality)
4. $|x+y| \leq |x| + |y|$ (triangle inequality)
5. $|x-z| \leq |x-y| + |y-z|$ (addition of vectors)



Proof: • 1 + 2 follow from definitions.
• 3 follows from C-S before!
• 4 is the same as for Δ-inequality:

$$|x+y|^2 = (x+y)\cdot(x+y) = x.x + 2x.y + y.y$$
$$\underset{③}{\leq} |x|^2 + 2|x||y| + |y|^2 = (|x|+|y|)^2.$$

(then take square roots).

• 5 follows from 4 setting $x \leftrightarrow x-y$ and $y \leftrightarrow y-z$. ☒

End L13

We now formally introduce functions, before discussing sets with a notion of distance, in which we will be able to define limits and other properties which hold in the usual sense in $\mathbb{R}, \mathbb{C}, \mathbb{R}^n$.

Functions and Cardinality

(or just simply f)

Definition: A function, $f : A \to B$, between sets A and B is an assignment $f(x) \in B$ for each $x \in A$. We call A the domain, B the codomain, $\{f(x) \mid x \in A\}$ the range of f, whose elements are called the values of f.

Examples:
- $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(n) = n+1$ for $n \in \mathbb{Z}$
- $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x^2$ for $x \in \mathbb{R}$.

We commonly use the following notations:

- If $E \subset A$ and $f : A \to B$, $f(E) = \{f(x) \mid x \in E\}$ is the image of E under f; thus $f(A) =$ range of f.
- If $f(A) = B$ we say that f is surjective/onto.
- If $E \subset B$ and $f : A \to B$, $f^{-1}(E) = \{x \in A \mid f(x) \in E\}$ is the pre-image of E under f; thus if $y \in B$ then the pre-image of y is $f^{-1}(y) = \{x \in A \mid f(x) = y\}$.
- If for every $y \in B$ there is at most one element of $f^{-1}(y)$ we say that f is injective/1-1. Equivalently, f is injective if $f(x) \neq f(\tilde{x})$ whenever $x \neq \tilde{x}$ for $x, \tilde{x} \in A$.
- We say that $f : A \to B$ is bijective if it is surjective and injective. One can then define the inverse of f, $f^{-1} : B \to A$, by setting $f^{-1}(y) = x \iff f(x) = y$.

We can now speak about the 'size' of sets:

Definition: If there is a bijection between sets A and B then we say that A and B have the same cardinality and write $A \sim B$.

We use cardinality to make sense of infinity:

Definition: For $n \in \mathbb{N}$ let $[n] = \{1, 2, \cdots, n\}$ ($[0] = \emptyset$). We say that a set A is:

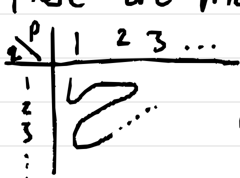
- Finite if $A \sim [n]$ for some $n \in \mathbb{N}$ ($\Leftrightarrow$ same # elements)
- Infinite if A is not finite.
- Countable if $A \sim \mathbb{N}$.
- At most countable if A is finite or countable.
- Uncountable if A is neither finite or countable.

We can now examine the size of familiar sets:

Examples:

- $\emptyset, \{0, 1\}, \{a, b, c, \dots, z\}$ are finite.
- $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \{\text{primes}\}$ are infinite.
- $\mathbb{Z}, \mathbb{Q}$ are countable. For $\mathbb{Z}$ the map $f : \mathbb{N} \to \mathbb{Z}$ defined by
$$f(n) = \begin{cases} n/2 , & n\text{-even} \\ -\frac{n-1}{2} , & n\text{-odd} \end{cases}$$
, is a bijection (check!)

There are many ways to show that $\mathbb{Q}$ is countable,



OR $\pm \frac{p}{q} \mapsto 2^p 3^q 5^{\pm 1 + 1}$ OR $\frac{p}{q} \mapsto (p, q) \in \mathbb{Z}^2$.

- We will show that $\mathbb{R}$ (and hence $\mathbb{C}, \mathbb{R}^n$) are uncountable, but this will require some more preliminaries.

Remark: Often I write $n \geq 1$ for elements of $\{n \in \mathbb{N} \mid n \neq 0\}$.

Definition: A <u>sequence</u> is a function defined on $\mathbb{N}\setminus\{0\}$. ($\hookrightarrow \{n \geq 1\}$)
If f is a sequence and $f(n) = x_n$ for each $n \geq 1$, then we write f as $(x_n)_{n\geq 1}$, $\{x_n\}_{n \geq 1}$ or just (x_n)/$x_1, x_2, \ldots$
If A is a set and $x_n \in A$ for all $n \geq 1$ then (x_n) is said to be a <u>sequence in A</u>.

We can use sequences to prove:

(So $\{\text{primes}\} \subset \mathbb{N}$ is!)

Theorem: Every subset of a countable set is at most countable.

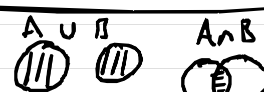
Proof: Let A be countable and $E \subset A$, we only need to show that E is countable if it is infinite. Since A is countable there is a sequence $(x_n) = A$ such that $x_n = x_m \Leftrightarrow n = m$ (i.e. distinct).

Consider $\{n \mid x_n \in E\} \subset \mathbb{N}$, by the well ordering principle there is some $n_1 \geq 1$ which is the smallest element s.t. $x_{n_1} \in E$. We then consider $\{n \mid x_n \in E\} \setminus \{x_{n_1}\} \subset \mathbb{N}$ and again choose its smallest element, $n_2 \geq 1$, such that $x_{n_2} \in E \setminus \{x_{n_1}\}$. Continuing inductively we choose $x_{n_k} \in \mathbb{N} \setminus \{x_{n_1}, \ldots, x_{n_{k-1}}\}$ by the same process.

We then define a bijection $f : \mathbb{N}\setminus\{0\} \to E$ by $f(k) = x_{n_k}$; which is injective as $x_{n_k} = x_{n_\ell} \Leftrightarrow n_k = n_\ell \Leftrightarrow k = \ell$. To see that f is surjective we note that if $y \in E$ then as $E \subset A$ there is some $m \geq 1$ such that $y = x_m$; the above procedure to define the x_{n_k} will select x_m after at most m steps. $\blacksquare$

//

We now formally introduce set operations:

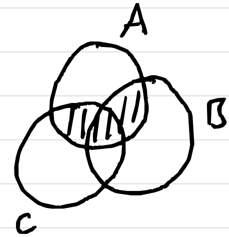


Definition: Let A (index set) and Ω (universe) be sets such that for each $\alpha \in A$ there is some $E_\alpha \subset \Omega$, then:

- The <u>union</u> of the E_α, $\bigcup_{\alpha \in A} E_\alpha$, is such that $x \in \bigcup_{\alpha \in A} E_\alpha \Leftrightarrow x \in E_\alpha$ for <u>some</u> $\alpha \in A$.
- The <u>intersection</u> of the E_α, $\bigcap_{\alpha \in A} E_\alpha$, is such that $x \in \bigcap_{\alpha \in A} E_\alpha \Leftrightarrow x \in E_\alpha$ for <u>all</u> $\alpha \in A$.

We have the following operations:

- $A \cup B = B \cup A$, $A \cap B = B \cap A$.
- $(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$.
- $A \subset A \cup B$, $A \cap B \subset A, B$.
- $A \cup \phi = A$, $A \cap \phi = \phi$.
- If $A \subset B$ then $A \cup B = B$, $A \cap B = A$.
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.



<u>Definition</u>: If $A \cap B = \phi$ we say that A and B are <u>disjoint</u>, and if $A \cap B \neq \phi$ we say that they <u>intersect</u>.

<u>Theorem</u>: A countable union of countable sets is countable. (by <u>countable union</u> we mean $\bigcup_{n \geq 1} E_n$ for a sequence of sets (E_n)).

<u>Proof</u>: Let $E_n = (x_k^n)$ for $n, k \geq 1$ be countable sets written as sequences. We can form a sequence

$$(a_n) = x_1^1, x_1^2, x_2^1, x_1^3, x_2^2, x_3^1, x_1^4, x_2^3, x_3^2, x_4^1, x_1^5, \dots$$

(indices) $\longrightarrow$ (<u>sum to 3</u>) (<u>sum to 4</u>) (<u>sum to 5</u>) (etc.)

which contains all elements of $\bigcup_{n \geq 1} E_n$ (note that some element will appear more than once unless all the E_n are disjoint). We can then construct a further sequence (hence a bijection to $\mathbb{N}$) by successively removing repeated terms, hence $\bigcup_{n \geq 1} E_n \sim \mathbb{N}$ $\boxtimes$

<u>Theorem</u>: Let A be a countable set and A^n the set of n-tuples, $(a_1, \dots, a_n)$, where $a_i \in A$ for each $i = 1, \dots, n$. Then A^n is countable. ($n \geq 1$).

<u>Proof</u>: As $A^1 = A$, A^1 is countable. Elements of A^n can be written as (b, a) where $b \in A^{n-1}$ and $a \in A$; for each fixed $b \in A^{n-1}$ the set $E_b = \{(b, a) \mid a \in A\}$ is thus countable. By the previous theorem, as $A^n = \bigcup_{b \in A^{n-1}} E_b$, we see that A^n is countable by induction. $\boxtimes$

Examples: • $\mathbb{N}^n, \mathbb{Z}^n$ are thus countable for each $n \geq 1$.

• Each $\frac{p}{q} \in \mathbb{Q}$ can be written as $(p,q) \in \mathbb{Z}^2$ and hence $\mathbb{Q}$ is countable.

• One can show that the <u>algebraic numbers</u> (solutions to polynomial equations with $\mathbb{Q}$-coefficients) are also countable using this result. $\boxed{\text{End L4}}$

So far we have not seen any uncountable sets:

<u>Theorem:</u> Let $A = \{(x_n) \mid x_n \in \{0,1\}\}$, then A is uncountable.
↳ (set of <u>binary</u> sequences)

<u>Proof:</u> Let $E \subset A$ be a countable subset and enumerate its elements (sequences) as (S_i). We will construct, from these sequences, a further sequence $s \in A \setminus E$; this shows that any countable subset of A is proper $(\neq A)$ and thus A cannot be countable (as $A = A$).

To construct s we choose its kth term to be $0/1$ if the first term of S_k is $1/0$ (flip the term); thus $s \neq S_k$ for any $k \geq 1$ as the kth term of s differs from the kth term of S_k! Hence $s \in A \setminus E$, and so A is uncountable. $\blacksquare$

$$E \begin{cases} S_1 = \overset{\nearrow 1}{0}, 0, 0, 0, \ldots \\ S_2 = 1, \overset{\nearrow 1}{0}, 0, 0, \ldots \\ S_3 = 0, 1, \overset{\nearrow 1}{0}, 0, \ldots \\ S_4 = 1, 1, 0, \overset{\nearrow 1}{0}, \ldots \\ \vdots \quad \vdots \quad \vdots \quad \vdots \end{cases} \Rightarrow s = 1, 1, 1, 1, \ldots \in A \setminus E.$$

(often called <u>Cantor's diagonalisation argument</u>)

<u>Remark:</u> This shows that $\mathbb{R}$ is uncountable using binary expansions, but we will see a different proof later.

Metric Spaces

We now study sets with a notion of distance:

Definition: A metric space is a set X and a function $d : X \times X \to [0,\infty)$ such that:

1. $d(x,y) \geq 0$ and $d(x,y) = 0 \iff x = y$.
2. $d(x,y) = d(y,x)$
3. $d(x,z) \leq d(x,y) + d(y,z)$ for all $x,y,z \in X$ (Δ ineq.),

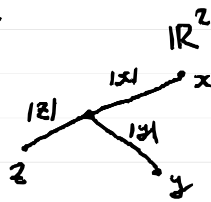
We call $d(x,y)$ the distance between x and y.

Examples: • $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{R}^n$ (hence $\mathbb{C}$) are all metric spaces with the usual distance $d(x,y) = |x-y|$.

• On $\mathbb{R}^2$ we can consider the distance

(different to $(\mathbb{R}^2, |\cdot|)$)

$$d(x,y) = \begin{cases} |x-y| , & x = \lambda y \text{ for } y \in \mathbb{R} \\ |x|+|y| , & x \neq \lambda y \text{ for } y \in \mathbb{R} \end{cases}$$



which is the 'Paris metro' metric.

Remarks: • We may write (X,d) to specify the metric.

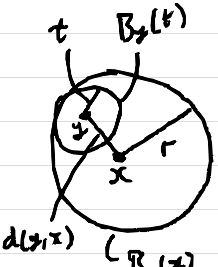
• We write intervals $(a,b) = \{x \in \mathbb{R} \mid a < x < b\}$ (for $a,b \in \mathbb{R} \cup \{\pm\infty\}$) $[a,b) = \{x \in \mathbb{R} \mid a \leq x < b\}$ etc.

• Subsets of metric spaces are metric spaces by restricting the metric (often called induced metric).

Definition: Let (X,d) be a metric space and $x \in X$, an open ball (or simply ball) around x, of radius $r > 0$, is the set $B_r(x) = \{y \in X \mid d(x,y) < r\}$ (sometimes a ball is called a neighbourhood of x). Given $E \subset X$, we say that $x \in E$ is an interior point of E if there is some $r > 0$ with $B_r(x) \subset E$, we say that E is open in X if every point of E is an interior point.

<u>Remarks:</u> • The set of interior points of $E \subset X$ is often denoted by E°; thus E is open in $X \iff E = E^{\circ}$.
<u>/Examples:</u>

• $\emptyset \subset X$ and X itself are open; for the former there are no points to check! Thus any metric space has a <u>topology</u> induced by its metric.



• Open balls are called such as they are open; to see this, let $B_r(x) \subset X$ for some $x \in X, r > 0$ be a ball. For $y \in B_r(x)$ we consider
$$t = r - d(x,y) > 0 \leftarrow (\text{since } y \in B_r(x))$$
and note that $B_t(y) \subset B_r(x)$ since if $z \in B_t(y)$ then
$$d(z,y) < t \Rightarrow d(z,x) \leq d(z,y) + d(y,z)$$
$$< t + d(y,z)$$
$$= r - d(x,z) + d(y,z) = r.$$
Thus, $z \in B_r(x)$ also so $B_t(y) \subset B_r(x)$ and so any ball is indeed open. ▨

//

We can also use open balls to make sense of limiting points:

<u>Definition:</u> Let (X,d) be a metric space and $E \subset X$. We say that $x \in X$ is a <u>limit point</u> of E if every neighbourhood of x contains a point $y \in E$ distinct from x (i.e. $y \neq x$); this is equivalent to saying $E \cap (B_r(x) \setminus \{x\}) \neq \emptyset$ for any $r > 0$. We say that $E \subset X$ is <u>closed</u> in X if E contains all of its limit points; points $p \in E$ that are not limit points of E are called <u>isolated</u>.

<u>Remarks:</u> • The set of limit points of $E \subset X$ is often denoted by E'; thus E is closed in $X \iff E' \subset E$.

• Limit points do <u>not</u> have to lie in the set; for example if $E = (a,b) \subset \mathbb{R} \Rightarrow E' = [a,b]$ so E is <u>not</u> closed in $\mathbb{R}$. (nor is $[a,b) \subset \mathbb{R}$, but $[a,b] \subset \mathbb{R}$ is!) (nor is $\{\frac{1}{n} \mid n \geq 1\} \subset \mathbb{R}$) (0 is a limit point).

We have the following (perhaps intuitive) result:

Theorem: Let (X,d) be a metric space and $E \subset X$. If $x \in E'$, then every neighbourhood of x contains infinitely many points of E.

Proof: Suppose there was some $r>0$ such that $B_r(x)$ has only finitely many points of E; ie. $B_r(x) \cap E = \{y_1, \dots, y_n\}$ (note that this cannot be empty as $x \in E'$). We set

$$\hat{r} = \min_{i=1,\dots,n} \{d(y_i, x)\} > 0 \quad (\text{as } y_i \neq x)$$

which is such that $\hat{r} < r$ (as $y_i \in B_r(x)$) and hence $B_{\hat{r}}(x) \subset B_r(x)$ and so $B_{\hat{r}}(x) \cap E \subset \{y_1, \dots, y_n\}$. However, by the choice of $\hat{r}$, $y_i \notin B_{\hat{r}}(x)$ for $i=1,\dots,n$ and so $B_{\hat{r}}(x) \cap E = \emptyset$; this contradicts $x \in E'$ ▨

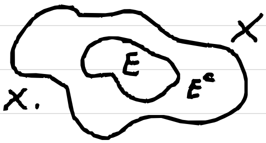
Remark: Any finite subset of a metric space has no limit points!

End L5

One useful notion for sets is:

(this notion holds for sets more generally of course)

Definition: Let (X,d) be a metric space and $E \subset X$, the <u>complement</u> of E in X is $E^c = \{x \in X \mid x \notin E\} = X \setminus E$.



Remarks:
- $X = E \cup E^c$, $(E^c)^c = E$, $X^c = \emptyset$, $\emptyset^c = X$.
- Complements depend on the 'ambient' metric space; for example if $[-1,1] \subset \mathbb{R}$ then $[-1,1]^c = (-\infty,-1) \cup (1,\infty) \subset \mathbb{R}$, but if $[-1,1] \subset \{x \in \mathbb{R} \mid x \geq -1\}$ then $[-1,1]^c = (1,\infty)$.

Theorem: Let (X,d) be a metric space and $\{E_\alpha\}_\alpha$ be a collection of subsets of X, then

$$\left(\bigcup_\alpha E_\alpha\right)^c = \bigcap_\alpha (E_\alpha^c).$$

Proof: If $x \in \left(\bigcup_\alpha E_\alpha\right)^c \iff x \notin \bigcup_\alpha E_\alpha \iff x \notin E_\alpha$ for every $\alpha \iff x \in \bigcap_\alpha (E_\alpha)^c$. ▨

We have the following characterisation of open sets:

Theorem: Let (X,d) be a metric space and $E \subset X$, then E is open $\iff E^c$ is closed.

Proof: $(\Rightarrow)$ If E is open and $x \in X$ is a limit point of E^c, then every neighbourhood of x contains a point of E^c. Thus, x cannot be an interior point of E, but since E is open $\Rightarrow x \notin E \iff x \in E^c$. Hence $(E^c)' \subset E^c$ and so E^c is closed.

$(\Leftarrow)$ If E^c is closed then $(E^c)' \subset E^c$, hence if $x \in E$ then x cannot be a limit point of E^c. Thus there is some neighbourhood of x disjoint from E^c (namely there is $r > 0$ such that $B_r(x) \cap E^c = \emptyset$) and hence x is an interior point of E (as $B_r(x) \subset E$). The choice of $x \in E$ was arbitrary and hence $E = E^\circ$ and so E is open. ▨

Remark: Since $(E^c)^c = E$, E is closed $\iff E^c$ is open.

We now see how open/closed sets behave under union/intersection:

Theorem: ① Any union of open sets is open.

② Any intersection of closed sets is closed

③ Finite intersections of open sets are open.

④ Finite unions of closed sets are closed.

Proof: Note first that, since $\left(\bigcup_\alpha E_\alpha\right)^c = \bigcap_\alpha (E_\alpha^c)$ and

E_α is open $\iff E_\alpha^c$ is closed by the previous two theorems, ① $\iff$ ② and ③ $\iff$ ④; we will just prove ① and ③.

For ①, we notice that if $\{E_\alpha\}_\alpha$ are open sets then as $x \in \bigcup_\alpha E_\alpha \Rightarrow x \in E_\alpha$ for some α and E_α is open,

there must exist $r>0$ such that $B_r(x) \subset E_\alpha \subset \bigcup_\alpha E_\alpha$, hence $\bigcup_\alpha E_\alpha$ is open.

For ③, we note that if $\{E_i\}_{i=1}^n$ are open sets then $x \in \bigcap_{i=1}^n E_i \Rightarrow x \in E_i$ for $i=1,\dots,n$. As each E_i is open there exists, for each $i=1,\dots,n$, some $r_i>0$ such that $B_{r_i}(x) \subset E_i$. Let $r = \min_{i=1,\dots,n}\{r_i\} > 0$ and note then that

$B_r(x) \subset B_{r_i}(x) \subset E_i$ for each $i=1,\dots,n \Rightarrow B_r(x) \in \bigcap_{i=1}^n E_i$, hence $\bigcap_{i=1}^n E_i$ is open. ▨

Remark: The requirement that the collections in ③ and ④ are finite is necessary! For example;

$$\bigcap_{n\geq 1} \underbrace{\left(-\tfrac{1}{n}, \tfrac{1}{n}\right)}_{\text{(open in } \mathbb{R})} = \{0\} \leftarrow (\underline{\text{not}} \text{ open in } \mathbb{R}!).$$

Another useful notion combines a set and its limit points:

Definition: Let (X,d) be a metric space and $E \subset X$, the <u>closure</u> of E in X is $\overline{E} = E \cup E'$.

Examples:
- $\overline{[a,b]} = \overline{(a,b)} = \overline{[a,b)} = \overline{(a,b]} \Rightarrow$ (<u>check!</u>) $\leftarrow \overline{\mathbb{Q}} = \mathbb{R}!$
- If $x \in X$ and $r>0$ then $\overline{B_r(x)} \subset \{y \in X \mid d(x,y) \leq r\}$.
  ↳ ($\supset$ fails in general; see HW 4).

Theorem: Let (X,d) be a metric space and $E \subset X$. Then:

1. $\overline{E}$ is closed.
2. $\overline{E} = E \iff E$ is closed.
3. If $F \subset X$ is closed and $E \subset F \Rightarrow \overline{E} \subset F$.

(So $\overline{E}$ is the 'smallest' closed set containing E)

Proof: For (1), if $x \in (\overline{E})^c$ then x is not a limit point of E and does not lie in E. Thus, there must be a neighbourhood of x contained in E^c, i.e. there is some $r>0$ such that $B_r(x) \subset E^c$. We will now show that $B_r(x) \subset (\overline{E})^c$ also, implying $\overline{E}$ is closed. As $B_r(x)$ is open, every point $y \in B_r(x)$ belongs to a neighbourhood disjoint from $E \Rightarrow y$ is not a limit point of E; thus, $B_r(x) \subset (E')^c$. We thus see that

$$B_r(x) \subset E^c \cap (E')^c = (E \cup E')^c = (\overline{E})^c.$$

For (2), $\overline{E} = E \iff E = E \cup E' \iff E' \subset E \iff E$ is closed.

For (3), F is closed $\iff F' \subset F$ and we note that if $x \in E'$ then as $E \subset F \Rightarrow$ for each $r>0$,

$$\emptyset \neq (B_r(x) \setminus \{x\}) \cap E \subset (B_r(x) \setminus \{x\}) \cap F \Rightarrow x \in F',$$

i.e. $E' \subset F'$. As $E \subset F$ and $E' \subset F$ (as $F' \subset F$) we have that $\overline{E} = E \cup E' \subset F$ as desired. ☒

Let's apply this to $\mathbb{R}$: (if we don't specify we always mean $(\mathbb{R}, |\cdot|)$)

Theorem: Let $E \subset \mathbb{R}$ be non-empty and bounded above, then $\sup E \in \overline{E}$. In particular if E is closed, then $\sup E \in E$.

