Part IB — Analysis and Topology

Based on lectures by Dr P. Russell

Michaelmas 2022

Contents

Part I
Generalizing continuity and convergence

§1 Three Examples of Convergence

§1.1 Convergence in $\mathbb{R}$

Let (x_n) be a sequence in $\mathbb{R}$ and $x \in \mathbb{R}$. We say (x_n) *converges* to x and write $x_n \to x$ if

$$\forall\, \epsilon > 0 \quad \exists\, N \quad \forall\, n \geq N \quad |x_n - x| < \epsilon.$$

Useful fact: $\forall\, a, b \in \mathbb{R}$ $|a + b| \leq |a| + |b|$ (Triangle Inequality).

Bolzano-Weierstrass Theorem (BWT) A bounded sequence in $\mathbb{R}$ must have a convergent subsequence (Proof by interval bisection).

Recall: A sequence (x_n) in $\mathbb{R}$ is Cauchy if

$$\forall\, \epsilon > 0 \quad \exists\, N \quad \forall\, m, n \geq N \quad |x_m - x_n| < \epsilon.$$

Exercise 1.1 (Easy)**.** Show convergent $\implies$ Cauchy.

General Principle of Convergence (GPC) Any Cauchy sequence in $\mathbb{R}$ converges.

> *Outline.* If (x_n) Cauchy then (x_n) bounded so by BWT has a convergent subsequence, say $x_{n_j} \to x$. But as (x_n) Cauchy, $x_n \to x$. $\square$

§1.2 Convergence in $\mathbb{R}^2$

Remark 1. This all works in $\mathbb{R}^n$

Let (z_n) be a sequence in $\mathbb{R}^2$ and $z \in \mathbb{R}^2$. What should $z_n \to z$ mean?

In $\mathbb{R}$: "As n gets large, z_n gets arbitrarily close to z."

What does 'close' mean in $\mathbb{R}^2$?

In $\mathbb{R}$: a, b close if $|a - b|$ small. In $\mathbb{R}^2$: Replace $|\cdot|$ by $\|\cdot\|$

Recall: If $z = (x, y)$ then $\|z\| = \sqrt{x^2 + y^2}$.

Triangle Inequality If $a, b \in \mathbb{R}^2$ then $\|a + b\| \leq \|a\| + \|b\|$.

Definition 1.1

Let (z_n) be a sequence in $\mathbb{R}^2$ and $z \in \mathbb{R}^2$. We say (z_n) **converges** to z and write $z_n \to z$ if $\forall\, \epsilon > 0\ \exists\, N\ \forall\, n \geq N\ \ \|z_n - z\| < \epsilon$.

Equivalently, $z_n \to z$ iff $\|z_n - z\| \to 0$ (convergence in $\mathbb{R}$).

Example 1.1

Let $(z_n), (w_n)$ be sequences in $\mathbb{R}^2$ with $z_n \to z, w_n \to w$. Then $z_n + w_n \to z + w$.

Proof.

$$\begin{aligned}\|(z_n + w_n) - (z + w)\| &\leq \|z_n - z\| + \|w_n - w\| \\ &\to 0 + 0 = 0 \text{ (by results from IA).}\end{aligned}$$

□

In fact, given convergence in $\mathbb{R}$, convergence in $\mathbb{R}^2$ is easy:

Proposition 1.1

Let (z_n) be a sequence in $\mathbb{R}^2$ and let $z \in \mathbb{R}^2$. Write $z_n = (x_n, y_n)$ and $z = (x, y)$. Then $z_n \to z$ iff $x_n \to x$ and $y_n \to y$.

Proof. ($\implies$): $|x_n - x|, |y_n - y| \leq \|z_n - z\|$. So if $\|z_n - z\| \to 0$ then $|x_n - x| \to 0$ and $|y_n - y| \to 0$.

($\impliedby$): If $|x_n - x| \to 0$ and $|y_n - y| \to 0$ then $\|z_n - z\| = \sqrt{(x_n - x)^2 + (y_n - y)^2} \to 0$ by results in $\mathbb{R}$. □

Definition 1.2 (Bounded Sequence)

A sequence (z_n) in $\mathbb{R}^2$ is **bounded** if $\exists\, M \in \mathbb{R}$ s.t. $\forall\, n$ $\|z_n\| \leq M$.

Theorem 1.2 (BWT in $\mathbb{R}^2$)

A bounded sequence in $\mathbb{R}^2$ must have a convergent subsequence.

Theorem 1.3 (GPC for $\mathbb{R}^2$)

Any Cauchy sequence in $\mathbb{R}^2$ converges.

Proof. Let (z_n) be a Cauchy sequence in $\mathbb{R}^2$. Write $z_n = (x_n, y_n)$. For all m, n, $|x_m - x_n| \leq \|z_m - z_n\|$ so (x_n) is a Cauchy sequence in $\mathbb{R}$, so converges by GPC. Similarly, (y_n) converges in $\mathbb{R}$. So by 1.1, (z_n) converges. □

Thought for the day What about continuity? Let $f : \mathbb{R}^2 \to \mathbb{R}$. What does it mean for f to be continuous? (Simple modification of defn for $\mathbb{R} \to \mathbb{R}$).

What can we do with it?

Big theorem in IA: If $f : \mathbb{R} \to \mathbb{R}$ is a continuous function on a closed bounded interval then f is bounded and attains its bounds.

Is there a similar theorem for $\mathbb{R}^2 \to \mathbb{R}$. What do we replace 'closed bounded interval' by? We proved the theorem using BWT. Why did it work? Why did we need a closed bounded interval to make it work? What can we do in $\mathbb{R}^2$?

§1.3 Convergence of Functions

Let $X \subset \mathbb{R}$[1], let $f_n : X \to \mathbb{R}$ $(n \geq 1)$ and let $f : X \to \mathbb{R}$. What does it mean for f_n to converge to f.

Obvious idea:

Definition 1.3 (Pointwise convergence)

Say (f_n) **converges pointwise** to f and write $f_n \to f$ pointwise if $\forall\, x \in X$ $f_n(x) \to f(x)$ as $n \to \infty$.

Pros

- Simple
- Easy to check
- Defined in terms of convergence in $\mathbb{R}$

Cons

- Doesn't preserve 'nice' properties.
- 'Doesn't feel right'.

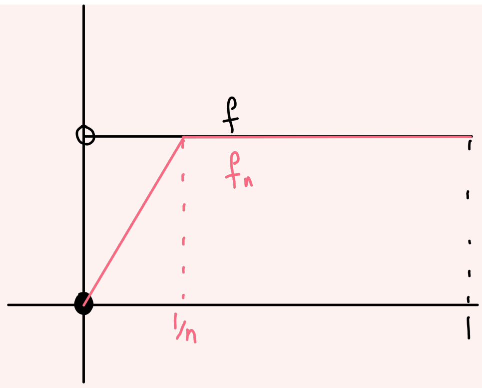
In all three examples, have $X = [0, 1]$, $f_n \to f$ pointwise.

Example 1.2 (Every f_n continuous but f not)

$$f_n(x) = \begin{cases} nx & x \leq \frac{1}{n} \\ 1 & x \geq \frac{1}{n} \end{cases}$$

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

[1] Mostly can think of $X = \mathbb{R}$ or some interval



Clearly f_n continuous for all n but f not. If $x = 0$, $\forall\, n$ $