Proof: If $\sup E \in E$ then there is nothing to show, so suppose $\sup E \notin E$. For each $\varepsilon > 0$ there must exist $x \in E$ such that

$$\sup E - \varepsilon < x < \sup E,$$

or else $\sup E - \varepsilon$ would be an upper bound for E. →

Such an $x \in (\sup E - \varepsilon, \sup E) \subset \underbrace{(\sup E - \varepsilon, \sup E + \varepsilon)}_{(B_\varepsilon(\sup E) \subset \mathbb{R})}$

and thus $B_\varepsilon(\sup E) \cap E \neq \emptyset$, as the choice of $\varepsilon > 0$ was arbitrary $\Rightarrow$ $\sup E$ is a limit point of E, hence $\sup E \in E' \subset \bar{E}$. ∎

$\boxed{\text{End L6}}$

Notice that when defining open sets, $E \subset X$, we specified that E was open in X. As any subset of a metric space, $Y \subset X$, is also a metric space one may wonder if there is a relation between being open in Y and being open in X.

Example: Viewing $\mathbb{R} = \mathbb{R} \times \{0\} \subset \mathbb{R}^2$ we see that $(a,b) \times \{0\}$ is open in $\mathbb{R} \times \{0\}$ but is not open in $\mathbb{R}^2$. Note however that $(a,b) \times \{0\} = \underbrace{((a,b) \times \mathbb{R})}_{\text{(open in } \mathbb{R}^2)} \cap (\mathbb{R} \times \{0\})$.

This motivates:

(just a careful way to specify where we are open)

Definition: Let (X,d) be a metric space and $E \subset Y \subset X$. We say that E is open relative to Y if E is open in Y; i.e. for each $x \in E$ there is some $r > 0$ such that

$$B_r(x) \cap Y \subset E.$$

(here $B_r(x)$ is the ball in X)

The remark in the above example holds generally:

Theorem: Let (X,d) be a metric space and $E \subset Y \subset X$, then E is open relative to Y $\iff$ $E = F \cap Y$ for some open $F \subset X$.

Proof: ($\Rightarrow$) If E is open relative to Y then for every $x \in E$ there is some $r_x > 0$ such that $B_{r_x}(x) \cap Y \subset E$. We then set

$$F = \bigcup_{x \in E} B_{r_x}(x)$$ — (open as a union of open sets)

$\longrightarrow$

($\Rightarrow$) and are done if $E = F \cap Y$. To see this we note that

$$x \in E \subset Y \text{ and } x \in B_{r_x}(x) \subset F \Rightarrow E \subset F \cap Y,$$

and if $y \in F \cap Y \Rightarrow y \in B_{r_x}(x)$ for some $x \in E$ and $y \in Y$

$$\Leftrightarrow y \in B_{r_x}(x) \cap Y \subset E \quad \text{(as } E \text{ is open relative to } Y\text{)}$$

$$\Rightarrow F \cap Y \subset E;$$

hence $E = F \cap Y$.

($\Leftarrow$) If $E = F \cap Y$ for $F \subset X$ open, if $x \in E = F \cap Y$ then there is some $r_x > 0$ such that $B_{r_x}(x) \subset F$ and hence as $x \in Y$ also, we have $B_{r_x}(x) \cap Y \subset F \cap Y = E$ and so E is open relative to Y. ▨

Compactness

We saw that 'openness' is <u>not</u> inherited by supersets; the following notion of 'smallness' for subsets is however:

<u>Definition:</u> Let (X, d) be a metric space and $E \subset X$, we say that a collection of open sets $\{F_\alpha\}_\alpha$ in X is an <u>open cover</u> of E if

$$E \subset \bigcup_\alpha F_\alpha .$$

A set $K \subset X$ is said to be <u>compact</u> in X if every open cover of K has a finite <u>subcover</u>; namely if $\{F_\alpha\}_\alpha$ is an open cover of F then there exist some $\{F_{\alpha_1}, \dots, F_{\alpha_n}\} \subset \{F_\alpha\}_\alpha$ which is an open cover of K, ie.

$$K \subset \bigcup_{i=1}^{n} F_{\alpha_i} .$$

<u>Remarks:</u>
- Every finite subset of a metric space is compact.
- This concept is of central importance in analysis, we will work with it a <u>lot</u>. It also makes sense topologically.

Theorem: Let (X,d) be a metric space and $K \subset Y \subset X$, then K is compact in $Y \iff K$ is compact in X.

Proof: $(\Rightarrow)$ If K is compact relative to Y and $\{F_\alpha\}_\alpha$ are open sets in X which cover K, then by the previous theorem the sets $F_\alpha \cap Y$ are open for each α and as $K \subset Y$,

$$K \subset \bigcup_\alpha (F_\alpha \cap Y)$$

and so $\{F_\alpha \cap Y\}_\alpha$ is an cover of K by open sets in Y. Thus there are $\{F_{\alpha_1} \cap Y, \dots, F_{\alpha_n} \cap Y\} \subset \{F_\alpha \cap Y\}_\alpha$ such that

$$K \subset \bigcup_{i=1}^n (F_{\alpha_i} \cap Y) \leftarrow (\text{as } K \text{ is compact in } Y).$$

but since $F_\alpha \cap Y \subset F_\alpha$ for each α,

$$K \subset \bigcup_{i=1}^n F_\alpha \quad \text{also},$$

hence K is compact in X.

$(\Leftarrow)$ If K is compact in X and $\{F_\alpha\}_\alpha$ are open sets in Y which cover K, then by the previous theorem there are open sets, $\{G_\alpha\}_\alpha$, in X such that $F_\alpha = G_\alpha \cap Y$ for each α. Since the $\{F_\alpha\}_\alpha$ cover K and $K \subset Y$ we have

$$K \subset \bigcup_\alpha (G_\alpha \cap Y) \subset \bigcup_\alpha G_\alpha,$$

and hence as K is compact in X there are some $\{G_{\alpha_1}, \dots, G_{\alpha_n}\} \subset \{G_\alpha\}_\alpha$ such that

$$K \subset \bigcup_{i=1}^n G_{\alpha_i}.$$

Since $K \subset Y$ also, we have that

$$K = K \cap Y \subset \left(\bigcup_{i=1}^n G_{\alpha_i}\right) \cap Y = \bigcup_{i=1}^n (G_{\alpha_i} \cap Y) = \bigcup_{i=1}^n F_{\alpha_i}$$

also, hence K is compact in Y. $\blacksquare$

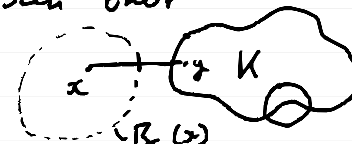
Remark: By this result we can just say 'compact' rather than compact in; thus saying 'let (X,d) be a compact metric space' makes sense as a sentence (as opposed to open).

We also have the following properties for compact sets:

Theorem: Compact sets are closed.

Proof: Let (X,d) be a metric space and $K \subset X$ compact. Suppose that $x \in K^c$ and for each $y \in K$ set $r_y = \frac{1}{2} d(x,y)$. As K is compact there are $y_1, \dots, y_n \in K$ such that

$$K \subset \bigcup_{i=1}^{n} B_{r_{y_i}}(y_i),$$



and notice that by the choices of the r_{y_i} we have that $B_{r_{y_i}}(x) \cap B_{r_{y_i}}(y_i) = \emptyset$ for each $i = 1, \dots, n$; hence

setting $r = \min_{i=1,\dots,n} \{r_{y_i}\} > 0$ we have $B_r(x) \cap B_{r_{y_i}}(y_i) = \emptyset$, and so $B_r(x) \cap K = \emptyset$. Hence $B_r(x) \subset K^c$ and so K^c is open. ∎

Theorem: Closed subsets of compact sets are compact.

Proof: Let (X,d) be a metric space and $F \subset K \subset X$ where F is closed and K is compact. Let $\{F_\alpha\}_\alpha$ be an open cover of F and notice that since $F \subset K$ we have that $\{F_\alpha\}_\alpha \cup F^c$ is an open cover of K; as K is compact there exist $\{F_{\alpha_1}, \dots, F_{\alpha_n}\} \subset \{F_\alpha\}_\alpha$ such that

$$K \subset \left(\bigcup_{i=1}^{n} F_{\alpha_i} \right) \cup F^c.$$

Since $F \subset K$ and $F \cap F^c = \emptyset \Rightarrow F \subset \bigcup_{i=1}^{n} F_{\alpha_i}$ and hence F is compact. ∎

Remark: This actually shows that if F is closed and K is compact then $F \cap K$ is compact (K compact $\Rightarrow$ closed so $F \cap K$ closed).

End L7

Compact sets also behave well under infinite intersections, just as closed sets did:

Theorem: Let (X,d) be a metric space and $\{K_\alpha\}_\alpha$ be a collection of compact sets. If $\bigcap_{i=1}^{n} K_{\alpha_i} \neq \phi$ for any finite subcollection $\{K_{\alpha_1},\dots,K_{\alpha_n}\} \subset \{K_\alpha\}_\alpha$, then $\bigcap_\alpha K_\alpha \neq \phi$.

(Finite intersection property)

Proof: If $\bigcap_\alpha K_\alpha = \phi$ then there is some α_0 such that no point of K_{α_0} lies in all of the $\{K_\alpha\}_\alpha \Rightarrow$ each $x \in K_{\alpha_0}$ is such that $x \in K_\alpha^c$ for some α. Thus,

$$K_{\alpha_0} \subset \bigcup_\alpha K_\alpha^c,$$

but as K_{α_0} is compact and the K_α^c are open, there must exist $\alpha_1,\dots,\alpha_n$ such that

$$K_{\alpha_0} \subset \bigcup_{i=1}^{n} K_{\alpha_i}^c.$$

However, this means $K_{\alpha_0} \cap K_{\alpha_1} \cap \dots \cap K_{\alpha_n} = \phi$ which contradicts our assumption; thus $\bigcap_\alpha K_\alpha \neq \phi$. ∎

Remark: If $\{K_n\}_{n\geq 1}$ is a sequence of non-empty compact sets such that $K_n \supset K_{n+1}$ (nested) for $n \geq 1 \Rightarrow \bigcap_{n\geq 1} K_n \neq \phi$.

Theorem: Let $E \subset K$ be an infinite subset of a compact set, then E has a limit point in K.

Proof: Suppose not, then every $x \in K$ is such that there is some $r_x > 0$ with $B_{r_x}(x)$ containing at most one point of E (namely x, if $x \in E$). Notice that

$$K \subset \bigcup_{x \in K} B_{r_x}(x)$$

and hence $\{B_{r_x}(x)\}_{x\in K}$ is an open cover of K. However, as E is infinite, $E \subset K$, and each ball contains at most one point in E, no finite subcover will cover E (or K); this contradicts the compactness of K. ∎

//

We now focus our attention to compact sets in Euclidean spaces:

Theorem: If $\{I_n\}_{n\geq 1}$ is a sequence of closed intervals in $\mathbb{R}$ such that $I_n \supset I_{n+1}$ for $n\geq 1$, then $\bigcap_{n\geq 1} I_n \neq \emptyset$.

Proof: Let $I_n = [a_n, b_n]$ for $n\geq 1$ and $E = \{a_n \mid n\geq 1\}$. As $b_n \leq b_1$ for $n\geq 1 \Rightarrow E$ is bounded above by b_1 and so, as $a_1 \in E \Rightarrow E \neq \emptyset$, the least upper bound property for $\mathbb{R} \Rightarrow \sup E$ exists. As $I_n \supset I_{n+1}$ for $n\geq 1$ we have

$$a_n \leq a_m \leq b_m \leq b_n \quad \text{for } m\geq n,$$

hence $\quad a_m \leq \sup E \leq b_m \Rightarrow \sup E \in I_m$ for $m\geq 1$,

i.e. $\quad \sup E \in \bigcap_{n\geq 1} I_n.$ ☒

This also generalises to $\mathbb{R}^n$ with the following notion:

Definition: An <u>n-cell</u> in $\mathbb{R}^n$ is a set $I \subset \mathbb{R}^n$ of the form

$$I = [a_1, b_1] \times \dots \times [a_n, b_n],$$

where $a_i \leq b_i$ for $i = 1, \dots, n$.

Theorem: If $\{I_n\}_{n\geq 1}$ is a sequence of n-cells in $\mathbb{R}^n$ such that $I_m \supset I_{m+1}$ for $m\geq 1$, then $\bigcap_{m\geq 1} I_m \neq \emptyset$.

Proof: Let $I_m = [a_1^m, b_1^m] \times \dots \times [a_n^m, a_n^m]$ for $m\geq 1$ and apply the previous theorem to find

$$x_i \in \bigcap_{m\geq 1} [a_i^m, b_i^m],$$

then $(x_1, \dots, x_n) \in \bigcap_{m\geq 1} I_m$. ☒

We will now work to classify compact sets in Euclid-ean spaces; first showing that n-cells are compact and then finding equivalent characterisations of compactness for subsets of $\mathbb{R}^n$.

Theorem: Every n-cell is compact.

Proof: Let $I = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ be an n-cell. We set

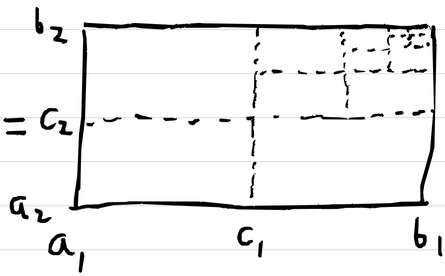
$$\delta = \left(\sum_{i=1}^{n} (b_i - a_i)^2 \right)^{\frac{1}{2}} \geq 0,$$

so that if $x, y \in I$, then $|x-y| \leq \delta$. For a contradiction suppose that $\{F_\alpha\}_\alpha$ is an open cover of I with no finite subcover.

Fix $c_i = \frac{a_i + b_i}{2}$ for $i = 1, \ldots, n$ (midpoint of $[a_i, b_i]$), the intervals $[a_i, c_i]$ and $[c_i, b_i]$ for $i = 1, \ldots, n$ give us 2^n possible n-cells (2 choices for n intervals) whose union covers I. At least one of these n-cells, I_1, has no finite subcover by $\{F_\alpha\}_\alpha$ (or I would be covered since $I_1 \subset I$). Continuing by induction we obtain a sequence of n-cells, $\{I_n\}_{n \geq 1}$, such that:

1. $I_n \supset I_{n+1}$ for $n \geq 1$.
2. I_n has no finite subcover by $\{F_\alpha\}_\alpha$.
3. If $x, y \in I_n$, then $|x-y| \leq 2^{-n}\delta$. → ($\frac{1}{2}$ size each step)

By the previous theorem (and 1) there is some $\check{x} \in \bigcap_{n \geq 1} I_n \subset I$, and since the $\{F_\alpha\}_\alpha$ cover I, $\check{x} \in F_\alpha$ for some α. As F_α is open there is $r > 0$ such that $B_r(\check{x}) \subset F_\alpha$. Choosing $n \geq 1$ sufficiently large so that $2^{-n}\delta < r$ (this is possible by the Archimedean property), as $\check{x} \in I_n$, 3 above implies that $I_n \subset F_\alpha$, contradicting 2! ☒



$I = $ (picture) $\subset \mathbb{R}^2$ ← (Picture of proof idea)

Theorem: Let $E \subset \mathbb{R}^n$, then the following are equivalent:

1. E is compact.
2. E is closed and bounded.
3. Every infinite subset of E has a limit point in E.

Remarks:

- By <u>bounded</u> we mean that there is some $D > 0$ such that $d(x,y) \leq D$ if $x, y \in E$; this notion makes sense for any metric space.

- The equivalence of (1) and (2) is often called the Heine-Borel theorem.

- In general metric spaces, it is <u>not</u> true that (2) $\Rightarrow$ (1) or (2) $\Rightarrow$ (3), but (1) $\Leftrightarrow$ (3) is true.
  $\hookrightarrow$ (see HWK 5) $\hookrightarrow$ (later on normed vector spaces).

Proof: For (2) $\Rightarrow$ (1), if E is closed and bounded then there is some n-cell, I, with $E \subset I$; hence E is compact as a closed subset of a compact set.

For (1) $\Rightarrow$ (3), we saw that every infinite subset of a compact set contains a limit point.

For (3) $\Rightarrow$ (2), if every infinite subset of E has a limit point in E then we first note that E must be bounded; if not then there must exist some sequence, $(x_n) \subset E$, with $|x_n| > n$ for each $n \geq 1$, this sequence is an infinite subset of E with no limit point. Hence, we see that E must be bounded, it remains to show that E is closed.

If E were not closed then there must exist some limit point of E, $x \in E^c$. Since x is a limit point of E there must exist some sequence, $(x_n) \subset E$, such that $|x - x_n| < \frac{1}{n}$ for $n \geq 1$. $\rightarrow$

As $(x_n) \subset E$ is an infinite subset of E there must exist some limit point in E. If $(x_n) \subset E$ had some limit point $y \in E$ such that $y \neq x$ then as

$$|x-y| = |x - x_n + x_n - y| \underset{(\Delta \text{ineq})}{\leq} |x - x_n| + |x_n - y|$$

and so

$$|x_n - y| \geq |x-y| - |x - x_n| > \frac{|x-y|}{2} > 0$$

for $n \geq 1$ sufficiently large (since $|x - x_n| < \frac{|x-y|}{2}$ for large n). But then y cannot be a limit point of (x_n); hence x is the only limit point of $(x_n) \Rightarrow x \in E$, contradicting our assumption that $x \in E^c$. Thus E is closed.

We thus have (2) $\Rightarrow$ (1) $\Rightarrow$ (3) $\Rightarrow$ (2), and are done. ∎

As n-cells are compact we also have:

Theorem: Every bounded infinite subset of $\mathbb{R}^n$ has a limit point.

Proof: If $E \subset \mathbb{R}^n$ is bounded then there is some n-cell, $I \subset \mathbb{R}^n$, such that $E \subset I$. As n-cells are compact and E is infinite then we saw that E must have a limit point in I. ∎

End d8

Perfect Sets, Uncountability of $\mathbb{R}$, and the Cantor Set

We will develop a notion for sets in metric spaces allowing us to construct uncountable subsets of Euclidean spaces:

Definition: Let (X,d) be a metric space and $P \subset X$, we say that P is *perfect* if $P = P'$; namely if P is closed and every point of P is a limit point of P. Equivalently, P is perfect if is closed and has no isolated points.

Examples: • $\emptyset$, $[a,b]$ ($a \neq b$), $\mathbb{R} \subset \mathbb{R}$ are all perfect, as are n-cells.

• Later we will construct a perfect subset of $\mathbb{R}$ which contains no line segment (Cantor set)!

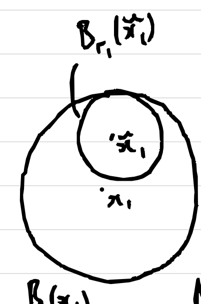
• $[0,1] \cap \mathbb{Q}$ <u>is</u> perfect as a subset of $\mathbb{Q}$ but <u>not</u> as a subset of $\mathbb{R}$!

Remark: The example above shows that being perfect depends on the ambient space (like openness).

Theorem: If $P \subset \mathbb{R}^n$ is nonempty and perfect, then P is uncountable.

Proof: As P is nonempty and perfect there must exist some limit point, hence P is infinite. Suppose that P were countable, then there is a sequence, $(x_n) \subset P$, such that $P = (x_n)$ and $x_n = x_m \iff m = n$.

We will construct a sequence of nonempty compact sets, $\{K_n\}_{n \geq 1}$, such that $K_{n+1} \subset K_n \subset P$ for $n \geq 1$, so that $\bigcap_{n \geq 1} K_n \neq \emptyset$, but with $x_n \notin K_n$ for $n \geq 1$, so that $\bigcap_{n \geq 1} K_n = \emptyset$; giving a contradiction.



Firstly, for $x_1 \in P$ there is some $\hat{x}_1 \in (B_1(x_1) \setminus \{x_1\}) \cap P$. Setting $r_1 = \frac{1}{2}\min\{|x_1 - \hat{x}_1|, 1 - |x_1 - \hat{x}_1|\}$ we see that $x_1 \notin B_{r_1}(\hat{x}_1)$ and $\overline{B_{r_1}(\hat{x}_1)} \subset B_1(x_1)$.

Next, there is some $\hat{x}_2 \in (B_{r_1}(\hat{x}_1) \setminus \{x_2\}) \cap P$. Setting $r_2 = \frac{1}{2}\min\{|\hat{x}_2 - x_2|, r_1 - |\hat{x}_2 - \hat{x}_1|\}$ we see that $x_2 \notin B_{r_2}(\hat{x}_2)$ and $\overline{B_{r_2}(\hat{x}_2)} \subset B_{r_1}(\hat{x}_1)$. Continuing this inductively we have points $\{\hat{x}_n\}_{n \geq 1} \subset P$ and radii $r_n > 0$ for each <u>$n \geq 1$ such</u> that $x_n \notin B_{r_n}(\hat{x}_n)$ and $\overline{B_{r_{n+1}}(\hat{x}_{n+1})} \subset B_{r_n}(\hat{x}_n)$ for $n \geq 1$.

Now note that since $\overline{B_{r_n}(x_n)} \subset B_1(x_1)$ for each $n \geq 1$, these sets are closed (by Hwk 4) and bounded; hence compact by Heine-Borel. As P is closed and the $\{\overline{B_{r_n}(x_n)}\}_{n\geq 1}$ are compact we see that $K_n = \overline{B_{r_n}(x_n)} \cap P$ is nonempty, compact, and such that both $K_{n+1} \subset K_n$ and $x_n \notin K_n$ for $n \geq 1$.

Thus, $\bigcap_{n\geq 1} K_n$ is a nonempty subset of P, but cannot contain any point of P: a contradiction. Hence P is uncountable. ☒

Remark: Since any closed interval $(a \neq b)$ in $\mathbb{R}$ is perfect, any open interval is uncountable; thus $\mathbb{R}$ is uncountable!

We now construct a perfect subset of $\mathbb{R}$ which does not contain any open interval; the Cantor set:

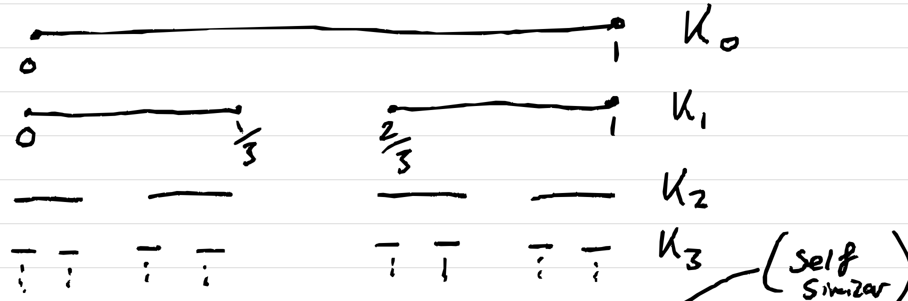
Firstly, let $K_0 = [0,1] \subset \mathbb{R}$ and remove the interval $(\frac{1}{3}, \frac{2}{3})$ to define $K_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$; i.e. remove the middle third. Remove the middle third of the intervals in K_1 to define the set $K_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{3}{9}] \cup [\frac{6}{9}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$. Inductively removing the middle third of each interval we obtain a sequence of nonempty compact sets $\{K_n\}_{n\geq 1}$ such that $K_{n+1} \subset K_n$ and K_n is the union of 2^n intervals of length 3^{-n} for each $n \geq 1$; the Cantor set,

$$C = \bigcap_{n\geq 1} K_n,$$ is then non-empty and compact.

To see that C is perfect, suppose that $x \in C$ and let $S \subset \mathbb{R}$ be any open neighbourhood of x. Then, for large $n \geq 1$ let $I_n \subset K_n \cap S$ be the interval of length 3^{-n} containing x, and set $x_n \neq x$ an endpoint of I_n ($x_n \in C$). Thus any neighbourhood of x contains infinitely many x_n and so x is a limit point of C; as $x \in C$ was arbitrary, $C = C'$ and so C is perfect.

If $(a,b) \subset C$ and $x \in (a,b)$ then by the construct-ion of C, for large $n \geq 1$ there is an interval $I_n \subset K_n$ such that $x \in I_n \subset (a,b)$ (since $|I_n| < 3^{-n}$). By the construction of C, I_{n+1} removes the middle third of I_n and hence $I_n \not\subset K_{n+1}$ which contradicts $I_n \subset (a,b) \subset C$. Hence C contains no open interval.

Visually,

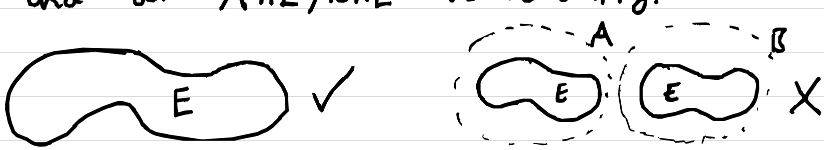


Remarks:
- This is an instance of a 'fractal' line object, many of which can be constructed by iterative procedures.
- An interesting property of C is that it is uncountable but has zero length (measure).
- One does not have to remove the same proportion of $[0,1]$ at each step, allowing for other constructions, see e.g. fat Cantor sets. $\boxed{\text{End L9}}$

Connectedness

We now define what it means for a set to be 'in one piece':

Definition: Let (X,d) be a metric space and $E \subset X$, then we say that E is connected in X if there do not exist disjoint sets $A, B \subset X$ open in X such that $E \subset A \cup B$, and both $A \cap E$, $B \cap E$ are nonempty.



Just as compactness is independent of the ambient space, connectedness is also:

<u>Theorem:</u> Let (X,d) be a metric space and $E \subset X$, then E is connected in $E \iff E$ is connected in X.

<u>Proof:</u> ($\Rightarrow$) If E were not connected in X, then by definition there are disjoint sets $A, B \subset X$ open in X such that $E = A \cup B$ and $A \cap E, B \cap E \neq \emptyset$. Since we saw that $A \cap E, B \cap E$ are open in $E \Rightarrow E$ is not connected in E; hence connected in $E \Rightarrow$ connected in X.

($\Leftarrow$) If E is not connected in E, again we have disjoint sets $C, D \subset E$ open in E such that $E \subset C \cup D$ and both $C \cap E, D \cap E \neq \emptyset$. We will construct, from C and D, sets $A, B \subset X$ open in X such that $A \cap B = \emptyset$ such that $A = C \cap E$ and $B = C \cap E$; hence $E \subset A \cup B$ and so E is not connected in X. Since C and D are open in E, for $x \in C$ and $y \in D$ there are $r_x, r_y > 0$ such that both $B_{r_x}(x) \cap E \subset C$ and $B_{r_y}(y) \cap E \subset D$; moreover, since $C \cap D = \emptyset$ we must have that $d(x,y) \geq \max\{r_x, r_y\}$ (or otherwise, if $d(x,y) < r_x$ or $r_y \Rightarrow y \in B_{r_x}(x)$ or $x \in B_{r_y}(y)$ ※). Let $\tilde{r}_x = \frac{r_x}{2}$ and $\tilde{r}_y = \frac{r_y}{2}$ for each $x \in C$ and $y \in D$, then if there was some $z \in B_{\tilde{r}_x}(x) \cap B_{\tilde{r}_y}(y)$ we would have,

$$d(x,y) \leq d(x,z) + d(z,x) < \frac{r_x}{2} + \frac{r_y}{2} \leq \max\{r_x, r_y\},$$

but this contradicts $d(x,y) \geq \max\{r_x, r_y\}$; hence we have $B_{\tilde{r}_x}(x) \cap B_{\tilde{r}_y}(y) = \emptyset$ for all $x \in C, y \in D$. Setting

$$A = \bigcup_{x \in C} B_{\tilde{r}_x}(x), \quad B = \bigcup_{y \in D} B_{\tilde{r}_y}(y),$$

we have open sets in X such that $A \cap B = \emptyset$, $E \subset A \cup B$, $A \cap E, B \cap E \neq \emptyset$ (and $A \cap E = C$, $B \cap E = D$) so that E is not connected in X. $\boxtimes$

Remarks: • By the previous result we can just say that a space is connected without reference to an ambient space (c.f. compactness).

• Let us call a set $E \subset X$ clopen in X if it is both closed and open; for example in any metric space, (X,d), both X, ϕ are clopen in X.

Theorem: Let (X,d) be a metric space, then X is connected $\Leftrightarrow$ ϕ, X are the only clopen sets in X.

Proof: ($\Leftarrow$) If X is not connected then there are disjoint, nonempty sets $A, B \subset X$ open in X such that $X = A \cup B$. Notice that $A^c = B$ and $B^c = A$ and hence both A and B are also closed, hence clopen. Moreover, neither $A, B = \phi, X$ since they are nonempty. Hence if the only clopen sets in X are $\phi, X \Rightarrow X$ connected.

($\Rightarrow$) If there is some clopen $A \neq \phi, X$ in X then, since $A^c \neq \phi, X$ and A^c is also clopen, we have that $X = A \cup A^c$ for nonempty sets $A, A^c \subset X$ open in X; hence X is not connected and so X connected $\Rightarrow$ the only clopen sets in X are ϕ, X. ☒

Example: $(0,1) \cup (2,3)$ is not connected since $(0,1)^c = (2,3)$ is open, hence both $(0,1)$ and $(2,3)$ are clopen.

Let's classify connected subsets of the line:

Theorem: $E \subset \mathbb{R}$ is connected $\Leftrightarrow$ whenever $x, y \in E$ and $x < y$, then if $x < z < y$ then $z \in E$ (i.e. $(x,y) \subset E$). Thus, if $E \subset \mathbb{R}$ is connected it must be one of

$$\mathbb{R}, (-\infty, b], (-\infty, b), [a, +\infty), (a, +\infty), [a, b], [a, b), (a, b], \text{ or } (a, b) \text{ for } a \leq b.$$

Proof: ($\Rightarrow$) If $x, y \in E$ with $x < y$ but there was some $x < z < y$ such that $z \notin E$ (i.e. $(x, y) \not\subset E$), then $E \subset (-\infty, z) \cup (z, \infty)$ and hence E is not connected; hence if E is connected $\Rightarrow$ property does hold.

($\Leftarrow$) If E is not connected then there are disjoint sets, $A, B \subset \mathbb{R}$, such that $A \cap E, B \cap E \neq \emptyset$, and $E \subset A \cup B$. Let $x \in A$, $y \in B$ and suppose without loss of generality that $x < y$ (or swap A, B roles). Setting $z = \sup(A \cap [x, y])$ (exists as $A \cap [x, y]$ bounded by y, $\neq \emptyset$) we see that $z \notin B$ (as $A \cap B = \emptyset$ and B open would imply $z - \varepsilon \in B$ bounds $A \cap [x, y]$ from above); hence $z < y$. Similarly, as A is open, $x < z$, and so $z \notin A$ (or otherwise z is not an upper bound for A since A is open). Thus $x < z < y$ but $z \notin A$ and $z \notin B$, hence $z \in A^c \cap B^c = (A \cup B)^c \subset E^c$ so $z \notin E$; hence if the property holds $\Rightarrow$ E is connected. $\square$

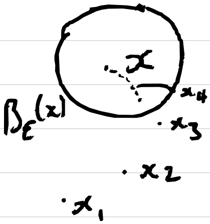
Examples:
- $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} = \{z \in \mathbb{C} \mid |z| = 1\}$ is connected. (we will justify later).
- If $A \subset \mathbb{R}^2$ is countable $\Rightarrow$ $\mathbb{R}^2 \setminus A$ is connected; so $\mathbb{R}^2 \setminus \mathbb{Q}^2$ is! (Thus it is hard to characterise connected subsets of $\mathbb{R}^2$).
- $\mathbb{R}^2 \setminus \{x\text{-axis}\}$ is disconnected.

Remark: The notion of connectedness, like compactness, makes sense for general topological spaces.

$\boxed{\text{End 210}}$

Sequences and Series

We now have built up the machinery to precisely define and study limits of sequences in metric spaces; as a consequence we are also able to make sense of infinite sums, i.e. series.



Definition: Let (X,d) be a metric space, then a sequence, $(x_n) \subset X$, is said to converge in X if there is a point $x \in X$ such that for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that if $n \geq N$ then $d(x_n, x) < \varepsilon$. We then say that x is the limit of (x_n) and write

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

If (x_n) does not converge then it diverges.

Remarks:

- $x_n \to x \iff$ for every $\varepsilon > 0$, $x_n \in B_\varepsilon(x)$ for all sufficiently large n. Sometimes we say eventually to mean there is some $N \in \mathbb{N}$ such that $n \geq N$ makes a property true; thus $x_n \to x \iff (x_n)$ eventually lies in every neighbourhood of x.

- We specify convergence in X since, $\left(\frac{1}{n}\right) \to 0$ in $\mathbb{R}$ but not in $\mathbb{R}\setminus\{0\}$ (as $0 \notin \mathbb{R}\setminus\{0\}$).

Examples:

- $\left(\frac{1}{n}\right) \to 0$ in $\mathbb{R}$ as noted above.

- (n^2) diverges and is unbounded.

- $(-1)^n$ and (i^n) diverge but are bounded.

- $\left(1 + \frac{(-1)^n}{n}\right) \to 1$ in both $\mathbb{Q}$ and $\mathbb{R}$.

- Any constant sequence converges.

- $(e^{i/n}) \to 1$ in both $\mathbb{C}$ and $S^1 \subset \mathbb{C}$.

- $\left(\frac{1}{n}, \frac{1}{n^2}\right) \to (0,0)$ in $\mathbb{R}^2$; equivalently $\left(\frac{1}{n} + \frac{i}{n^2}\right) \to 0$ in $\mathbb{C}$.

We have the following properties:

Theorem: Let (X,d) be a metric space and $(x_n) \subset X$ a sequence. Then:

1. Limits are unique.
2. Convergent sequences are bounded.
3. $x_n \to x \iff$ every neighbourhood of x contains all but finitely many of the (x_n).
4. If $E \subset X$ and x is a limit point of E, then there is a sequence $(x_n) \subset E$ such that $x_n \to x$.

Proof: For 1, if there exist $x, \tilde{x} \in X$ such that both $x_n \to x$ and $x_n \to \tilde{x}$, then for each $\varepsilon > 0$ there exist $N, \hat{N} \in \mathbb{N}$ such that
$$\begin{cases} d(x_n, x) < \varepsilon/2 & \text{for } n \geq N \\ d(x_n, \tilde{x}) < \varepsilon/2 & \text{for } n \geq \hat{N}. \end{cases}$$
Thus, for $n \geq \max\{N, \hat{N}\}$ we have
$$d(x, \tilde{x}) \leq d(x, x_n) + d(x_n, \tilde{x}) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon,$$
hence $d(x, \tilde{x}) < \varepsilon$ for all $\varepsilon > 0 \Rightarrow d(x, \tilde{x}) = 0$ so $x = \tilde{x}$.

For 2 if $x_n \to x$ for $x \in X$ then there is some $N \in \mathbb{N}$ such that $d(x_n, x) \leq 1$ for $n \geq N$. Setting
$$M = \max\{1, d(x_1, x), \dots, d(x_{N-1}, x)\} > 0.$$
We see that $d(x_n, x) \leq M$ for all $n \geq 1$, hence (x_n) is bounded.

$(\Rightarrow)$ For 3, if $x_n \to x$ and $B_r(x)$ for $r > 0$ is any neighbourhood of x then there is some $N \in \mathbb{N}$ such that we have $x_n \in B_r(x)$ for $n \geq N$; hence at most $N-1$ of the (x_n) lie outside of $B_r(x)$.

$(\Leftarrow)$ On the other hand if every $B_r(x)$ for $r > 0$ contains all but finitely many of the (x_n) then there is $N \in \mathbb{N}$ such that $x_n \in B_r(x)$ for $n \geq N$; hence $x_n \to x$.

For 4, if $E \subset X$ and x is a limit point of E then for each $n \geq 1$ we have that the set $(B_{\frac{1}{n}}(x) \setminus \{x\}) \cap E$ is nonempty; hence we may choose $x_n \in (B_{\frac{1}{n}}(x) \setminus \{x\}) \cap E$ for each $n \geq 1$. We observe that for each $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that $\frac{1}{N} < \varepsilon$ (Archimedean property) and hence if $n \geq N$ we have

$$d(x_n, x) \leq \frac{1}{n} \leq \frac{1}{N} < \varepsilon,$$

hence the sequence (x_n) converges to x in X. $\boxtimes$

Since in $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^n$ we have algebraic operations (addition and scalar multiplication), we can see how they relate to limits:

<u>Theorem:</u> Suppose $(x_n), (y_n) \subset \mathbb{C}$ are sequences with $x_n \to x$ and $y_n \to y$. Then:

1. $(x_n + y_n) \to x + y$.
2. $(x_n y_n) \to xy$.
3. $(1/x_n) \to 1/x$ if $x_n, x \neq 0$ for all $n \geq 1$.

<u>Proof:</u> For 1, given $\varepsilon > 0$ there are $N, \tilde{N} \in \mathbb{N}$ such that

$$\begin{cases} |x_n - x| < \varepsilon/2 & \text{for } n \geq N, \\ |y_n - y| < \varepsilon/2 & \text{for } n \geq \tilde{N}. \end{cases}$$

Thus, for $n \geq \max\{N, \tilde{N}\}$ we have

$$\begin{aligned} |(x_n + y_n) - (x - y)| &= |(x_n - x) + (y_n - y)| \\ &\leq |x_n - x| + |y_n - y| < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon; \end{aligned}$$

hence $(x_n + y_n) \to x + y$.

For 2, we first note that if $c \in \mathbb{C} \setminus \{0\}$ and $\varepsilon > 0$ then choosing $N \in \mathbb{N}$ so that $|x_n - x| < \varepsilon/c$ for $n \geq N$, we see that $(cx_n) \to cx$ for any $c \in \mathbb{C}$ (the case where $c = 0$ is immediate since $(0 \cdot x_n) = (0)$, which is a constant 0 sequence, hence convergent to 0).

Note that we have the identity

$$x_n y_n - xy = (x_n - x)(y_n - y) + x(y_n - y) + y(x_n - x). \quad (+)$$

Then, given $\varepsilon > 0$ there $N, \tilde{N} \in \mathbb{N}$ such that

$$\begin{cases} |x_n - x| < \sqrt{\varepsilon} & \text{for } n \geq N, \\ |y_n - y| < \sqrt{\varepsilon} & \text{for } n \geq \tilde{N}, \end{cases}$$

we see that if $n \geq \max\{N, \tilde{N}\}$ then $|(x_n - x)(y_n - y)| < \varepsilon$; hence $((x_n - x)(y_n - y)) \to 0$. Thus, from the previous reasoning and ① we have $\begin{cases} (x(y_n - y)) \to 0, \\ (y(x_n - x)) \to 0, \end{cases}$

and so by $(+)$ and ① we have

$$(x_n y_n - xy) = ((x_n - x)(y_n - y) + x(y_n - y) + y(x_n - x)) \to 0;$$

hence $(x_n y_n - xy) \to 0$ so by ① $(x_n y_n) \to xy$.

For ③, we note that

$$\left| \frac{1}{x_n} - \frac{1}{x} \right| = \frac{|x_n - x|}{|x_n||x|} \quad (\#);$$

(We get $|x_n| > \frac{|x|}{2}$ and $|x_n - x| < \frac{|x|^2}{2} \cdot \varepsilon$ eventually to show $(\#) < \varepsilon$.)

We want to show that given $\varepsilon > 0$ we can choose some $N \in \mathbb{N}$ such that $(\#) < \varepsilon$ for $n \geq N$. Firstly, since $(x_n) \to x$ there is some $N \in \mathbb{N}$ such that

$$|x_n - x| < \frac{|x|}{2} \text{ for } n \geq N.$$

(reverse Δ ineq.) Noting the reverse triangle inequality,

$$||x| - |y|| \leq |x - y|;$$

which follows from the usual triangle inequality since

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

Now as $||x_n| - |x|| \leq |x_n - x| < \frac{|x|}{2}$ for $n \geq N$ we see that

$$-\frac{|x|}{2} < |x_n| - |x| < \frac{|x|}{2},$$

and so $|x_n| > \frac{|x|}{2}$ for $n \geq N$. Finally, again as $x_n \to x$ there is some $\hat{N} \in \mathbb{N}$ such that

$$|x_n - x| < \frac{|x|^2}{2} \cdot \varepsilon \text{ for } n \geq \hat{N};$$

for $n \geq \max\{N, \hat{N}\}$ we have $(\#) < \varepsilon$ so $\left(\frac{1}{x_n}\right) \to \frac{1}{x}$. $\blacksquare$

<u>Theorem</u>: Suppose $(x_n), (y_n) \subset \mathbb{R}^k$ and $(c_n) \subset \mathbb{R}$ are sequences with $x_n \to x$, $y_n \to y$ in $\mathbb{R}^k$, and $c_n \to c$ in $\mathbb{R}$. Then:

1. If $x_n = (a_n^1, \dots, a_n^k) \in \mathbb{R}^k$ for $n \geq 1$ then $x_n \to x = (a^1, \dots, a^n) \iff a_n^i \to a^i$ for $i = 1, \dots, k$.
2. $(x_n + y_n) \to x + y$, $(x_n \cdot y_n) \to x \cdot y$, and $(c_n x_n) \to cx$.

<u>Proof</u>: Note that ② follows from ① and the previous theorem, so we just prove ①.

($\Rightarrow$) If $x_n \to x = (a_1, \dots, a_n)$ then noting that
$$|a_n^i - a^i| \leq \left(\sum_{i=1}^{k} |a_n^i - a^i|^2 \right)^{\frac{1}{2}} = |x_n - x|,$$
we see that $a_i^n \to a_i$ for $i = 1, \dots, k$.

($\Leftarrow$) If $a_n^i \to a^i$ for $i = 1, \dots, k$, then for each $\varepsilon > 0$ there exist $N_1, \dots, N_k \in \mathbb{N}$ such that
$$|a_n^i - a^i| < \frac{\varepsilon}{\sqrt{k}} \quad \text{if } n \geq N_i \text{ for } i = 1, \dots, k.$$
Then if $n \geq \max\{N_1, \dots, N_k\}$ we see that
$$|x_n - x| = \left(\sum_{i=1}^{k} |a_n^i - a^i|^2 \right)^{\frac{1}{2}} < \left(\sum_{i=1}^{k} \frac{\varepsilon^2}{k} \right)^{\frac{1}{2}} = \varepsilon;$$
hence $x_n \to x$. $\boxed{\text{End L11}}$ $\blacksquare$

<u>Subsequences</u>

Note that while $((-1)^n)$ diverges, the sequences, (± 1), consisting of the even/odd terms do converge.

<u>Definition</u>: Let (X, d) be a metric space and $(x_n) \subset X$ a sequence. (indexed by k not n) For a sequence $(n_k) \subset \mathbb{N}$ such that $1 \leq n_k < n_{k+1}$ for each $k \geq 1$, the sequence $(x_{n_k}) \subset (x_n)$ is called a <u>subsequence</u> of (x_n). If (x_{n_k}) converges in X then its limit is called a <u>subsequential limit</u>.

Theorem: Let (X,d) be a metric space and $(x_n) \subset X$ a sequence, then $x_n \to x \iff$ every subsequence of (x_n) converges to x.

Proof: ($\Leftarrow$) If every subsequence of (x_n) converges to x, then the subsequence (x_n) itself also converges to x (i.e. $n_k = k$ for $k \geq 1$).

($\Rightarrow$) If $x_n \to x$ and $(n_k) \subset \mathbb{N}$ is such that $1 \leq n_k < n_{k+1}$ for each $k \geq 1$, then $n_k \geq k$ by induction (check). If $\varepsilon > 0$ then there is some $N \in \mathbb{N}$ such that $d(x_k, x) < \varepsilon$ for $k \geq N$; as $n_k \geq k$ this shows that $d(x_{n_k}, x) < \varepsilon$ whenever $n_k \geq k \geq N$ and hence $x_{n_k} \to x$. $\blacksquare$

Theorem: Let (X,d) be a compact metric space and $(x_n) \subset X$ a sequence, then (x_n) has a convergent subsequence. In particular, every bounded sequence in $\mathbb{R}^k$ has a convergent subsequence (this is the <u>Bolzano-Weierstrass thm</u>)

Proof: If the range of (x_n) is finite (ie. $\{x_n \mid n \geq 1\}$ is finite) then at least one point $x \in \{x_n \mid n \geq 1\}$ appears in (x_n) infinitely many times; thus there is a sequence $(n_k) \subset \mathbb{N}$ with $1 \leq n_k < n_{k+1}$ and $x_{n_k} = x$ for each $k \geq 1$. Then (x_{n_k}) is a subsequence of (x_n) converging to x.

If the range of (x_n) is infinite, then as X is compact the infinite set $\{x_n \mid n \geq 1\} \subset X$ has a limit point $x \in X$. Similarly to previous reasoning (4 of sequence properties) we can choose $x_{n_k} \in (x_n)$ such that $d(x_{n_k}, x) < 1/k$ for some sequence $(n_k) \subset \mathbb{N}$ with $1 \leq n_k < n_{k+1}$ for $k \geq 1$; then (x_{n_k}) is a subsequence of (x_n) converging to x.

If $(x_n) \subset \mathbb{R}^k$ is bounded then $(x_n) \subset I$ for some compact n-cell, $I \subset \mathbb{R}^n$, then we can apply the above. $\blacksquare$

Theorem: Let (X,d) be a metric space and $(x_n) \subset X$ a sequence, then the set of subsequential limits of (x_n) is closed.

Proof: Let E be the set of subsequential limits of (x_n) and x be a limit point of E; we want to show that $x \in E$ from which we see that E is closed. If $E = \emptyset$ then there is nothing to show, so we may suppose that $E \neq \emptyset$ and hence E is infinite.

To show that $x \in E$ it suffices to construct some subsequence of (x_n) converging to x. First fix $n_1 \geq 1$ such that $x_{n_1} \neq x$ (exists since E is infinite) and note that for each $k \geq 1$ there is some $x^k \in E$ such that $d(x^k, x) < 2^{-k} \cdot d(x_{n_1}, x)$ (such $x^k \in E$ exist since x is a limit point of E). As $x^k \in E$ for each $k \geq 1$ we may inductively choose $n_k > n_{k-1} \geq 1$ such that $d(x_{n_k}, x^k) < 2^{-k} \cdot d(x_{n_1}, x)$ also (since for $x^k \in E$ there must be some subsequence of (x_n) converging to x^k). Thus,

$$\begin{aligned} d(x_{n_k}, x) &\leq d(x_{n_k}, x^k) + d(x^k, x) \\ &< 2^{-k} \cdot d(x_{n_1}, x) + 2^{-k} d(x_{n_1}, x) \\ &= 2^{1-k} \cdot d(x_{n_1}, x), \end{aligned}$$

and hence the subsequence $x_{n_k} \to x$ (for $\varepsilon > 0$ choose $k \geq 1$ s.t. $2^{1-k} \cdot d(x_{n_1}, x) < \varepsilon$). We have thus found a subsequence of $x_n \to x \Rightarrow x \in E$. ∎

Cauchy Sequences

We now make precise the idea that $\mathbb{Q}$ has 'gaps' while $\mathbb{R}$ does not:

Definition: Let (X,d) be a metric space and $(x_n) \subset X$ a sequence, then we say that (x_n) is a <u>Cauchy Sequence</u> if for every $\varepsilon > 0$ there exists some $N \in \mathbb{N}$ such that if $n, m \geq N$ then $d(x_n, x_m) < \varepsilon$.

Theorem: A convergent sequence is a Cauchy Sequence.

Proof: Let (X,d) be a metric space and $x_n \to x$ a convergent sequence in X. For each $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that $d(x_n, x) < \frac{\varepsilon}{2}$ for $n \geq N$. Thus, if $n, m \geq N$ we have

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

and hence (x_n) is a Cauchy Sequence. ∎

Theorem: Let (X,d) be a compact metric space, then every Cauchy sequence converges in X. Moreover, every Cauchy Sequence is bounded (even if X is not compact) and so every Cauchy sequence in $\mathbb{R}^k$ converges.

Proof: If $(x_n) \subset X$ is Cauchy then for each $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that $d(x_n, x_m) < \varepsilon/2$ for $n, m \geq N$. Also, since X is compact we saw that there is a subsequence (x_{n_k}) of (x_n) such that $x_{n_k} \to x \in X$; hence there is some $\tilde{N} \in \mathbb{N}$ such that $d(x_{n_k}, x) < \varepsilon/2$ for $k \geq \tilde{N}$. Thus, for $n \geq N$ we have

$$d(x_n, x) \leq d(x_n, x_{n_{\tilde{N}}}) + d(x_{n_{\tilde{N}}}, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon;$$

and hence $x_n \to x$.

Moreover, even if X is not compact but $(x_n) \subset X$ is Cauchy, then since $d(x_n, x_m) < 1$ for $n, m \geq N_1$ (for some $N_1 \in \mathbb{N}$), we see that setting

$$M = \max\{1, d(x_1, x_N), \dots, d(x_{N-1}, x_N)\} > 0$$

we have $d(x_n, x_N) < M$ for all $n \geq 1 \Rightarrow (x_n)$ is bounded.

Finally, if $(x_n) \subset \mathbb{R}^k$ is Cauchy then by the above reasoning it is bounded. Hence there is some compact n-cell, $I \subset \mathbb{R}^k$, with $(x_n) \subset I$. By the first part of the theorem $\Rightarrow (x_n)$ converges. ∎

Definition: A metric space is said to be complete if every Cauchy sequence is convergent.

Remarks:
- By the previous theorem, every compact metric space is complete. The converse is not true since $\mathbb{R}^n$ was seen to be complete; however there is a notion called totally bounded which, with completene-ss, does imply compactness.
- $\mathbb{Q}$ is not complete; thus completeness is detecting the 'gaps', i.e. irrationals, that are not in $\mathbb{Q}$.

End L12

Monotone Sequences

We saw that convergent sequences are bounded but that bounde-d sequences do not necessarily converge (e.g. $((-1)^n)$, (i^n) etc.). In $\mathbb{R}$ however there is a criteria that guarantees bounded sequences converge:

Definition: A sequence $(x_n) \subset \mathbb{R}$ is monotone if it is increasing or decreasing; i.e. if $x_n \leq x_{n+1}$ or $x_{n+1} \leq x_n$ for all $n \geq 1$. We use the terms strictly increasing/decre-asing if $x_n < x_{n+1}$ / $x_{n+1} < x_n$ for all $n \geq 1$.

Theorem: A monotone sequence converges if and only if it is bounded.

Proof: We already saw that convergent sequences are bounded. If $(x_n) \subset \mathbb{R}$ is a monotone bounded sequence which is increasing (the decreasing case follows by considering $(-x_n)$) then the set $E = \{x_n \mid n \in \mathbb{N}\}$ is bounded and thus $\sup E \in \mathbb{R}$ exists; thus $x_n \leq \sup E$ for $n \geq 1$. For $\varepsilon > 0$ there must exist some $N \in \mathbb{N}$ such that $x_n \in (\sup E - \varepsilon, \sup E + \varepsilon)$ for $n \geq N$ (else $\sup E - \varepsilon$ is an upper bound for E); hence we see that $x_n \to \sup E$. ∎

Theorem: (*Squeeze / Sandwich*) If $(x_n), (y_n), (z_n) \subset \mathbb{R}$ are sequences with $x_n \to x$, $y_n \to y$, and $z_n \to z$, then if $x_n \leq y_n \leq z_n$ for $n \geq 1$ we have $x \leq y \leq z$.

Proof: We prove that $x \leq y$ since the other case will follow by identical reasoning. Suppose that $y < x$ and fix $N, \tilde{N} \in \mathbb{N}$ such that both $\begin{cases} |x_n - x| < \frac{x-y}{2} & \text{for } n \geq N, \\ |y_n - y| < \frac{x-y}{2} & \text{for } n \geq \tilde{N}. \end{cases}$

Thus for $n \geq \max\{N, \tilde{N}\}$ we have both

$$|x_n - x| < \frac{y-x}{2} \Rightarrow x_n - x > \frac{y-x}{2} \Rightarrow x_n > \frac{x+y}{2},$$

$$|y_n - y| < \frac{y-x}{2} \Rightarrow y_n - y < \frac{x-y}{2} \Rightarrow y_n < \frac{x+y}{2};$$

hence $y_n < \frac{x+y}{2} < x_n$ for $n \geq \max\{N, \tilde{N}\}$, a contradiction. ∎

Recall that the extended real line, $\mathbb{R} \cup \{\pm\infty\}$, had the property that every subset had a supremum and an infimum:

Definition: We say that a sequence $(x_n) \subset \mathbb{R}$ <u>diverges to $\pm\infty$</u>, and write $x_n \to \pm\infty$, if for each $M \in \mathbb{R}$ there is some $N \in \mathbb{R}$ such that $x_n \geq M$ / $x_n \leq M$ for $n \geq N$.

Definition: Let $(x_n) \subset \mathbb{R}$ be a sequence and E the set of all subsequential limits (including $\pm\infty$) of (x_n). We then denote the <u>upper and lower limits</u> of (x_n) to be

$$\limsup_{n \to \infty} (x_n) = \sup E, \text{ and, } \liminf_{n \to \infty} (x_n) = \inf E;$$

We sometimes call these <u>lim sup/lim inf</u> of a sequence.

Remark: Equivalently, $\limsup_{n \to \infty} (x_n) = \lim_{n \to \infty} \left(\sup_{m \geq n} \{x_m\} \right)$, $\liminf_{n \to \infty} (x_n) = \lim_{n \to \infty} \left(\inf_{m \geq n} \{x_m\} \right)$.

Visually,

Theorem: Let $(x_n) \subset \mathbb{R}$ be a sequence and E the set of all subsequential limits (including $\pm\infty$) of (x_n).

① Then $\limsup_{n\to\infty}(x_n), \liminf_{n\to\infty}(x_n) \in E$; namely the limsup and liminf of a sequence are subsequential limits.

② If we have $\limsup_{n\to\infty}(x_n) < x$ then there is some $N \in \mathbb{N}$ such that $x_n < x$ for all $n \geq N$ (and similarly for $\liminf_{n\to\infty}(x_n) > x$).

Moreover, limsup/liminf are the only numbers satisfying ①+②.

Proof: For ① we consider cases for limsups (liminfs follow similarly):

Case 1: $\limsup_{n\to\infty}(x_n) = +\infty$, thus the set E is not bounded above and hence (x_n) is not bounded above; we can then find $x_{n_k} \to +\infty$ so $\limsup_{n\to\infty}(x_n) \in E$.

Case 2: $\limsup_{n\to\infty}(x_n) = -\infty$, then if $M \in \mathbb{R}$ there is some $N \in \mathbb{N}$ such that $x_n < M$ for all $n \geq N$ (else we could find $x_{n_k} \to x$ with $x \geq M$ by the previous theorem); thus $x_n \to -\infty$ so $\limsup_{n\to\infty}(x_n) \in E$.

Case 3: $\limsup_{n\to\infty}(x_n) \in \mathbb{R}$, thus E is bounded above and we saw that E is closed. As $E \subset \mathbb{R}$ we know that $\sup E \in \overline{E}$, but $\overline{E} = E$ so $\limsup_{n\to\infty}(x_n) \in E$.

This proves ①.

For ②, if $x_n \geq x$ for infinitely many $n \in \mathbb{N}$ then there is a sequence $(n_k) \subset \mathbb{N}$ such that $x_{n_k} \geq x$ for $k \geq 1$. Any subsequence of (x_{n_k}) is a subsequence of (x_n) also, but any subsequential limit of (x_{n_k}) is $\geq x$; hence we have $\limsup_{n\to\infty}(x_n) \geq x$, a contradiction.

If $y \in \mathbb{R} \cup \{\pm\infty\}$ satisfies ①+② also, then $y \in E$ so we have $y \leq \limsup_{n\to\infty}(x_n) = \sup E$ so if $y < x < \limsup_{n\to\infty}(x_n)$ then ② $\Rightarrow x_n < x < \limsup_{n\to\infty}(x_n)$ eventually in n, a contradiction to ①. ☒

Remarks: • $\limsup_{n\to\infty}(-x_n) = -\liminf_{n\to\infty}(x_n)$ $(-(\pm\infty) = \mp\infty)$; we use this in the previous proof to only work with limsup.

• If $(x_n),(y_n)\subset\mathbb{R}$ are sequences with $x_n \le y_n$ for $n\ge 1$ then both $\begin{cases}\liminf_{n\to\infty}(x_n) \le \liminf_{n\to\infty}(y_n), \\ \limsup_{n\to\infty}(x_n) \le \limsup_{n\to\infty}(y_n).\end{cases}$ (Check!)

• $x_n \to x \iff x = \limsup_{n\to\infty}(x_n) = \liminf_{n\to\infty}(x_n)$

Examples: • Let $\mathbb{Q} = \{x_n\}_{n\ge1}$, then $\begin{cases}\limsup(x_n) = +\infty, \\ \liminf(x_n) = -\infty.\end{cases}$

• If $(y_n) = \left((-1)^n\left(1+\frac{1}{n}\right)\right)$, then $\begin{cases}\limsup(x_n) = 1, \\ \liminf(x_n) = -1.\end{cases}$

$\boxed{\text{End L13}}$

Series

We use techniques from sequences to make sense of infinite sums:

Definition: Given a sequence $(x_n)\subset\mathbb{C}$ we write the sum of the first n terms or the <u>nth partial sum</u> as

$$S_n = \sum_{k=1}^{n} x_k = x_1 + \dots + x_n .$$

If the sequence $(S_n)\subset\mathbb{C}$ of partial sums converges we say that the <u>series</u> $\sum_{k=1}^{\infty} x_k = \sum x_k$ converges; thus if $S_n \to S$ we write $\sum_{k=1}^{\infty} x_k = S$. If $(S_n)\subset\mathbb{C}$ diverges we say that the series <u>diverges</u>.

Remarks: Sequences and series are the same concept with different notation, but it helps to take both perspectives.

$$(x_n) \text{ sequence} \longleftrightarrow \sum_{k=1}^{\infty}(x_k - x_{k-1}) \text{ Series} \quad (x_0 = 0)$$

$(S_n = x_n$ for $n \ge 2.$

By this remark we immediately have:

Theorem: $\sum x_k$ converges $\iff$ there is some $N \in \mathbb{N}$ such that $\left| \sum_{k=n}^{m} x_k \right| < \varepsilon$ for $n, m \geq N$.

Proof: This follows since convergent $\iff$ Cauchy in $\mathbb{R}^2 = \mathbb{C}$. $\blacksquare$

Remarks:
- $\sum x_k$ converges $\Rightarrow x_n \to 0$ by setting $n = m$ above!
- $x_n \to 0 \not\Rightarrow \sum x_n$ converges (as we will see).

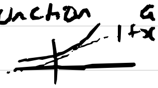
Just as for the above result, many results for series follow from their sequence counterparts. We will just summarise some interesting results and refer to Chapter 3 of Rudin for more detail.

Examples:
- If $0 \leq x < 1$ then $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$ but $\sum_{k=0}^{\infty} x^k$ diverges if $x \geq 1$; note that
$$S_n = \sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1-x} \quad \left(\begin{array}{l} S_n = 1 + \dots + x^n \\ x S_n = x + \dots + x^{n+1} \end{array} \right),$$
hence $S_n \to \frac{1}{1-x}$ since $x^{n+1} \to 0$ for $0 \leq x < 1$. Setting $x = 1$ we see that $S_n = n \to +\infty$.

- We let $e = \sum_{k=0}^{\infty} \frac{1}{k!}$ $\left(k! = (k)(k-1)(k-2)\dots(2)(1),\ 0! = 1 \right)$ which converges since (S_n) is increasing and bounded since:
$$0 \leq S_n = 1 + 1 + \frac{1}{1 \cdot 2} + \frac{1}{1 \cdot 2 \cdot 3} + \dots + \frac{1}{1 \cdot \dots \cdot n}$$
$$< 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^{n-1}} < 1 + \frac{1}{1 - \frac{1}{2}} = 3.$$

Moreover, one can show that $e = \lim_{n \to \infty} \left(\left(1 + \frac{1}{n}\right)^n \right)$.

From Calculus we know e is the unique base, $a > 1$, s.t. the function a^x has a $\underbrace{\text{tangent line}}_{\hookrightarrow (1+x)}$ of gradient $= 1$ at 0.



Theorem: e is irrational

Proof: If $e = p/q$ for $p, q \in \mathbb{N}$ then we notice that

$$e - S_n = \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \dots$$

$$\left(\sum_{k=0}^{n} \frac{1}{k!}\right) \nearrow \quad < \frac{1}{(n+1)!}\left(1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + \dots\right) = \frac{1}{(n+1)!}\left(\frac{1}{1-\frac{1}{n+1}}\right)$$

$$= \frac{1}{n!\cdot n};$$

thus $0 < q!(e - S_q) < \frac{1}{q}$. As $q!e \in \mathbb{N}$ (if $e = \frac{p}{q}$) and

$$q!S_q = q!\left(1 + 1 + \frac{1}{2} + \frac{1}{3!} + \dots + \frac{1}{q!}\right) \in \mathbb{N},$$

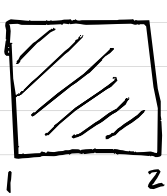
we have that $q!(e - S_q) \in \mathbb{N}$; this implies (as $q \geq 1$) there is some $k \in \mathbb{N}$ such that $0 < k < 1$, a contradiction. ■

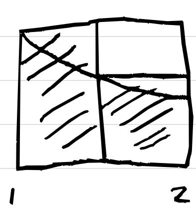
One important example of a divergent series is the <u>harmonic</u> series, $\sum \frac{1}{n}$; to see why this diverges (even though $\frac{1}{n} \to 0$) we write

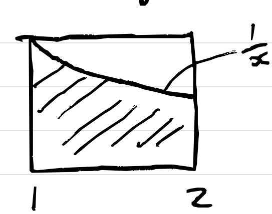
$$1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{\geq \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{1}{2}} + \dots$$

Which implies $S_{2^n} \geq 1 + \frac{n}{2}$ so $S_{2^n} \to +\infty \Rightarrow \sum \frac{1}{n}$ diverges!

We note however that $\sum \frac{(-1)^{n+1}}{n}$ does converge! Visually

$S_1 =$ 

, $S_3 =$ 

, ..., 

(actually $\sum \frac{(-1)^{n+1}}{n} = \log(2)$!).

Notice that if we rearrange the terms of $\sum \frac{(-1)^{n+1}}{n}$ we can write

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots \text{ as } \underbrace{1 - \frac{1}{2}}_{\frac{1}{2}} - \frac{1}{4} + \underbrace{\frac{1}{3} - \frac{1}{6}}_{\frac{1}{6}} + \cdots$$

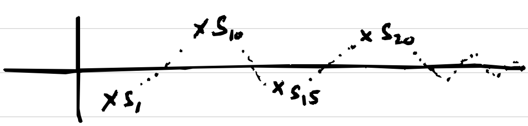
So that the rearranged sum $\frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \cdots\right) = \frac{1}{2}\log(2)$!

Thus we have rearranged the order of the sum and gotten a different value. In fact this can be done for <u>any</u> series $\sum x_n$ which converges whenever $\sum |x_n|$ diverges; known as <u>Riemann rearrangement</u>:

<u>Theorem:</u> Suppose $\sum x_n$ converges but $\sum |x_n|$ diverges, then for each $M \in \mathbb{R}$ there is some bijection $\sigma : \mathbb{N} \to \mathbb{N}$, the <u>rearrangement</u>, such that $\sum x_{\sigma(n)} = M$.

<u>Proof:</u> Let $A = \{x_n > 0\}$, $B = \{x_n < 0\}$ (i.e. the positive and negative terms in the series); we then have that $\sum_{a_n \in A} a_n$ and $\sum_{b_n \in B} (-b_n)$ are infinite or $\sum |x_n| = \sum_{a_n \in A} a_n + \sum_{b_n \in B} (-b_n) < \infty$ if both are finite, and $\sum x_n$ would diverge if only one was finite.

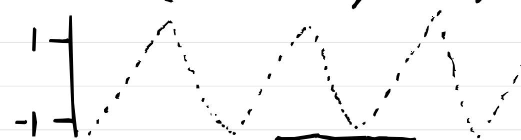
Now we define the rearrangement by taking the collection of positive terms such that their sum is $> M$, then taking the collection of negative terms so that the sum is $< M$. We then keep repeating this procedure of over and then under approximating M by alternating positive and negative terms. The deviation from M in this procedure is bounded by the terms $x_n \to 0$; hence this rearrangement of the series converges to M! $\blacksquare$



M ← (<u>Picture of Proof idea</u>)

Remark: By adapting the proof one can rearrange any such series to diverge to $\pm\infty$ or so that it does not approach any fixed limit ($\mathbb{R}$ or $\pm\infty$); visually:

and



End d14

Continuity

We now make precise the notion of continuity for functions defined between general metric spaces, capturing sensitivity/dependence of the output on the inputs; we first define limits of functions:

Definition: Let (X, d_X), (Y, d_Y) be metric spaces, $E \subset X$, with $f: E \to Y$, and z a limit point of E. We write $f(x) \to y$ as $x \to z$ or $\lim_{x \to z} f(x) = y$ if there is $y \in Y$ so that for each $\varepsilon > 0$ there is some $\delta > 0$ such that

(ε-δ)

$$d_X(x, z) < \delta \implies d_Y(f(x), y) < \varepsilon$$

Examples: • $\lim_{x \to 1} \left(\frac{1}{x}\right) = 1$.

(f need not be defined at z)

• $\lim_{x \to a} (x^2) = a^2$, • $\lim_{x \to 0} \left(\frac{1}{x}\right)$ does not exist!

Remarks: • $\lim_{x \to z} f(x) = y \iff \lim_{n \to \infty} f(x_n) = y$ for every $(x_n) \subset E$ with $x_n \neq x$ for $n \geq 1$ and $x_n \to z$.

[($\Rightarrow$) follows as $x_n \in B_\delta(z)$ eventually, ($\Leftarrow$) by contrapositive as we can build a sequence w/ $\lim_{n \to \infty} f(x_n) \neq y$ if $\lim_{x \to z} f(x) \neq y$].

• By the above, limits are unique (as we saw).

We also define algebraic operations on complex functions:

Definition: Let (X,d) be a metric space, $E \subset X$, and $f,g : X \to \mathbb{C}$. We define:

- $f+g : E \to \mathbb{C}$ by $(f+g)(x) = f(x)+g(x)$.
- $fg : E \to \mathbb{C}$ by $(fg)(x) = f(x)g(x)$.
- $f/g : E \setminus \{g=0\} \to \mathbb{C}$ by $(f/g)(x) = f(x)/g(x)$.
- $f \geq g$ if $f,g : E \to \mathbb{R}$ and $f(x) \geq g(x)$ for all $x \in E$.

Also, if $f,g : E \to \mathbb{R}^k$, $\lambda \in \mathbb{R}$ we can similarly define:

- $f+g$, $f \cdot g$, λf.

Remark: The limit laws, where well defined, also hold for functions! (e.g. $\lim_{x\to z}(f+g)(x) = A+B$ if $\lim_{x\to z} f(x) = A$ and $\lim_{x \to z} g(x) = B$).

Definition: $(\varepsilon-\delta)$ Let (X,d_X), (Y,d_Y) be metric spaces, $E \subset X$, $z \in E$, and $f : E \to Y$. We say that f is <u>continuous at z</u> if for each $\varepsilon > 0$ there is some $\delta > 0$ such that

$$d_X(x,z) < \delta \Rightarrow d_Y(f(x), f(z)) < \varepsilon.$$

(f <u>needs</u> to be defined at z)

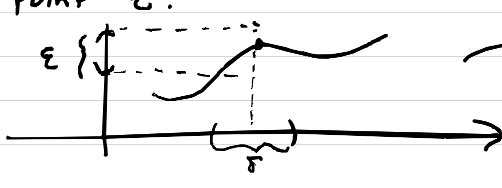
If f is continuous at every point of E, we say that f is <u>continuous on E</u>.

Remark: If z is a limit point of E then

$$f \text{ cts at } z \iff \lim_{x \to z} f(x) = f(z).$$

(continuous)

Also, if z is isolated in E then f is continuous at z since there is some neighbourhood in E only containing the point z.



(Cty for $f : \mathbb{R} \to \mathbb{R}$)

We also have another characterisation, useful in topology:

<u>Theorem:</u> Let (X, d_X), (Y, d_Y) be metric spaces, then $f: X \to Y$ is continuous $\iff$ $f^{-1}(V)$ is open in X for every $V \subset Y$ open in Y.

<u>Proof:</u>

$(\Rightarrow)$ If $f: X \to Y$ is cts and $V \subset Y$ is open in Y, we let $x \in f^{-1}(V)$ so that $f(x) \in V$. Since V is open in Y there is some $\varepsilon > 0$ such that $B_\varepsilon^Y(f(x)) \subset V$, and as f is cts there is some $\delta > 0$ so that $d_Y(f(x), f(z)) < \varepsilon$ whenever $d_X(x, z) < \delta$; hence $f(z) \in V$ so $z \in f^{-1}(V)$ and thus $B_\delta^X(x) \subset f^{-1}(V)$.

$(\Leftarrow)$ If $f^{-1}(V)$ is open in X whenever V is open in Y and $x \in X$, then for each $\varepsilon > 0$ we know that $B_\varepsilon^Y(f(x))$ is open in Y. By the assumption $f^{-1}(B_\varepsilon^Y(f(x)))$ is open in X and so there is some $\delta > 0$ such that (as $x \in f^{-1}(B_\varepsilon^Y(f(x)))$) $B_\delta(x) \subset f^{-1}(B_\varepsilon^Y(f(x)))$; hence

$$d_X(z, x) < \delta \implies d_Y(f(z), f(x)) < \varepsilon$$

so that f is cts at x. ☐

<u>Remark:</u> As U^c is closed if U is open we also have that $f: X \to Y$ is cts $\iff$ $f^{-1}(C)$ is closed in X if C is closed in Y.

Continuity also behaves well under composition and algebraic operations:

<u>Theorem:</u> Compositions of cts functions are cts.

<u>Proof:</u> Let (X, d_X), (Y, d_Y), (Z, d_Z) be metric spaces, $f: X \to Y$ and $g: f(X) \to Z$ cts and let $h = g \circ f : X \to Z$. For $x \in X$ as g is cts at $f(x) \in Y$ for every $\varepsilon > 0$ there is some $\eta > 0$ such that $d_Y(y, f(x)) < \eta \implies d_Z(g(y), g(f(x))) < \varepsilon$.

$\longrightarrow$

As f is cts at $x \in X$, for this $\varepsilon > 0$ there is some $\delta > 0$ such that $d_X(w, x) < \delta \Rightarrow d_Y(f(w), f(x)) < \varepsilon$. Thus, as

$$d_X(w,x) < \delta \Rightarrow d_Y(f(w), f(x)) < \varepsilon \Rightarrow d_Z(\overbrace{g(f(w))}^{h(w)}, \overbrace{g(f(x))}^{h(x)}) < \varepsilon;$$

and hence $h = f \circ g$ is cts. ∎

<u>Theorem:</u> ① If $f, g : X \to \mathbb{C}$ are cts then so are $f+g$, fg, $\frac{f}{g}$ $(g \neq 0)$.
② If $f : X \to \mathbb{R}^k$ is such that $f(x) = (f_1(x), \dots, f_k(x))$ for $f_1, \dots, f_k : X \to \mathbb{R}$, then f is cts $\Leftrightarrow$ $f_1, \dots, f_k$ are cts.
③ If $f, g : X \to \mathbb{R}^k$ are cts then so are $f+g$, $f \cdot g$, and λf for $\lambda \in \mathbb{R}$.

<u>proof:</u> For ①, this follows by the limit laws at limit points and the fact that functions are cts at isolated points.

For ②, we note that for each $j = 1, \dots, k$ we have that for $x, y \in X$, $|f_j(x) - f_j(y)| \leq \left(\sum_{i=1}^{k} |f_i(x) - f_i(y)|^2 \right)^{\frac{1}{2}} = |f(x) - f(y)|$; hence f cts $\Leftrightarrow$ f_j are for all $j = 1, \dots, k$ (take $\frac{\varepsilon}{\sqrt{k}}$ for $\Leftarrow$).

For ③, this follows by combining ① and ②. ∎

<u>Remarks:</u>
- We call the functions $\{f_i\}_{i=1}^{k}$ the <u>components</u> of f.
- $f+g : X \to \mathbb{R}^k$ but $f \cdot g : X \to \mathbb{R}$.

<u>Examples:</u>
- $\phi_i : \mathbb{R}^k \to \mathbb{R}$ defined by $\phi_i(x) = x_i$ is cts (the ith <u>coordinate function</u>).
- Inductively applying ① we see that every <u>polynomial</u>
$$P(x) = \sum c_{n_1, \dots, n_k} x_1^{n_1} \cdots x_k^{n_k} \text{ is cts;}$$
where $P : \mathbb{R}^k \to \mathbb{C}$, $c_{n_1, \dots, n_k} \in \mathbb{C}$ are <u>coefficients</u>, $n_1, \dots, n_k \in \mathbb{N}$, and the sum is finite. <u>Rational functions</u>, where defined, are cts (ratios of polynomials).

- $x \mapsto |x|$ is cts as $\big||x|-|y|\big| \leq |x-y|$ (saw before) and hence if $f: X \to \mathbb{R}^n$ is cts then $x \mapsto |f(x)|$ is also.

End L15

Continuity and Compactness

(i.e. cts maps preserve compactness)

Theorem: Let $f: X \to Y$ be a cts map from a compact metric space, then $f(X)$ is compact.

Proof: Let $\{U_\alpha\}_\alpha$ be an open cover of $f(X)$, as f is cts $\{f^{-1}(U_\alpha)\}_\alpha$ is an open cover of X. As X is compact there exists a finite subcover $\{f^{-1}(U_{\alpha_i})\}_{i=1}^n$ such that

$$X \subset \bigcup_{i=1}^n f^{-1}(U_{\alpha_i}).$$

As $f(f^{-1}(U_{\alpha_i})) \subset U_{\alpha_i}$ for each $i=1,\dots,n$ we have

$$f(X) \subset f\Big(\bigcup_{i=1}^n f^{-1}(U_{\alpha_i})\Big) = \bigcup_{i=1}^n f(f^{-1}(U_{\alpha_i})) \subset \bigcup_{i=1}^n U_{\alpha_i};$$

thus $\{U_{\alpha_i}\}_{i=1}^n$ is a finite subcover for $f(X)$. ☒

Remark: We use the fact that the image of the union is the union of the image in the above proof (which is not too tricky to check directly).

Theorem: Let $f: X \to Y$ be a cts bijection from a compact metric space, then $f^{-1}: Y \to X$ (defined by $f^{-1}(f(x)) = x$) is cts.

Proof: Let $U \subset X$ be open, we will show that $f(U)$ is open in Y which shows that f^{-1} is cts as $f^{-1}(f(U)) = U$. We see that

$$U \text{ open} \Leftrightarrow U^c \text{ closed} \underset{\text{as } X \text{ is}}{\Rightarrow} U^c \text{ compact} \underset{\text{as } f \text{ cts}}{\Rightarrow} f(U^c) \text{ compact} \Rightarrow f(U^c) \text{ closed}.$$

As f is a bijection we have that

$$Y \setminus f(U^c) = f(U) \text{ so that } f(U^c)^c = f(U)$$

and hence as $f(U^c)$ is closed, $f(U^c)^c = f(U)$ is open. ☒

Remark: Both theorems fail if X is not compact, for example $f: \mathbb{R} \to \mathbb{R}$ given by $f(x) = x$ shows $f(\mathbb{R})$ not compact and $g: [0, 2\pi) \to S^1 \subset \mathbb{C}$ given by $g(\theta) = e^{2\pi i \theta}$ shows that g^{-1} is defined but not cts.

In Euclidean space we have:

Theorem: If $f: X \to \mathbb{R}^n$ is a cts map from a compact metric space, then $f(X)$ is closed and bounded.

Proof: By the first theorem above, $f(X)$ is compact in $\mathbb{R}^n$, hence closed and bounded by the Heine-Borel theorem. $\blacksquare$

Specifically in $\mathbb{R}$ we have the so called Extreme Value theorem:

Theorem: If $f: X \to \mathbb{R}$ is a cts map from a compact metric space, then there are $x, y \in X$ such that
$$f(x) = \sup f(X) \text{ and } f(y) = \inf f(X).$$

Proof: We may assume $X \neq \emptyset$ (or there is nothing to show). By the last result, $f(X)$ is bounded and so by the least upper bound property for $\mathbb{R}$, both $\sup f(X)$ and $\inf f(X)$ exist. Moreover, as $f(X)$ is closed we have $\overline{f(X)} = f(X)$ and so $\sup f(X), \inf f(X) \in \overline{f(X)} = f(X)$. $\blacksquare$

Remark: This is equivalent to saying that there are $x, y \in X$ such that $f(y) \leq f(z) \leq f(x)$ for all $z \in X$; i.e. f attains its maximum at x and minimum at y.

Uniform Continuity

Notice in the definition of continuity that it was specified at a given point; namely for $\varepsilon > 0$ there was a $\delta > 0$ depending on the point chosen so that the definition held. If we can choose one $\delta > 0$ that works for all points we have:

Definition: Let (X, d_X), (Y, d_Y) be metric spaces and $f: X \to Y$. We say that f is <u>Uniformly Continuous</u> on X if for each $\varepsilon > 0$ there is $\delta > 0$ such that

$$d_X(x, y) < \delta \Rightarrow d_Y(f(x), f(y)) < \varepsilon.$$

Remark: Uniform cty $\Rightarrow$ cty, cty $\not\Rightarrow$ uniform cty (e.g. x^2).

Example: $f: \mathbb{R} \to \mathbb{R}$ with $f(x) = ax + b$ (linear fn) is uniformly cts on $\mathbb{R}$; to see this, if $\varepsilon > 0$ then setting $\delta = \varepsilon/(|a|+1)$

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| = |a||x - y| < \frac{|a|\varepsilon}{|a|+1} < \varepsilon.$$

Theorem: If $f: X \to Y$ is cts and X is compact, then f is Uniformly cts.

Proof: For $\varepsilon > 0$, and since f is cts, if $x \in X$ we can find $\delta_x > 0$ such that

$$d_X(x, y) < \delta_x \Rightarrow d_Y(f(x), f(y)) < \frac{\varepsilon}{2}.$$

We then have an open cover, $\{B_{\delta_x}(x)\}_{x \in X}$, of X; as X is compact there is a finite subcover, $\{B_{\delta_{x_i}}(x_i)\}_{i=1}^{n}$, for some $\{x_i\}_{i=1}^{n} \subset X$. Set

$$\delta = \frac{1}{2}\min\{\delta_{x_1}, \dots, \delta_{x_n}\} > 0$$

So that if $x, y \in X$ then $y \in B_{\delta_{x_i}}(x_i)$ for some $i = 1, \dots, n$, hence

$$d_X(x, y) < \delta \Rightarrow d_X(x, x_i) \leq d_X(x, y) + d_X(y, x_i) < \delta + \frac{\delta_{x_i}}{2} < \delta_{x_i},$$

and so

$$d_Y(f(x), f(y)) \leq d_Y(f(x), f(x_i)) + d_Y(f(y), f(x_i)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus f is uniformly cts on X. ∎

Remark: Can also prove this by contradiction (<u>check</u>).

We finally emphasise why compactness is essential in these results:

Theorem: Let $E \subset \mathbb{R}$ be noncompact, then:

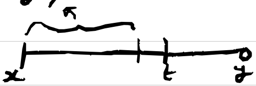
1. there is an unbounded cts function on E.
2. there is a cts bounded function on E with no maximum.
3. if E is bounded there is a cts function on E which is not uniformly cts.

Proof: We first assume E is bounded and prove ① → ③. For ①, there must exist a limit point y of E which is not in E. Define $f: E \to \mathbb{R}$ by $f(x) = \frac{1}{x-y}$ which is cts but not bounded.

For ③, f defined above is not uniformly continuous since for any $\delta > 0$ and $x \in E$ with $|x-y| < \delta$ we have

$$|f(x) - f(t)| = \left| \frac{1}{x-y} - \frac{1}{t-y} \right| = \left| \frac{t-x}{(x-y)(t-y)} \right| > \frac{1}{2|t-y|} > 1$$

by choosing $|t-y|$ small and $|t-x| > |x-y|/2$.

↳ (Visually, )

For ②, any y as above define $g: E \to \mathbb{R}$ by

$$g(x) = \frac{1}{1+(x-y)^2}$$

which is cts on E and bounded. We see that $\sup_{x \in E}(g(x)) = 1$ but $g(x) < 1$ if $x \in E$ so that g has no maximum!

Now if E is unbounded then $f(x) = x \Rightarrow$ ① and $h(x) = \frac{x^2}{1+x^2} \Rightarrow$ ② since $\sup_{x \in E}(h(x)) = 1$ but $h(x) < 1$ for any $x \in E$. ☒

Remark: ③ fails if E is unbounded by considering any function on $\mathbb{N}/\mathbb{Z}$ (every such function is uniformly cts!) End 2.16

Continuity and Connectedness

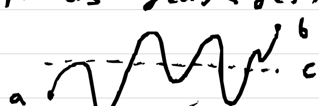
Theorem: If $f: X \to Y$ is cts and X is connected, then $f(X)$ is connected.

Proof: If $f(X)$ is not connected then there are disjoint open nonempty sets, $A, B \subset Y$, such that $f(X) \subset A \cup B$. As f is cts, both $f^{-1}(A), f^{-1}(B)$ are open nonempty sets in X; moreover they are disjoint as A and B are disjoint. If $x \in X$ then $f(x) \in f(X) = A \cup B$ and hence $x \in f^{-1}(A) \cup f^{-1}(B)$, but this implies $X \subset f^{-1}(A) \cup f^{-1}(B)$, a contradiction. ▨

This allows us to prove the <u>intermediate value theorem</u>:

Theorem: If $f: [a, b] \to \mathbb{R}$ is cts and $c \in (f(a), f(b))$ (or $(f(b), f(a))$) then there is $x \in (a, b)$ such that $f(x) = c$.

Proof: By the last result $f([a,b])$ is connected as $[a,b]$ is. As $c \in (f(a), f(b))$ (w.l.o.g.) and $f(a) < c < f(b)$ the characterisation for connected sets in $\mathbb{R} \Rightarrow c \in f([a,b])$; hence there is some $x \in [a,b]$ such that $c = f(x)$. Finally we note that $x \neq a \neq b$ as $f(a) < f(x) < f(b)$. ▨

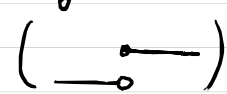
(Visually: 

) (Cts fns drawn without lifting pencil/chalk)

Discontinuities

We will see that the converse of the IVT does not hold; i.e. there are functions taking every value between two given numbers that fail to be cts:

Definition: If $f: X \to Y$ is not continuous at $x \in X$ we say that f has a <u>discontinuity</u> at x.

()

Example: $H(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geq 0 \end{cases}$, has a discontinuity at 0.
(Heaviside fn)

The failure of cty here can be classified:

<u>Definition:</u> Let $f:(a,b)\to Y$ and $a\leq x\leq b$ we write
$$f(x^+)=\lim_{t\to x^+} f(t) \quad\text{and}\quad f(x^-)=\lim_{t\to x^-} f(t)$$
for the <u>right and left hand limits of f at x</u> respectively, if

$f(t_n)\to f(x^+)$ for all sequences $(t_n)\subset(x,b)$, $t_n\to x$,

$f(t_n)\to f(x^-)$ for all sequences $(t_n)\subset(a,x)$, $t_n\to x$.

<u>Remark:</u> $\lim_{t\to x} f(t)$ exists $\iff f(x^+)=f(x^-)=\lim_{t\to x} f(t)$.

<u>Example:</u> $H(0^+)=1$, $H(0^-)=0$ so $\lim_{t\to 0} H(t)$ DNE.

<u>Definition:</u> If $f:(a,b)\to Y$ is discts at $x\in[a,b]$ then:
- x is a discty of the <u>first kind</u> if $f(x^+)$, $f(x^-)$ exist.
- x is a discty of the <u>second kind</u> otherwise.

<u>Examples:</u>
- 0 is a discty for H of the first kind as $H(0^+)\neq H(0^-)$.
- $f(x)=\begin{cases}0, & \text{if } x\neq 0\\ 1, & \text{if } x=0\end{cases}$  has a discty of the first kind at 0 even though $f(0^+)=f(0^-)$ but $\neq f(0)=1$.
- $g(x)=\begin{cases}0, & \text{if } x\in\mathbb{Q}\\ 1, & \text{if } x\notin\mathbb{Q}\end{cases}$ has a discty of the second kind at every $x\in\mathbb{R}$ (as $g(x^\pm)$ DNE).
- $h(x)=\begin{cases}x, & x\in\mathbb{Q}\\ 0, & x\notin\mathbb{Q}\end{cases}$ <u>is</u> cts at 0 (check) but has a discty of second kind at every $x\neq 0$.

- Assuming $\sin(x)$ is well defined and cts we see that $\sin(\frac{1}{x})$ is cts on $\{x \neq 0\}$. Defining

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

then $f(0^{\pm})$ DNE so 0 is a discty of the second kind. Note that f attains every value in $[-1, 1]$ however, so the converse of IVT fails!

Monotone functions

For our final topic in cty we study functions that never decrease or increase on $\mathbb{R}$:

Definition: We say that $f: (a,b) \to \mathbb{R}$ is monotone if it is either,
- increasing, i.e. $f(x) \leq f(y)$ for $a < x \leq y < b$.
- decreasing, i.e. $f(x) \geq f(y)$ for $a < x \leq y < b$.

Remark: If f is both increasing and decreasing $\Rightarrow$ f is constant.

Theorem: Let $f: (a,b) \to \mathbb{R}$ be monotone increasing, then $f(x^+)$ and $f(x^-)$ exist for every $x \in (a,b)$. Moreover:
- $\sup_{a<t<x} f(t) = f(x^-) \leq f(x) \leq f(x^+) = \inf_{x<t<b} f(t)$.
- $f(x^+) \leq f(y^-)$ for $a < x < y < b$.

Proof: Note that $E = \{f(t) \mid a < t < x\}$ is bounded above by $f(x)$ and hence $\sup_{a<t<x} f(t)$ exists. For each $\varepsilon > 0$ there must exist some $\delta > 0$ such that

$$\sup_{a<t<x} f(t) - \varepsilon < f(x-\delta) < \sup_{a<t<x} f(t)$$

for $a < x-\delta < x$ (or would not be the sup). As f is monotone increasing, if $x-\delta < y < x$ then $f(x-\delta) \leq f(y) \leq \sup_{a<t<x} f(t)$ so that

$$\sup_{a<t<x} f(t) - \varepsilon < f(y) < \sup_{a<t<x} f(t) \Rightarrow \sup_{a<t<x} f(t) = f(x^-).$$

Similarly we can see that $\inf_{x<t<b} f(t) = f(x^+)$; thus if $a<x<y<b$ we have

$$f(x^+) = \inf_{x<t<b} f(t) \leq \inf_{x<t<y} f(t) \leq \sup_{x<t<y} f(t) \leq \sup_{a<t<y} f(t) = f(y^-).$$ ☐

Remark: Similar results hold for monotone decreasing functions; thus monotone functions only have discty of the first kind.

Theorem: Monotone functions have at most countably many discontinuities.

Proof: Without loss of generality let $f: (a,b) \to \mathbb{R}$ be monotone increasing and let $E = \{\text{discty pts of } f\}$. If $x \in E$ then $f(x^-) \neq f(x^+)$ and so there is some $r_x \in \mathbb{Q}$ such that $f(x^-) < r_x < f(x^+)$. Now if $x \neq y$ and $y \in E$ then by the previous result:

$$f(x^-) \leq r_x < f(x^+) \leq f(y^-) < r_y < f(y^+),$$

and thus $g: E \to \mathbb{Q}$ defined by $g(x) = r_x$ is injective. We thus see that $g: E \to g(E)$ is a bijection from E to a subset of $\mathbb{Q}$, hence E is at most countable. ☐

We finish with a construction of a monotone function with prescribed discontinuities: let $E \subset (a,b)$ be countable (e.g. $\mathbb{Q} \cap (a,b)$) and $(x_n) = E$ be an enumeration; then, for any sequence, $(a_n) \subset \mathbb{R}$, of positive numbers with $\sum a_n$ converging let $f: (a,b) \to \mathbb{R}$ be

$$f(x) = \sum_{\{n \mid x_n < x\}} a_n.$$

One can check that f satisfies:
- f is monotone increasing
- $f(x_n^+) - f(x_n^-) = a_n > 0$
- f cts on $(a,b) \setminus E$.

In fact, $f(x^-) = f(x)$ so that f is cts from the left. If we sum over $x_n \leq x$ instead we get $f(x^+) = f(x)$ i.e. cts from the right.

End of 17

Normed Vector Spaces

Notice that $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{R}^n$ have the structure of vector spaces in addition to being metric spaces. Spaces of this type are of great use in studying PDEs/functional analysis:

Definition: A norm on a vector space, V, over $\mathbb{C}$ is a map $\|\cdot\|: V \to [0,\infty)$ satisfying:

- $\|x\| = 0 \iff x = 0$.
- $\|\lambda x\| = |\lambda|\|x\|$ for any $\lambda \in \mathbb{C}$ and $x \in V$.
- $\|x+y\| \leq \|x\| + \|y\|$ for any $x, y \in V$.

We then call $(V, \|\cdot\|)$ a normed vector space.

Examples: • $(\mathbb{R}, |\cdot|)$, $(\mathbb{Q}, |\cdot|)$, $(\mathbb{R}^n, |\cdot|)$ are all normed vector spaces

• Let $S = \{\text{sequences } (x_n) \subset \mathbb{R}\}$. We define norms,

(Δ ineq. for $1 \leq p < \infty$ is called Minkowski's ineq.)

$$\begin{cases} \|(x_n)\|_p = \left(\sum |x_n|^p \right)^{\frac{1}{p}} & \text{for } 1 \leq p < \infty, \\ \|(x_n)\|_\infty = \sup\limits_{n \in \mathbb{N}} |x_n|. \end{cases}$$

We then let

(So called little ℓ^p/sequence spaces)

$$\begin{cases} \ell^p = \{(x_n) \in S \mid \|(x_n)\|_p < \infty\} & \text{for } 1 \leq p < \infty, \\ \ell^\infty = \{(x_n) \in S \mid \|(x_n)\|_\infty < \infty\}; \end{cases}$$

which define normed vector spaces.

• If X is a metric space we let

$$\mathcal{C}(X) = \{ f: X \to \mathbb{C} \mid f \text{ is cts and bounded on } X \}$$

and define the norm $\|\cdot\|: \mathcal{C}(X) \to [0,\infty)$ by setting

$$\|f\| = \sup_{x \in X} |f(x)| \text{ for each } f \in \mathcal{C}(X);$$

(Supremum norm)

the EVT shows that $\|f\| < \infty$ for any cts fn if X is compact.

↳ (We will work frequently with this space!)

↳ (induced metric is $d(f,g) = \|f-g\|$ for $f, g \in \mathcal{C}(X)$)

Remark: Any norm defines a metric since we can let $d: V \times V \to [0,\infty)$ by setting
$$d(x,y) = \|x-y\| \text{ for } x,y \in V.$$
We call this the <u>induced metric</u> on V. If V is complete with respect to its induced metric we say that V is a <u>Banach space</u>.

Example: $(\mathbb{R}, |\cdot|)$ is a Banach space, $(\mathbb{Q}, |\cdot|)$ is not.

Definition: Let $(x_n) \subset V$ be a sequence in a normed vector space, the series $\sum x_n$ is said to <u>converge</u> in V if the sequence (S_N) of partial sums, $S_N = \sum_{i=1}^{N} x_i$, converges to some $x \in X$; i.e. if we have
$$\lim_{N\to\infty} \|x - \sum_{i=1}^{N} x_i\| = 0.$$

Theorem: $(V, \|\cdot\|)$ is a Banach space $\iff$ $\sum x_n$ converges in V whenever $\sum \|x_n\|$ converges in $\mathbb{R}$.

Proof: $(\Rightarrow)$ If $(V, \|\cdot\|)$ is Banach and $\sum \|x_n\|$ converges, then for $\varepsilon > 0$ by setting
$$S_N = \sum_{i=1}^{N} x_i, \quad T_N = \sum_{i=1}^{N} \|x_i\|$$
we have that $(T_N) \subset \mathbb{R}$ is Cauchy and so
$$\|S_m - S_n\| = \left\| \sum_{i=n+1}^{m} s_i \right\| \leq \sum_{i=n+1}^{m} \|x_i\| = |T_m - T_n| < \varepsilon$$
for $m > n$ sufficiently large; hence $(S_N) \subset V$ is Cauchy and thus convergent as V is complete.

$(\Leftarrow)$ If $\sum x_n$ converges whenever $\sum \|x_n\|$ does, let $(y_n) \subset V$ be a Cauchy sequence, so that for each $k \geq 1$ there is an $N_k \in \mathbb{N}$ such that $\|y_m - y_n\| < 2^{-k}$ if $m,n \geq N_k$. Setting $x_1 = y_{N_1}$ and for $j > 1$, $x_j = y_{N_j} - y_{N_{j-1}}$ we have both
$$\|x_j\| = \|y_{N_j} - y_{N_{j-1}}\| < 2^{-j+1} \text{ and } S_k = \sum_{i=1}^{k} x_i = y_{N_k}.$$
$\rightarrow$

$(\Leftarrow)$ Now we note that for each $k \geq 1$ we have
$$\sum_{i=1}^{k} \|x_i\| \leq \|x_1\| + \sum_{i=2}^{k} 2^{-i+1} < \|x_1\| + \sum \tfrac{1}{2^i} = \|x_1\| + 1;$$
hence $\left(\sum_{i=1}^{k} \|x_i\|\right) \subset \mathbb{R}$ is bounded and increasing, thus converges. By the assumption $\Rightarrow \sum x_i$ converges, but then $S_k = \sum_{i=1}^{k} x_i = y_{N_k}$ converges to some $y \in V$. Thus (y_n) is a cauchy sequence in V with a convergent subsequence, an $\varepsilon/2$ argument shows that (y_n) must converge also; hence V is a Banach space. ∎

We know that any finite dimensional vector space is isomorphic to some Euclidean space, we want to see if the norm structure is also preserved:

<u>Definition:</u> Two norms, $\|\cdot\|_1$ and $\|\cdot\|_2$, on a vector space are said to be <u>equivalent</u> if there exist $A, B > 0$ such that
$$A\|x\|_1 \leq \|x\|_2 \leq B\|x\|_1 \text{ for all } x \in V.$$

<u>Theorem:</u> All norms on finite dimensional vector spaces are equivalent.

<u>Proof:</u> We first note that equivalence of norms is an equivalence relation. Let us then choose a basis $\{e_i\}_{i=1}^{n}$ for V and define a norm
$$\left\|\sum_{j=1}^{n} a_j e_j\right\|_1 = \sum_{j=1}^{n} |a_j| \text{ for } \{a_j\}_{j=1}^{n} \subset \mathbb{R};$$
it thus suffices to show any other norm, $\|\cdot\|_2$, is equivalent to this norm. Note that if $x = 0$ there is nothing to show and if $A\|u\|_1 \leq \|u\|_2 \leq B\|u\|_1$ for all $u \in V$ with $\|u\|_1 = 1$ we can set $u = \frac{x}{\|x\|_1}$ for $x \in V \setminus \{0\}$ to get $A\|x\|_1 \leq \|x\|_2 \leq B\|x\|_1$; we will thus just show $A\|u\|_1 \leq \|u\|_2 \leq B\|u\|_1$ for $u \in V$ with $\|u\|_1 = 1$.

By the reverse Δ ineq. we have for each $x, y \in V$ that
$$\big|\|x\|_2 - \|y\|_2\big| \leq \|x - y\|_2$$
and so if $x = \sum_{i=1}^{n} a_i e_i$ and $y = \sum_{i=1}^{n} b_i e_i$ then
$$\big|\|x\|_2 - \|y\|_2\big| \leq \|x - y\|_2 \leq \sum_{i=1}^{n} |a_i - b_i| \|e_i\|_2 = \|x - y\|_1 \cdot \max_{1 \leq i \leq n}\{\|e_i\|_2\}.$$

Thus for each $\varepsilon > 0$ by choosing $\delta = \dfrac{\varepsilon}{\max\limits_{1 \le i \le n}\{\|e_i\|_2\}}$ we see that

$$\|x-y\|_1 < \delta \implies \big|\|x\|_2 - \|y\|_2\big| < \varepsilon;$$

hence the map $\|\cdot\|_2 : (V, \|\cdot\|_1) \to [0,\infty)$ is cts.

Finally we will show that $S = \{x \in V \mid \|x\|_1 = 1\}$ is compact so that by the EVT its image under the cts map $\|.\|_2$ is bounded, implying $A\|u\|_1 \le \|u\|_2 \le B\|u\|_1$ for each $u \in S$. By considering

$$T = \{(a_1, \dots, a_n) \in \mathbb{R}^n \mid \sum_{i=1}^{n} |a_i| = 1\}$$

and the map $f : T \to S$ defined by $f(a_1, \dots, a_n) = \sum_{i=1}^{n} a_i e_i$, since T is closed and bounded (check) and thus compact by the Heine-Borel theorem, it suffices to check that f is cts. For $\varepsilon > 0$ we let $\delta = \frac{\varepsilon}{\sqrt{n}}$ and note that

$$|(a_1, \dots, a_n) - (b_1, \dots, b_n)| < \frac{\varepsilon}{\sqrt{n}} \implies \left\|\sum_{i=1}^{n} (a_i - b_i)e_i\right\|_1 \le \left(\sum_{i=1}^{n} |a_i - b_i|^2\right)^{\frac{1}{2}} = \varepsilon,$$

so that f is cts and hence S is compact. $\boxed{}$

<u>Remarks:</u>

- This theorem is saying that all finite dimensional normed vector spaces are the 'same' as $\mathbb{R}^n$ with the usual absolute value. Hence the only 'interesting' normed vector spaces are those of infinite dimension!

- Our proof hinged on the fact that the unit sphere was compact in a finite dimensional vector space. In fact, one can show that the unit sphere is compact if and only if the space is of finite dimension. Moreover there exist norms which are <u>not</u> equivalent on infinite dimensional spaces! For more, see functional analysis.

$\boxed{\text{End L18}}$

Differentiation

We are now able to recover familiar theory from Calculus for One-Variable functions using the precise formulation of limits:

Definition: We say that $f:[a,b]\to\mathbb{R}$ is differentiable at $x\in[a,b]$ if
$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=\lim_{t\to x}\frac{f(t)-f(x)}{t-x} \text{ exists,}$$
and we call $f'(x)$ the derivative of f at x. If f is differentiable at every $x\in E\subset[a,b]$ we say that f is differentiable on E, and let $f':E\to\mathbb{R}$ be the derivative as a function.

With this we recover all of the main results from Calculus I:

Theorem: If $f:[a,b]\to\mathbb{R}$ is differentiable at $x\in[a,b]$ then it is cts at x.

Proof: If $t\neq x$ then $f(t)-f(x)=\dfrac{f(t)-f(x)}{(t-x)}(t-x)\to f'(x)\cdot 0=0$ as $t\to x$; hence f is cts at x. $\blacksquare$

Remark: $|x|$ is cts at 0 but not differentiable at 0.

Theorem: If $f,g:[a,b]\to\mathbb{R}$ are diff'ble at x then:
1. $(f+g)'(x)=f'(x)+g'(x)$.
2. $(fg)'(x)=f'(x)g(x)+f(x)g'(x)$. (Product rule)

Proof: ① follows from the limit laws.
② We note that
$$f(t)g(t)-f(x)g(x)=f(t)\big(g(t)-g(x)\big)+g(x)\big(f(t)-f(x)\big)$$
So that dividing by $(t-x)$ and letting $t\to x$ we are done. $\blacksquare$

Examples:
- Using ②, $(x^n)'=nx^{n-1}$ for $n\in\mathbb{Z}$ (with $x\neq 0$ for $n\leq 0$), hence polynomials, rational fns are diff'ble.

Theorem: If $f:[a,b]\to\mathbb{R}$ is cts, $f'(x)$ exists for some $x\in[a,b]$ and $g:I\to\mathbb{R}$ is diff'ble at $f(x)\in I$ then

$$(g\circ f)'(x) = f'(x)\cdot g'(f(x)). \quad (\text{chain rule})$$

Proof: By definition there are $u(t), v(s)\to 0$ as $t\to x$, $s\to f(x)$ so that

$$\begin{cases} f(t)-f(x) = (t-x)(f'(x)+u(t)), \\ g(s)-g(f(x)) = (s-f(x))(g'(f(x))+v(s)), \end{cases}$$

and thus setting $s=f(t)$ we have

$$\begin{aligned} g(f(t))-g(f(x)) &= (f(t)-f(x))(g'(f(x))+v(f(t))) \\ &= (t-x)(f'(x)+u(t))(g'(f(x))+v(f(t))). \end{aligned}$$

Dividing both sides by $(t-x)$ and letting $t\to x$ we are done. ∎

Examples: • $\left(\frac{f}{g}\right)' = (f\cdot(g)^{-1})' = f'(g)^{-1} + f(-1(g)^{-2})g'$

$$= \frac{f'g-fg'}{g^2} \quad (\underline{\text{Quotient rule}})$$

• Let

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

then for $x\neq 0$ we have

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

At $x=0$ we see that if $t\neq 0$ then

$$\left|\frac{f(t)-f(0)}{t}\right| = \left|t\sin\left(\tfrac{1}{t}\right)\right| \leq |t| \Rightarrow f'(0)=0;$$

hence f is diff'ble on $\mathbb{R}$! Note however that f' is <u>not</u> cts as

$$\lim_{x\to 0}\left(\cos\left(\tfrac{1}{x}\right)\right) \text{ DNE!}$$

(in later notation this shows that f' exists but $f\notin C'$).

Definition: We say that $f : X \to \mathbb{R}$ has a local maximum at $x \in X$ if there is some $\delta > 0$ such that
$$d(x, y) < \delta \Rightarrow f(y) \leq f(x).$$
Minima are defined analogously. Extrema = max/minima.

Theorem: If $f : [a,b] \to \mathbb{R}$ has a local extrema at $x \in (x, b)$ then if $f'(x)$ exists, $f'(x) = 0$.

(Fermat)

Proof: Suppose (w.l.o.g.) x is a local maximum, thus there is some $\delta > 0$ such that $y \in (x - \delta, x + \delta) \Rightarrow f(y) \leq f(x)$. Hence,
$$\begin{cases} \dfrac{f(t) - f(x)}{t - x} \geq 0 & \text{if } t \in (x - \delta, x), \\ \dfrac{f(t) - f(x)}{t - x} \leq 0 & \text{if } t \in (x, x + \delta), \end{cases}$$
and so $f'(x)$, if it exists, must equal zero. ∎

Theorem: If $f, g : [a,b] \to \mathbb{R}$ are cts functions which are diffble on (a,b), then there is some $x \in (a,b)$ such that
$$(f(b) - f(a))\, g'(x) = (g(b) - g(a))\, f'(x).$$

In particular, if $g(x) = x$ then
$$f'(x) = \frac{f(b) - f(a)}{b - a} \quad (\text{Mean value theorem}).$$

Proof: Let $h(x) = (f(b) - f(a))\, g(x) - (g(b) - g(a))\, f(x)$ so that h is cts on $[a,b]$ and diffble on (a,b). If h is constant then $h'(x) = 0$ for all $x \in (a,b)$, so we're done. If h is not constant, then as $h(a) = h(b) = f(b)g(a) - g(b)f(a)$ there must be some local extrema $x \in (a,b)$ so that by the previous result $h'(x) = 0$. ∎

Remark: This shows that
$$\begin{cases} f' \geq 0 \Rightarrow f \text{ increasing.} \\ f' = 0 \Rightarrow f \text{ constant.} \\ f' \leq 0 \Rightarrow f \text{ decreasing.} \end{cases}$$

Theorem: Let $f, g : (a,b) \to \mathbb{R}$ be diff'ble, $(a, b \in \mathbb{R} \cup \{\pm\infty\})$ $g(x) \neq 0$ for $x \in (a,b)$.

(L'Hospital's rule) If $\frac{f'(x)}{g'(x)} \to A \in \mathbb{R} \cup \{\pm\infty\}$ as $x \to a$ or $x \to b$, and either $f(x), g(x) \to 0$ or $g(x) \to \pm\infty$ as $x \to a$, then

$$\frac{f(x)}{g(x)} \to A \in \mathbb{R} \cup \{\pm\infty\} \text{ as } x \to a \text{ or } x \to b.$$

Proof: We consider the cases $A \neq +\infty$ and $x \to a$ first; the case $x \to b$ is easy to adapt. Fix some $B \in \mathbb{R}$ with $A < B$ and choose some $r \in (A, B)$. We then have that for some $c \in (a, b)$

$$x \in (a, c) \implies \frac{f'(x)}{g'(x)} < r;$$

hence if $a < x < y < c$ the MVT implies there is $t \in (x, y)$ with

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < r. \quad (+)$$

If $f(x), g(x) \to 0$ as $x \to a \implies \frac{f(y)}{g(y)} < r < B$ for $y \in (a, c)$.
Similarly, if $g(x) \to +\infty$ ($-\infty$ similar) then for smaller choice of $x \in (a, c)$ we have $g(x) > 0$ and $g(x) > g(y)$. Multiply $(+)$ by $(g(x) - g(y))/g(x)$ to get

$$\frac{f(x) - f(y)}{g(x)} < r\left(\frac{g(x) - g(y)}{g(x)}\right) \implies \frac{f(x)}{g(x)} < B$$

for $x \in (a, c)$ sufficiently close to a. Thus we have, in either case, that $\frac{f(x)}{g(x)} < B$ whenever $x \in (a, \hat{c})$ for some $\hat{c}$.

If $A = -\infty$ we are then done. If $-\infty < A \leq +\infty$ then we can similarly find $\check{\hat{c}}$ such that if $\breve{B} \in \mathbb{R}$ with $\breve{B} < A$

then $\frac{f(x)}{g(x)} > \breve{B}$ for $x \in (a, \check{\hat{c}})$;

the remaining cases then follow. $\square$

(to prove the $x \to b$ case we just consider (c, b).)

$\boxed{\text{End 219}}$

Continuity of Derivatives

We saw already that f being differentiable does not imply that f' is cts. However we can see that f' satisfies the conclusions of the IVT:

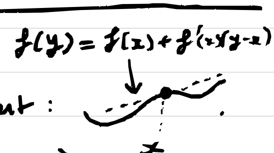
Theorem: (Darboux's theorem) If $f:[a,b]\to\mathbb{R}$ is differentiable and $c\in(f'(a),f'(b))$ (or $(f'(b),f'(a))$) then there is some $x\in(a,b)$ such that $f'(x)=c$.

Proof: Let $g(x)=f(x)-ct$ so that both $\begin{cases} g'(a)=f'(a)-c<0, \\ g'(b)=f'(b)-c>0, \end{cases}$ and thus a,b are not extrema for g. Hence, by the EVT g has a local extrema $x\in(a,b)$ which by Fermat's theorem $\Rightarrow$ $g'(x)=0 \iff f'(x)=c$. $\blacksquare$

Remark: This tells us that the derivative of a function cannot have any discontinuities of the first kind (i.e. no jumps).

Polynomial Approximation

We can use MVT to approximate a function by its tangent:



More generally:

Theorem: Let $f:[a,b]\to\mathbb{R}$, $n\geq 1$, $f^{(n-1)}$ ((n-1)th derivative) be cts on $[a,b]$, and $f^{(n)}$ defined on (a,b). Then if $\alpha,\beta\in[a,b]$ with $\alpha\neq\beta$,

$$f(\beta)=\sum_{k=0}^{n-1}\frac{f^{(k)}(\alpha)}{k!}(\beta-\alpha)^k+\frac{f^{(n)}(x)}{n!}(\beta-\alpha)^n$$

for some $x\in(\alpha,\beta)$. We call

$$P(t)=\sum_{k=0}^{n-1}\frac{f^{k}(\alpha)}{k!}(t-\alpha)^k$$

the $(n-1)$th order Taylor Polynomial of f at α.

Remark: This is saying that f can be approximated by a polynomial of degree $(n-1)$, namely $P(t)$, with the error of this approximation controlled by the nth derivative of f, $f^{(n)}$.

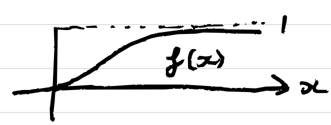
Proof: First we let $M \in \mathbb{R}$ be such that
$$f(\beta) = P(\beta) + M(\beta - \alpha)^n,$$
and define $g(t) = f(t) - P(t) - M(t-\alpha)^n$ on $[a,b]$; we will show that $M = f^{(n)}(x)/n!$ for some $x \in (\alpha, \beta)$. We compute that
$$g^{(n)}(t) = f^{(n)}(t) - \underbrace{P^{(n)}(t)}_{=0} - M \cdot n! = f^{(n)}(t) - M \cdot n!$$
and hence we are done if we find $x \in (\alpha, \beta)$ with $g^{(n)}(x) = 0$. By construction, $P^{(k)}(\alpha) = f^{(k)}(\alpha)$ for $k = 0, \dots, n-1$ and hence $g^{(k)}(\alpha) = f^{(k)}(\alpha) - P^{(k)}(\alpha) = 0$ for such k; moreover, by the choice of M we have $g(\alpha) = g(\beta) = 0$. We now iteratively apply the MVT to produce $x_{k+1} \in (\alpha, x_k)$ such that $g^{(k)}(x_k) = 0$ for each $k = 0, \dots, n$; hence we have some $x \in (\alpha, x_{n-1}) \subset (\alpha, \beta)$ such that $g^{(n)}(x) = 0$ and thus we have $M = f^{(n)}(x)/n!$ as desired. ☒

Definition: We say that $f : [a,b] \to \mathbb{R}$ is <u>real analytic</u> if $f^{(n)}$ exists on (a,b) for every $n \geq 1$ (i.e. <u>smooth</u>) and such that for each $y \in (a,b)$ we have
$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(y)}{k!}(x-y)^k \quad \left(\text{i.e. if } f \text{ is a limit of its Taylor Polynomials}\right)$$

<u>Examples:</u>
- Polynomials, e^x, trigonometric fns, logarithms are!

- $|x|$ is not as it is not diff'ble.

- let
$$f(x) = \begin{cases} e^{-\frac{1}{x}}, & \text{for } x > 0, \\ 0, & \text{for } x \leq 0, \end{cases}$$



then $f^{(n)}$ exists on $\mathbb{R}$ for every $n \geq 1$ (<u>check</u>), and in fact $f^{(n)}(0) = 0$. Hence for each $n \geq 1$, any Taylor polynomial of f at 0 is 0; but $f(\varepsilon) > 0$ for all small $\varepsilon > 0$ so we see that f is <u>not</u> analytic!

Multivariable differentiation

The definition of the derivative for functions $f:[a,b]\to\mathbb{R}^n$ or to $\mathbb{C}$ makes sense provided we interpret norms and points correctly. All the rules (sum, product, chain, diff'ble $\Rightarrow$ cts) hold with correct interpretation (e.g. if $\mathbb{R}^n$ then $f\cdot g$ dot product for product rule).

Remark: If $f:[a,b]\to\mathbb{C}$ we can write $f=f_1+if_2$ for $f_1,f_2:[a,b]\to\mathbb{R}$; thus $f'=f_1'+if_2'$ so that f is diff'ble $\Leftrightarrow$ f_1,f_2 are. Similarly, if $f:[a,b]\to\mathbb{R}^n$ then $f=(f_1,\dots,f_n)$ for $f_1,\dots,f_n:[a,b]\to\mathbb{R}$; thus $f'=(f_1',\dots,f_n')$ so that f is diff'ble $\Leftrightarrow$ $f_1,\dots,f_n$ are.

We now see that MVT and its consequences fail however:

Examples:

- If $f:\mathbb{R}\to\mathbb{C}$ is defined by $f(\theta)=e^{i\theta}=\cos(\theta)+i\sin(\theta)$ then $f(2n\pi)=f(0)=1$ for each $n\geq 1$ but we have $f'(\theta)=-\sin(\theta)+i\cos(\theta)\Rightarrow|f'(\theta)|=1$ for any $\theta\in\mathbb{R}$. Hence
$$f(2\pi)-f(0)\neq 2\pi f'(\theta) \text{ for } \underline{\text{any}}\ \theta\in(0,2\pi);$$
So MVT fails!

- If $g:(0,1)\to\mathbb{C}$ is defined by $g(x)=x+x^2e^{\frac{i}{x^2}}$ and $f:\mathbb{R}\to\mathbb{C}$ is defined by $f(x)=x$ then
$$\lim_{x\to 0}\frac{f(x)}{g(x)}=1,$$
but we have $g'(x)=1+\left(2x-\frac{2i}{x}\right)e^{\frac{i}{x^2}}$ so that
$$|g'(x)|\geq -1+\left|2x-\frac{2i}{x}\right|\geq -1+\frac{2}{x}\quad(\text{as } x\in(0,1)).$$
Thus
$$\left|\frac{f'(x)}{g'(x)}\right|=\frac{1}{|g'(x)|}\leq\frac{x}{2-x}\Rightarrow\lim_{x\to 0}\frac{f'(x)}{g'(x)}=0.$$

We therefore see that L'Hospital's rule fails too!

Note that the MVT shows that for diffble $f:[a,b]\to\mathbb{R}$

$$|f(b)-f(a)| \leq (b-a)\sup_{a<x<b}|f'(x)|;$$

even though MVT does not hold for multivariable functions we have an analogue of the above:

<u>Theorem:</u> If $f:[a,b]\to\mathbb{R}^n$ is a cts function and diffble on (a,b) then there is $x\in(a,b)$ such that

$$|f(b)-f(a)| \leq (b-a)|f'(x)|.$$

<u>Proof:</u> Let $z = f(b)-f(a)$ and define $g:[a,b]\to\mathbb{R}$ by setting $g(t)=z\cdot f(t)$; g is then a cts function which is diffble on (a,b) so the MVT implies

$$g(b)-g(a) = (b-a)g'(x) = (b-a)(z\cdot f'(x))$$

for some $x\in(a,b)$. Note that

$$g(b)-g(a) = z\cdot(f(b)-f(a)) = z\cdot z = |z|^2,$$

and thus combining the above we see that

$$|z|^2 \leq (b-a)(z\cdot f'(x)).$$

By the Cauchy-Schwarz inequality we see that

$$|z\cdot f'(x)| \leq |z|\,|f'(x)|$$

and so

$$|z|^2 \leq (b-a)|z\cdot f'(x)| \leq (b-a)|f'(x)|;$$

thus we have (even if $z=0$)

$$|z| \leq (b-a)|f'(x)|$$

or that $|f(b)-f(a)| \leq (b-a)|f'(x)|$ for some $x\in(a,b)$, as desired. ∎

$\boxed{\text{End L20}}$

Sequences and series of functions

Definition: Given a sequence (f_n) of $\mathbb{C}$ valued functions on a metric space (X,d) such that $\lim_{n\to\infty} f_n(x)$ exists for every $x \in X$ we define the limit, $f: X \to \mathbb{C}$, by setting

$$f(x) = \lim_{n\to\infty} f_n(x) \quad \text{for each } x \in X.$$

We then say that (f_n) converges pointwise to f. Similarly, if $\sum_n f_n(x)$ converges for every $x \in X$ we define the sum, $f: X \to \mathbb{C}$, by setting

$$f(x) = \sum_n f_n(x) \quad \text{for each } x \in X.$$

We want to understand whether limits/sums of functions preserve the properties of the sequence; e.g. if (f_n) is a sequence of cts/diff'ble fcs, is the limit/sum cts/diff'ble? Moreover, can we relate (f_n') to f'?

Recall that f is cts at $x \iff f(x) = \lim_{t\to x} f(t)$, and thus asking whether the limit of cts fcs is cts is asking if

$$\lim_{t\to x}\lim_{n\to\infty} f_n(t) = \lim_{n\to\infty}\lim_{t\to x} f_n(t);$$

namely, can we 'swap' limits? We see now that pointwise convergence is not sufficient:

Examples: • Let $s_{m,n} = \dfrac{m}{m+n}$ for each $m,n \geq 1$. Then,

$$\begin{cases} \lim_{m\to\infty} s_{m,n} = 1 \implies \lim_{n\to\infty}\lim_{m\to\infty} s_{m,n} = 1, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \neq \\ \lim_{n\to\infty} s_{m,n} = 0 \implies \lim_{m\to\infty}\lim_{n\to\infty} s_{m,n} = 0. \end{cases}$$

So ptwise convergence of fcs is not enough to guarantee cty of limits!

- For $x \in \mathbb{R}$ and $n \geq 0$ let $f_n(x) = \frac{x^2}{(1+x^2)^n}$ — cts — and set

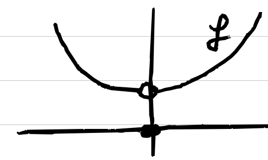
$$f(x) = \sum_{n=0}^{\infty} f_n(x) = \sum_{n=0}^{\infty} \frac{x^2}{(1+x^2)^n};$$

as $f_n(0) = 0$ we have $f(0) = 0$. If $x \neq 0$ then we have

$$f(x) = x^2 + \frac{x^2}{(1+x^2)} + \frac{x^2}{(1+x^2)^2} + \dots = 1 + x^2,$$

hence

$$f(x) = \begin{cases} 0, & \text{for } x = 0 \\ 1+x^2, & \text{for } x \neq 0 \end{cases}$$



Thus f is not cts!

- For $x \in \mathbb{R}$ and $n \geq 1$ let $g_n(x) = \frac{\sin(nx)}{\sqrt{n}}$ so that $g(x) = \lim_{n\to\infty} g_n(x) = 0$ but $g_n'(x) = \sqrt{n}\cos(nx)$ is such that $\lim_{n\to\infty} g_n'(0) = \sqrt{n}$ so that g_n' does not converge pointwise to g'.

Uniform Convergence

We introduce a stronger notion of convergence for functions:

Definition: We say that a sequence (f_n) of $\mathbb{C}$ valued functions converges uniformly to f on $E \subset X$, for a metric space (X,d), if for every $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that

$$n \geq N \Rightarrow |f_n(x) - f(x)| < \varepsilon \text{ for every } x \in E.$$

We say that $\sum_n f_n$ converges uniformly to f if the partial sums $S_N = \sum_{n \geq 1}^{N} f_n$ converge uniformly.

Remark: (f_n) converges uniformly to $f \Rightarrow (f_n)$ converges ptwise to f.

We sometimes write $f_n \xrightarrow{\text{unif}} f$ or $f_n \xrightarrow{\text{ptwise}} f$ to abbreviate.

Theorem: A sequence (f_n) of $\mathbb{C}$ valued functions on $E \subset X$, for a metric space (X,d), converge uniformly on E if and only if it is <u>uniformly Cauchy</u>; namely for each $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that

$$m, n \geq N \implies |f_m(x) - f_n(x)| < \varepsilon \text{ for all } x \in E.$$

Proof: ($\Rightarrow$) If $f_n \xrightarrow{\text{unif}} f$ and $\varepsilon > 0$ then there is some $N \in \mathbb{N}$ such that $n \geq N \implies |f_n(x) - f(x)| \leq \frac{\varepsilon}{2}$ for all $x \in E$, hence if $m, n \geq N$ then

$$|f_m(x) - f_n(x)| \leq |f_m(x) - f(x)| + |f(x) - f_n(x)|$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

for all $x \in E$; thus (f_n) is uniformly Cauchy.

($\Leftarrow$) If (f_n) is uniformly Cauchy and $\varepsilon > 0$ then there is some $N \in \mathbb{N}$ such that $m, n \geq N \implies |f_m(x) - f_n(x)| < \frac{\varepsilon}{2}$ for all $x \in E$. Noting that the sequence $(f_m(x))$ is Cauchy in $\mathbb{C}$ for each $x \in E \implies f(x) = \lim_{m\to\infty} f_m(x)$ exists for each $x \in E$. Combining these facts we have that if $m, n \geq N$ then for each $x \in E$

$$|f_m(x) - f_n(x)| < \frac{\varepsilon}{2} \implies -\frac{\varepsilon}{2} < f_m(x) - f_n(x) < \frac{\varepsilon}{2},$$

So that $-\frac{\varepsilon}{2} \leq \lim_{m\to\infty} \left(f_m(x) - f_n(x)\right) = f(x) - f_n(x) \leq \frac{\varepsilon}{2}$ and so $|f(x) - f_n(x)| < \varepsilon$ for all $x \in E$; thus $f_n \xrightarrow{\text{unif}} f$. ∎

Theorem: Suppose that $f_n \xrightarrow{\text{ptwise}} f$ on $E \subset X$, for a metric space (X,d), then $f_n \xrightarrow{\text{unif}} f \iff M_n = \sup_{x\in E} |f_n(x) - f(x)| \to 0$ as $n \to \infty$.

Proof: If $f_n \xrightarrow{\text{unif}} f$ then $M_n \to 0$ by definition. If $M_n \to 0$ then for each $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that $M_n < \varepsilon$ for $n \geq N$. Hence for $n \geq N$ we have

$$|f_n(x) - f(x)| \leq \sup_{x\in E} |f_n(x) - f(x)| = M_n < \varepsilon \text{ for all } x \in E,$$

so $f_n \xrightarrow{\text{unif}} f$. ∎

<u>Example:</u> The functions $f_n(x) = \frac{1}{nx+1}$ on $(0,1) \subset \mathbb{R}$ for $n \geq 1$ are such that $f_n \xrightarrow{\text{Ptwise}} 0$ but for each $n \geq 1$ we have $|1 - f_n(x)| = \left|1 - \frac{1}{nx+1}\right| = \left|\frac{nx}{nx+1}\right| = \frac{x}{x+1}$

so we see that $\sup_{x \in (0,1)} |1 - f_n(x)| \not\to 0$ so $f_n \overset{\text{unif}}{\not\to} 0$.

<u>Theorem:</u> (<u>Weierstrass M-test</u>) If (f_n) is a sequence of $\mathbb{C}$ valued functions on $E \subset X$, for a metric space (X,d), with $|f_n(x)| \leq M_n$ for each $x \in E, n \geq 1$, then $\sum_n f_n$ converges uniformly if $\sum_n M_n$ converges.

<u>Proof:</u> As $\sum_n M_n$ converges in $\mathbb{R}$, its partial sums, (S_N), are Cauchy in $\mathbb{R}$. Hence for each $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that $m, n \geq N$ implies by the Δ inequality that

$$\left|\sum_{i=m}^{n} f_i(x)\right| \leq \sum_{i=m}^{n} |f_i(x)| \leq \sum_{i=m}^{n} M_i < \varepsilon$$

So that the partial sums of $\sum_n f_n$ are uniformly Cauchy, and hence uniformly convergent by the first result above. $\boxtimes$

<u>Remark:</u> The converse statement fails in general! To see this we can choose the constant functions on $\mathbb{R}$ defined by

$$f_n(x) = \frac{(-1)^{n+1}}{n} \text{ for } n \geq 1;$$

then $|f_n(x)| = \frac{1}{n}$ so that

$$\sum_n f_n(x) = \log(2) \quad \text{for all } x \in \mathbb{R}$$

but $\sum_n \frac{1}{n}$ diverges! Hence $\sum_n f_n$ converges uniformly but $\sum_n M_n$ does not converge, so converse fails!

One could also consider 'sliding' bump functions

$$f_n = \frac{1}{n}\chi_{(n,n+1)} = \begin{cases} \frac{1}{n}, & \text{on } (n,n+1) \\ 0, & \text{o/w} \end{cases}.$$

$\boxed{\text{End L21}}$

We now see how uniform convergence guarantees cty of limits:

Theorem: Suppose $f_n \xrightarrow{\text{Unif}} f$ ($\mathbb{C}$ Valued) on $E \subset X$, for a metric space (X, d), x is a limit point of E and $\lim_{t \to x} f_n(t) = A_n$. Then (A_n) converges and

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

Moreover, if the (f_n) are cts $\Rightarrow$ f is cts.

Proof: As $f_n \xrightarrow{\text{Unif}} f$ the sequence (f_n) is uniformly Cauchy; hence for each $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that

$$n, m \geq N \Rightarrow |f_n(t) - f_m(t)| < \varepsilon \text{ for each } t \in E.$$

We then can send $t \to x$ to see that

$$n, m \geq N \Rightarrow |A_n - A_m| < \varepsilon,$$

so that (A_n) is Cauchy in $\mathbb{C}$ and thus converges to some $A \in \mathbb{C}$. Note that for each $n \geq 1$ and $t \in E$ we have

$$|f(t) - A| \leq |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A|,$$

by applying Δ ineq. twice. For $\varepsilon > 0$ we first choose $n \geq 1$ such that both

$$|f(t) - f_n(t)| < \frac{\varepsilon}{3} \text{ for all } t \in E \quad (\text{Since } f_n \xrightarrow{\text{Unif}} f),$$

and

$$|A_n - A| < \frac{\varepsilon}{3} \quad (\text{Since } A_n \to A \text{ as } n \to \infty).$$

Finally, we choose some neighbourhood, U, of x in X such that

$$|f_n(t) - A_n| < \frac{\varepsilon}{3} \text{ for all } t \in (E \cap U) \setminus \{x\} \quad (\text{Since } f_n(t) \to A_n \text{ as } t \to x).$$

Combining the above we have that

$$|f(t) - A| < \varepsilon \text{ for all } t \in (E \cap U) \setminus \{x\},$$

thus we have the desired conclusions. $\blacksquare$

Remark: Previous examples show that $f_n \xrightarrow{\text{Ptwise}} f$ and f cts do not guarantee $f_n \xrightarrow{\text{unif}} f$!

We can however guarantee the converse on compact sets:

Theorem: (Dini's theorem) If $f_n \xrightarrow{\text{Ptwise}} f$ on a compact set K, the $(f_n), f$ ($\mathbb{C}$ valued) are cts, and $f_n \geq f_{n+1}$ (or with $\leq$) for each $n \geq 1$, then $f_n \xrightarrow{\text{unif}} f$ on K.

Proof: Consider the functions $g_n = f_n - f$ for each $n \geq 1$, then the (g_n) are cts with $g_n \xrightarrow{\text{Ptwise}} 0$ and $g_n \geq g_{n+1}$; we show that $g_n \xrightarrow{\text{unif}} 0$, which implies $f_n \xrightarrow{\text{unif}} f$. For $\varepsilon > 0$ let

$$K_n = g_n^{-1}([\varepsilon, \infty)) \text{ (closed as } g_n \text{ is cts)}$$

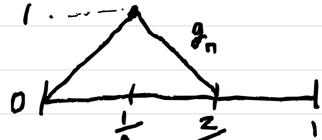
for each $n \geq 1$; as K is compact $\Rightarrow K_n$ compact also. If $x \in K_{n+1}$ then $\varepsilon \leq g_{n+1}(x) \leq g_n(x)$, so $x \in K_n$ also, thus $K_{n+1} \subset K_n$ for each $n \geq 1$. Now for each $x \in K$ we have $g_n(x) \to 0$ so that $x \notin K_n$ for all $n \geq 1$ sufficiently large (since $g_n(x) < \varepsilon$ eventually in n); thus we have

$$\bigcap_{n \geq 1} K_n = \emptyset \text{ and } K_{n+1} \subset K_n \text{ for } n \geq 1 \Rightarrow K_n = \emptyset \text{ for some } n \geq 1.$$

To conclude we must have some $N \in \mathbb{N}$ such that $K_n = \emptyset$ for $n \geq N \Rightarrow g_n(x) < \varepsilon$ for all $x \in K$ and $n \geq N$; as we must have $g_n \geq 0$ for every $n \geq 1$ (else $g_n \stackrel{\text{Ptwise}}{\not\to} 0$ since we have $g_{n+1} \leq g_n$). This implies $g_n \xrightarrow{\text{unif}} 0 \iff f_n \xrightarrow{\text{unif}} f$. $\blacksquare$

Remark: We need compactness as the examples $f_n(x) = \frac{1}{nx+1}$ on $(0,1)$ show ($f_n \geq f_{n+1}$ holds). Similarly the monotone requirement is necessary, consider the 'sliding hump' functions on $[0,1]$:

$$g_n(x) = \begin{cases} nx, & \text{for } 0 \leq x \leq \frac{1}{n} \\ 2 - nx, & \text{for } \frac{1}{n} \leq x \leq \frac{2}{n} \\ 0, & \text{for } \frac{2}{n} < x \end{cases}$$



These are not monotone, $g_n \xrightarrow{\text{Ptwise}} 0$ and $g_n \stackrel{\text{unif}}{\not\to} 0$ (as $g_n(\frac{1}{n}) = 1$ for all $n \geq 1$).

Recall, we defined $\mathscr{C}(X)$ to be the set of cts bounded ($\mathbb{C}$ valued) fcs on a metric space (X,d), which become a normed vector space with the supremum norm, $\|f\| = \sup_{x\in X} |f(x)|$. This norm induced a metric, $d(f,g) = \|f-g\|$ on $\mathscr{C}(X)$. Since $f_n \overset{\text{unif}}{\rightrightarrows} f \iff \sup_{x\in X} |f_n - f| \to 0$ we see that $f_n \to f$ in $\mathscr{C}(X) \iff f_n \overset{\text{unif}}{\rightrightarrows} f$ on X. Moreover:

Theorem: $\mathscr{C}(X)$ is a complete metric space.

Proof: If $(f_n) \subset \mathscr{C}(X)$ is a Cauchy sequence it must be uniformly Cauchy $\iff$ there is some $f: X \to \mathbb{C}$ such that $f_n \overset{\text{unif}}{\rightrightarrows} f$ on X. As the (f_n) are cts we have that f is also cts. We also know that f is bounded since $f_n \overset{\text{unif}}{\rightrightarrows} f \Rightarrow |f_n(x) - f(x)| < 1$ for all $x \in X$ and some $n \geq 1$ sufficiently large. Hence $f \in \mathscr{C}(X)$ and since $f_n \overset{\text{unif}}{\rightrightarrows} f$ we have $d(f_n, f) \to 0$ as $n \to \infty$. ∎

We now discuss how uniform convergence is related to differentiation. We have already seen that $g_n(x) = \frac{\sin(nx)}{\sqrt{n}}$ are s.t. $g_n \xrightarrow{\text{pwise}} 0$ but g_n' does not converge. Similarly even if the derivatives converge, the limiting function may fail to be differentiable; for example consider

$$h_n(x) = x^n \text{ on } [0,1], \text{ then } h_n'(x) = nx^{n-1} \text{ so that}$$

$$h_n'(x) \xrightarrow{\text{pwise}} h(x) = \begin{cases} 0, & \text{for } 0 \leq x < 1 \\ 1, & \text{for } x = 1 \end{cases}$$

(one could also consider $\sqrt{x^2 + \frac{1}{n}} \to |x|$).

Theorem: Let (f_n) be a sequence of $\mathbb{C}$ valued diff'ble fcs on $[a,b]$ such that $(f_n(x_0))$ converges for some $x_0 \in [a,b]$. If (f_n') converges uniformly on $[a,b]$, then (f_n) converges uniformly on $[a,b]$ to a function f and $f'(x) = \lim_{n\to\infty} f_n'(x)$ for all $x \in [a,b]$.

Proof: Let $\varepsilon > 0$, since $(f_n(x_0))$ converges it is Cauchy so there is some $N \in \mathbb{N}$ such that

$$m, n \geq N \Rightarrow |f_m(x_0) - f_n(x_0)| < \frac{\varepsilon}{2},$$

and since (f_n') is uniformly convergent it is uniformly Cauchy and so →

$(f_n \xrightarrow{\text{unif}} f)$

Potentially taking N larger we also have

$$m,n \geq N \implies |f_m'(t) - f_n'(t)| < \frac{\varepsilon}{2(b-a)} \text{ for all } t \in [a,b].$$

By MVT, for any $x,t \in [a,b]$ $(x \neq t)$ there is some $c \in (x,t)$ (or $c \in (t,x)$) such that if $m,n \geq N$ then

$$(+) \quad |(f_m(x) - f_n(x)) - (f_m(t) - f_n(t))| = |x-t||f_m'(c) - f_n'(c)|$$
$$\leq \frac{\varepsilon|x-t|}{2(b-a)} \leq \frac{\varepsilon}{2}.$$

Hence for each $x \in [a,b]$ we have for $m,n \geq N$ that

$$|f_m(x) - f_n(x)| \leq |(f_m(x) - f_n(x)) - (f_m(x_0) - f_n(x_0))|$$
$$+ |f_m(x_0) - f_n(x_0)|$$
$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon;$$

so that (f_n) is uniformly Cauchy and hence uniformly convergent on $[a,b]$. Let the limit of (f_n) be f and, for $x \in [a,b]$ fixed, define for $t \in [a,b] \setminus \{x\}$ functions

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t-x}, \quad \phi(t) = \frac{f(t) - f(x)}{t-x};$$

$(f_n' \xrightarrow{\text{Ptwise}} f')$

note then that $\lim_{t \to x} \phi_n(t) = f_n'(x)$ for each $n \geq 1$. By $(+)$ we have

$$|\phi_m(t) - \phi_m(t)| \leq \frac{\varepsilon}{2(b-a)} \text{ for } m,n \geq N,$$

hence (ϕ_n) converges uniformly on $[a,b] \setminus \{x\}$ (as it is uniformly Cauchy). As $f_n \xrightarrow{\text{Unif}} f$ we have that $\phi_n \xrightarrow{\text{Ptwise}} \phi$ on $[a,b] \setminus \{x\}$, and as x is a limit point of $[a,b] \setminus \{x\}$ by an earlier result we know that

$$\lim_{t \to x} \phi(t) = \lim_{t \to x} \lim_{n \to \infty} \phi_n(t) = \lim_{n \to \infty} \lim_{t \to x} \phi_n(t) = \lim_{n \to \infty} f_n'(x),$$

so that $f'(x) = \lim_{t \to x} \phi(t) = \lim_{n \to \infty} f_n'(x)$ for all $x \in [a,b]$. ▣

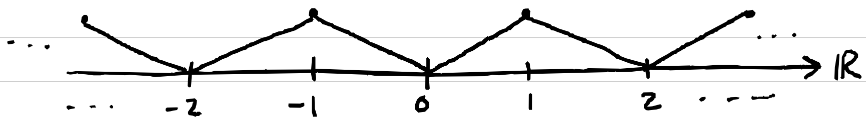
<u>Remark:</u> If we don't have $(f_n(x_0))$ convergent for some $x_0 \in [a,b]$ then (f_n) may not even converge; e.g. consider $f_n(x) = n$ on $[0,1]$, then $f_n' = 0$ for all $n \geq 1$ but f_n diverge! $\boxed{\text{End L22}}$

A continuous but nowhere differentiable function

We will use the theory we have built up to see how wildly beha-ved cts functions can be in general; Constructions of this type were said to belong to a 'gallery of monsters' by Poincaré (fractals).

Theorem: There is a cts function $f : \mathbb{R} \to \mathbb{R}$ which is nowhere diff'ble.

Proof: Let us extend $|x|$ on $[-1, 1]$ to a periodic function, φ, on $\mathbb{R}$: (so that $\varphi(x+2) = \varphi(x)$ for any $x \in \mathbb{R}$)

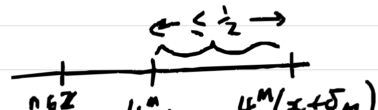


By the reverse Δ inequality we have (check) for $s, t \in \mathbb{R}$ that $|\varphi(x) - \varphi(y)| \leq |x - y| \Rightarrow \varphi$ is cts on $\mathbb{R}$.

We then define $f_n(x) = \left(\frac{3}{4}\right)^n \varphi(4^n x)$ for $x \in \mathbb{R}$, so that $|f_n(x)| \leq \left(\frac{3}{4}\right)^n$ for each $n \geq 1$ and $x \in \mathbb{R}$. By the Weierstrass M-test we have that $\sum_n f_n$ converges uniformly to

$$f(x) = \sum_n f_n(x) = \sum_n \left(\tfrac{3}{4}\right)^n \varphi(4^n x) \text{ Since } \sum_n \left(\tfrac{3}{4}\right)^n < \infty.$$

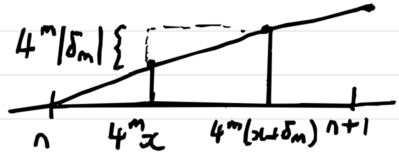
Moreover, since the partial sums are cts we have from their uniform convergence that f is also cts; we will show that it is nowhere diff'ble.



For each $x \in \mathbb{R}$ and $m \geq 1$ let

$$\delta_m = \pm \tfrac{1}{2} \cdot 4^{-m}$$ where $\pm$ is chosen so that the interval between $4^m x$ and $4^m(x + \delta_m)$ contains no integer (as $4^m |\delta_m| = \frac{1}{2}$). We now set $\gamma_n = \frac{1}{\delta_m}\left(\varphi(4^n(x+\delta_m)) - \varphi(4^n x)\right)$ for $n \in \mathbb{N}$, so that if $n > m$ we have $\gamma_n = 0$ (since then $4^n \delta_m$ is even) and if $0 \leq n \leq m$ then $|\gamma_n| \leq 4^n$ (since $|\varphi(x) - \varphi(y)| \leq |x - y|$). We also note that $|\gamma_m| = 4^m$ (since no integer is between $4^m x$ and $4^m(x + \delta_m)$ we have $|\varphi(4^m(x+\delta_m)) - \varphi(4^m x)| = 4^m |\delta_m| = \frac{1}{2}$).

Visually,



Finally we have that f is not diff'ble at x since

$$\frac{f(x+\delta_m)-f(x)}{\delta_m} = \sum_n \left(\frac{3}{4}\right)^n \frac{\varphi(4^n(x+\delta_m))-\varphi(4^n x)}{\delta_m}$$

$$= \sum_n \left(\frac{3}{4}\right)^n \gamma_n = \sum_{n=0}^{m} \left(\frac{3}{4}\right)^n \gamma_m$$

$$= \left(\frac{3}{4}\right)^m \gamma_m + \sum_{n=0}^{m-1} \left(\frac{3}{4}\right)^n \gamma_n$$

and so by reverse Δ ineq. we have

$$\left|\frac{f(x+\delta_m)-f(x)}{\delta_m}\right| = \left|\left(\frac{3}{4}\right)^m \gamma_m - \left(-\sum_{n=0}^{m-1}\left(\frac{3}{4}\right)^n \gamma_n\right)\right|$$

$$\geq \left|\left(\frac{3}{4}\right)^m \gamma_m\right| - \left|\sum_{n=0}^{m-1}\left(\frac{3}{4}\right)^n \gamma_n\right|.$$

Recalling $|\gamma_m| = 4^m$ and $|\gamma_n| \leq 4^n$ for $0 \leq n \leq m$ we have

$$\left|\frac{f(x+\delta_m)-f(x)}{\delta_m}\right| \geq 3^m - \sum_{n=0}^{m-1} 3^n = 3^m - \left(\frac{3^m-1}{2}\right) = \frac{3^m+1}{2},$$

but $3^m \to \infty$ as $m \to \infty \implies f'(x)$ DNE. As $x \in \mathbb{R}$ was arbitrary $\implies f$ not diff'ble anywhere! $\blacksquare$

Equicontinuity

The Bolzano-Weierstrass theorem guarantees that bounded sequences in $\mathbb{C}$ have convergent subsequences. One could ask whether an analogous result holds for bounded sequences of functions:

Definition: We say that a sequence (f_n) of $\mathbb{C}$ valued fns on a metric space, (X,d), are:

- Pointwise bounded on X if there is some $\phi : X \to \mathbb{R}$ such that $|f_n| \leq \phi$ for all $n \geq 1$.

- Uniformly bounded on X if there is some $M \geq 0$ such that $|f_n| \leq M$ for all $n \geq 1$.

<u>Remarks:</u> • Uniformly convergent $\Rightarrow$ uniformly bounded.

• If X is countable one can use Cantor's diagona-lisation argument to find a subsequence converging pointwise on X if the sequence is pointwise bounded.

• If a sequence is uniformly bounded it does not necessarily contain a pointwise convergent subsequence. This is shown in the text using the dominated convergence theorem on the sequence $(\sin(nx))$ on $[0, 2\pi]$.

One could also ask if convergent uniformly bounded sequences of fns contain uniformly convergent subsequences; but this also fails:

<u>Example:</u> Let $f_n(x) = \dfrac{x^2}{x^2+(1-nx)^2}$ on $[0,1]$ for $n \geq 1$. We then see that $|f_n(x)| \leq 1$ so that (f_n) is uniformly bounded, and moreover $f_n(x) \to 0$ for all $x \in [0,1]$. However, we have $f_n(1/n) = 1$ for all $n \geq 1$; so no subsequence converges uniformly.

Let us address the first remark above:

<u>Theorem:</u> If (f_n) is a pointwise bounded sequence of $\mathbb{C}$ valued functions on a countable metric space, (X,d), then there is a subsequence (f_{n_k}) such that $(f_{n_k}(x))$ converges for every $x \in X$.

<u>Proof:</u> Let us enumerate $X = (x_i)$, as $(f_n(x_1))$ is bounded there is a subsequence (f_k^1) of (f_n) such that $(f_k^1(x_1))$ converges as $k \to \infty$. As $(f_k^1(x_2))$ is bounded there is a subsequence (f_k^2) of (f_k^1) such that $(f_k^2(x_2))$ converges as $k \to \infty$. We inductively conti-nue this process, generating a subsequence $(f_k^{\ell+1})$ of (f_k^ℓ) such that $(f_k^{\ell+1}(x_{\ell+1}))$ converges as $k \to \infty$. We then choose the 'diagonal' subsequence (f_k^k) of (f_n) which is such that $(f_k^k(x_i))$ converges for each $i \geq 1$ (since (f_k^k) is a subsequence of (f_k^i) for $k \geq i$ by construction). Relabling $f_k^k = f_{n_k}$ for each $k \geq 1$ we are done. ∎

Visually,

$$\begin{array}{c|ccccc} x_1 & f_1^1 & f_2^1 & f_3^1 & f_4^1 & \cdots \\ x_2 & f_1^2 & f_2^2 & f_3^2 & f_4^2 & \cdots \\ x_3 & f_1^3 & f_2^3 & f_3^3 & f_4^3 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

(ith row converges on x_i)

(Each row is a subsequence of the one above)

(We choose the diagonal subsequence so that $(f_k^k(x_i))$ is convergent for each $i \geq 1$.)

We have seen that pointwise and uniform boundedness are usually not enough to extract 'well behaved' subsequences. We thus introduce:

Definition: A collection, $\mathcal{F}$, of $\mathbb{C}$ valued functions on a set $E \subset K$, for a metric space (X,d), is said to be <u>equicontinuous</u> on E if for each $\varepsilon > 0$ there is some $\delta > 0$ such that

$$x,y \in E,\ d(x,y) < \delta \Rightarrow |f(x) - f(y)| < \varepsilon \text{ for every } f \in \mathcal{F}.$$

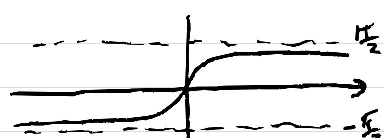
(Sometimes called uniform equicontinuity)

Remark: If $\mathcal{F}$ is equicts $\Rightarrow$ every $f \in \mathcal{F}$ is uniformly cts.

Examples:
- If $\mathcal{F} = (f_n)$ for diffble fns on $[0,1]$ with (f_n') uniformly bounded $\Rightarrow \mathcal{F}$ is equicts; if $|f_n'(x)| \leq M$ for all $n \geq 1$, $x \in [0,1]$ and $\varepsilon > 0$, set $\delta = \frac{\varepsilon}{M+1}$ so that by MVT we have
$$|x-y| < \frac{\varepsilon}{M+1} \Rightarrow |f_n(x) - f_n(y)| \leq |x-y| \cdot \sup_{t \in [0,1]} |f'(t)| < \frac{\varepsilon M}{M+1} < \varepsilon.$$

- We saw that $\mathcal{G} = \left(\frac{x^2}{x^2 + (1-nx)^2}\right)$ on $[0,1]$ was uniformly bdd, ptwise $\to 0$, but had no uniformly convergent subsequence. $\mathcal{G}$ is also <u>not</u> equicts as $g_n(\frac{1}{n}) = 1$ but $g_n(0) = 0$ for every $n \geq 1$ so no $\delta > 0$ works for $\varepsilon \in (0,1)$!

- $\mathcal{H} = (\arctan(nx))$ is not equicts since $\arctan(nx) \to \pm\frac{\pi}{2}$ if $x >/< 0$.



We will see that there are strong relations between equicty and uniform convergence.

$\boxed{\text{End L23}}$

<u>Theorem:</u> Let (K,d) be compact and $(f_n) \subset \mathcal{C}(K)$, then if (f_n) converges uniformly on K, (f_n) is equicts on K.

<u>Proof:</u> Let $\varepsilon > 0$ and note that since (f_n) converges uniformly, and hence uniformly Cauchy, there is some $N \in \mathbb{N}$ such that

$$m, n \geq N \Rightarrow \|f_n - f_m\| < \frac{\varepsilon}{3}.$$

Now as cts fns on compact sets are uniformly cts there is some $\delta_i > 0$ for each $i \geq 1$ such that

$$d(x,y) < \delta_i \Rightarrow |f_i(x) - f_i(y)| < \frac{\varepsilon}{3}.$$

Combining the above, for $n \geq N$ and $\delta_N > 0$ we have

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)|$$
$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \quad \text{if } d(x,y) < \delta_N.$$

Hence, if we set $\delta = \min\{\delta_1, \dots, \delta_N\}$ we have

$$d(x,y) < \delta \Rightarrow |f_n(x) - f_n(y)| < \varepsilon \text{ for all } n \geq 1,$$

so that (f_n) is equicts. $\boxtimes$

<u>Theorem:</u> Let (K,d) be compact and $(f_n) \in \mathcal{C}(K)$, then if (f_n) is pointwise bounded and equicts on K then both:

(<u>Arzelà-Ascoli</u>)

① (f_n) is uniformly bounded on K.

② (f_n) has a uniformly convergent subsequence.

<u>Proof:</u> For ①, we choose $\delta > 0$ from equicty of (f_n) so that

$$d(x,y) < \delta \Rightarrow |f_n(x) - f_n(y)| < 1 \text{ for all } n \geq 1.$$

By compactness of K there are $x_1, \dots, x_L \in K$ such that

$$K \subset \bigcup_{i=1}^{L} B_\delta(x_i),$$

and since (f_n) is pointwise bdd there are $M_1, \dots, M_L > 0$ such that $|f_n(x_i)| \leq M_i$ for each $i = 1, \dots, L$; we then have $|f_n(x)| \leq M + 1$ for all $x \in K$ and $n \geq 1$ where $M = \max\{M_1, \dots, M_L\}$.

(Compact Sets are Separable)

For 2, we first show that K contains an at most countable subset $E \subset X$ which is dense; i.e. if $U \subset K$ is open then $E \cap U \neq \emptyset$. If K is at most countable then this holds automatically. If not, since K is compact, for each $n \geq 1$ there is a finite set $\{x_i^n\}_{i=1}^{L_n} \subset K$ such that $K \subset \bigcup_{i=1}^{L_n} B_{\frac{1}{n}}(x_i)$;

We then set $E = \bigcup_{n \geq 1} \{x_i^n\}_{i=1}^{L_n}$, which is at most countable. If $U \subset K$ is open then for each $x \in U$ there is some $\delta > 0$ such that $B_\delta(x) \subset U$ and hence for $n \geq \frac{1}{\delta}$ (Archimedean property) there is some $x_i^n \in E$ such that $d(x, x_i^n) < \frac{1}{n} \leq \delta$ so that $x_i^n \in U$ also; hence $E \cap U \neq \emptyset$ so E is dense in K (i.e. K is seperable).

Now as (f_n) is pointwise bounded on $E \subset K$ and E is at most countable, there is a pointwise convergent subsequence, (f_{n_k}), of (f_n) on E; set $g_k = f_{n_k}$ for each $k \geq 1$, we will show that (g_k) is uniformly convergent. Let $\varepsilon > 0$ and choose $\delta > 0$ from equicty of (g_k) such that

$$d(x,y) < \delta \Rightarrow |g_k(x) - g_k(y)| < \frac{\varepsilon}{3} \text{ for all } k \geq 1.$$

For $n \geq \frac{1}{\delta}$ again we have $K \subset \bigcup_{i=1}^{L_n} B_\delta(x_i^n)$ for $\{x_i^n\}_{i=1}^{L_n} \subset E$ by construction of E. As (g_k) is pointwise convergent on E there is some $N \in \mathbb{N}$ such that

$$l, m \geq N \Rightarrow |g_l(x_i^n) - g_m(x_i^n)| < \frac{\varepsilon}{3} \text{ for } i = 1, \dots, L_n.$$

Now, if $x \in K$ then $x \in B_\delta(x_i^n)$ for some $i = 1, \dots, L_n$ and so for $l, m \geq N$ we have

$$|g_l(x) - g_m(x)| \leq |g_l(x) - g_l(x_i^n)| + |g_l(x_i^n) - g_m(x_i^n)| + |g_m(x_i^n) - g_m(x)|$$
$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon;$$

thus (g_k) is uniformly Cauchy and hence uniformly convergent as $\mathcal{C}(K)$ is complete. ☒

Remarks: • Compactness of K is necessary here; for example if $\mathcal{F} = f$ (i.e. a constant sequence) then equicty is equivalent to uniform cty, however every affine function $f(x) = ax+b$ on $\mathbb{R}$ is uniformly cts but certainly <u>not</u> uniformly bounded!

• We saw equicty is necessary as $g = \left(\frac{x^2}{x^2+(1-nx)^2}\right)$ is uniformly, hence pointwise, bdd but has no uniformly convergent subsequence.

• Pointwise boundedness is necessary here; for example $\mathcal{H} = (n)$ are equicts on $[0,1]$ but not pointwise bounded (hence not uniformly bdd) and certainly has no convergent subsequence (let alone uniformly convergent)!

The Arzelà-Ascoli theorem has deep implications in the study of differential equations and its proof involves a combination of almost all of the concepts we have seen in this course. As a concrete application we have the following:

We saw that (f_n) diff'ble with (f_n') uniformly bdd on $[0,1]$ were equicts by the MVT (as $|f(x)-f(y)| \leq |x-y| \sup_{z\in[0,1]} |f'(z)|$). If the (f_n) are uniformly <u>Hölder</u> cts on a compact metric space (X,d); i.e. if (f_n) are $\mathbb{C}$ valued and for some $\alpha \in (0,1)$ and $M > 0$ we have

$$|f_n(x) - f_n(y)| \leq M\, d(x,y)^\alpha \text{ for } n \geq 1, x,y \in X,$$

then (f_n) is equicts (if $\varepsilon > 0$ let $\delta = \left(\frac{\varepsilon}{M}\right)^\alpha \Rightarrow M d(x,y)^\alpha < \varepsilon$) and so Arzela-Ascoli applies if (f_n) is pointwise bounded to guarantee that (f_n) is both uniformly bounded <u>and</u> contains a uniformly convergent subsequence!

[We refer to <u>the α-Hölder space</u> $C^{0,\alpha}(X)$ as all α-Hölder cts fns on X, thus if X is compact then $C^{0,\alpha}(X) \hookrightarrow C(X)$ is a so called '<u>compact embedding</u>' as every bdd seq. has a convergent subseq.]

[End L24] [End of examinable material]

Construction of $\mathbb{R}$

Early on we stated the following existence theorem:

Theorem: There exists an ordered field, $\mathbb{R}$, which has the least upper bound property. Moreover, $\mathbb{Q} \subset \mathbb{R}$.

We used this without proof throughout the course, we will now prove it using the so called Cauchy Completion of $\mathbb{Q}$; one can also equivalently construct $\mathbb{R}$ from $\mathbb{Q}$ by use of Dedekind cuts as is done in Rudin chapter 1.

Proof: We consider the set, C, of all Cauchy sequences in $\mathbb{Q}$; recall that every Cauchy sequence is bounded. The set C satisfies all of the field axioms, except for the existence of multiplicative inverses. Precisely, if $(x_n), (y_n) \in C$ then we define $+, \cdot$ on C by

$$(x_n) + (y_n) = (x_n + y_n) \text{ and } (x_n) \cdot (y_n) = (x_n \cdot y_n),$$

(defining $\mathbb{R}$) the boundedness of $(x_n)(y_n) \Rightarrow (x_n \cdot y_n)$ is Cauchy in particular, with $0 = (0) \in C$, $1 = (1) \in C$ are the additive and multiplicative identities, and $(-x_n) = -(x_n)$. We do not have multiplicative inverses with this operation since for example

$$(1, 0, \ldots) \cdot (0, 1, 0, \ldots) = (0).$$

We resolve the lack of multiplicative inverses by identifying sequences whose terms difference goes to zero; namely we say that $(x_n) \sim (0)$ ((x_n) equivalent to (0)) if $\lim_{n \to \infty} |x_n| = 0$ and $(x_n) \sim (y_n)$ if $(x_n - y_n) \sim 0$. The equivalence classes

$$[(x_n)] = \{ (y_n) \in C \mid (x_n) \sim (y_n) \}$$

can then be added/multiplied (as $(x_n) \sim (0) \Rightarrow (x_n) + (y_n) \sim (y_n)$ and $(x_n)(y_n) \sim (0)$). Also if $[(x_n)] \neq (0)$ then $\lim_{n \to \infty} |x_n| > 0$ and so for some $N \in \mathbb{N}$ we have $|x_n| > 0$ for $n \geq N$; setting $y_n = 0$ for $n < N$ and $y_n = x_n^{-1}$ for $n \geq N$ we have $[(x_n)(y_n)] = [(1)]$. Noting that $[(1)] \neq [(0)]$ as $\lim_{n \to \infty} |1 - 0| = 1 > 0$ the set of equivalence classes of Cauchy sequences forms a field; we define this to be $\mathbb{R}$.

$\mathbb{Q} \subset \mathbb{R}$ + ordered

We can view $\mathbb{Q} \subset \mathbb{R}$ by identifying $x \in \mathbb{Q}$ with the constant sequence $(x) \in \mathcal{C}$ so that $[(x)] \in \mathbb{R}$. We need to show that $\mathbb{R}$ is ordered and satisfies the least upper bound property. We extend the absolute value, $|\cdot|$, on $\mathbb{Q}$ to $\mathbb{R}$ by setting

$$|[(x_n)]| = [(|x_n|)] \text{ for all } [(x_n)] \in \mathbb{R}.$$

One can then check that $\mathbb{R}_{>0} = \{x \in \mathbb{R} \setminus 0 \mid |x| = x\}$ contains $\mathbb{Q}_{>0} = \{x \in \mathbb{Q} \setminus \{0\} \mid |x| = x\}$ by the above map (similarly for <0) and hence $\mathbb{R}$ is ordered (and this order agrees with that of $\mathbb{Q}$); we say that $x > y$ for $x, y \in \mathbb{R}$ if $x - y \in \mathbb{R}_{>0}$. This also shows that $\mathbb{R}$ satisfies the Archimedian property!

$\mathbb{R}$ is Complete

To see that $\mathbb{R}$ has the least upper bound property we will first show that $\mathbb{R}$ is complete wrt $|\cdot|$; namely if (x_n) is Cauchy in $\mathbb{R}$ then $x_n \to x$ for some $x \in \mathbb{R}$. First we see that every Cauchy sequence $(x_n) \subset \mathbb{Q}$ converges to $x = [(x_n)] \in \mathbb{R}$ since if $\varepsilon > 0$ then for some $N \in \mathbb{N}$ we have $|x_n - x_m| < \varepsilon$, and so by definition if $n \geq N$ then we have that $|x - x_N| = |[(x_n - x_N)]| = [|(x_n - x_N)|] < \varepsilon$; hence $x_n \to x$.
Next, we have that $\mathbb{Q}$ is dense in $\mathbb{R}$ since if $x \in \mathbb{R}$ and $\varepsilon \in \mathbb{Q}_{>0}$ then $x = [(x_n)]$ and so there is some $N \in \mathbb{N}$ such that we have $|x - x_N| \leq \varepsilon$ where $x_N \in \mathbb{Q}$ (same reasoning as above as (x_n) is Cauchy in $\mathbb{Q}$). Now, if $(x_n) \subset \mathbb{R}$ is Cauchy by the density of $\mathbb{Q}$ in $\mathbb{R}$ we can find $(r_n) \subset \mathbb{Q}$ such that $|x_n - r_n| < \frac{1}{n}$ for all $n \geq 1$. The sequence $(r_n) \subset \mathbb{Q}$ is then Cauchy by the Δ ineq. and hence by the reasoning above $r_n \to [(r_n)] \in \mathbb{R}$. By construction $\lim_{n \to \infty} |x_n - r_n| = 0$ and so $(x_n - r_n) \sim (0)$, hence applying the Δ ineq. again we have that $x_n \to [(r_n)]$ also.

Finally, if $E \subset \mathbb{R}$ is nonempty and bounded above by $y_0 \in \mathbb{R}$, let $x_0 \in \mathbb{R}$ be any non-upper bound for E. We then recursively define

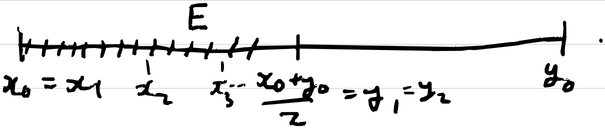
$$y_{n+1} = \begin{cases} \frac{y_n + x_n}{2}, & \text{if } \frac{y_n + x_n}{2} \text{ is an upper bound for } E, \\ y_n, & \text{otherwise.} \end{cases}$$

$\mathbb{R}$ has l.u.b property

and

$$x_{n+1} = \begin{cases} \frac{y_n + x_n}{2}, & \text{if } \frac{y_n+x_n}{2} \text{ is a non-upper bound for } E, \\ x_n, & \text{otherwise.} \end{cases}$$

Visually,



We thus have two sequences $(x_n), (y_n) \subset \mathbb{R}$ where

$$x_n \leq x_{n+1}, \quad y_n \geq y_{n+1}, \text{ and } \quad x_n \leq y_n \text{ for all } n \geq 1,$$

and inductively $|y_n - x_n| \leq 2^{-n}|y_0 - x_0|$. Moreover both are Cauchy as for $m \leq n$ we have

$$\begin{aligned} |y_m - y_n| &= |y_m - y_{m+1} + y_{m+1} - \dots + y_{n-1} - y_n| \\ &\leq \left|\frac{y_m - x_m}{2}\right| + \dots + \left|\frac{y_{n-1} - x_{n-1}}{2}\right| \\ &\leq \left(2^{-m-1} + \dots + 2^{-n}\right)(y_0 - x_0) \\ &= 2^{-m}\left(1 - \frac{1}{2^{m-n+1}}\right)(y_0 - x_0) \leq 2^{-m}(y_0 - x_0); \end{aligned}$$

and similarly for $|x_m - x_n| \leq 2^{-m}(y_0 - x_0)$. As $(x_n), (y_n)$ are Cauchy and $\mathbb{R}$ is complete, hence they converge to some limit, s (as $|y_n - x_n| \leq 2^{-n}|y_0 - x_0|$ their limits are the same). By construction each y_n is an upper bound for E and so $x \leq s$ for all $x \in E$ (else $y_n \not\to s$) and similarly as each x_n is not an upper bound for E we have that $x_n \leq u$ for any upper bound u of E, hence $s \leq u$. Therefore we see that s is the least upper bound for E; as E was arbitrary we see that $\mathbb{R}$ has the least upper bound property. ∎

The statement that $\mathbb{R}$ has the least upper bound property is often called the <u>axiom of completeness</u> which in the proof we saw followed from the completeness of $\mathbb{R}$ (and Archimedean property). It is also equivalent to monotone convergence theorem, nested intersection property (+AP), Bolzano-Weierstrass theorem, IVT, and the fact every infinite decimal sequence converges. $\boxed{\text{End L25}}$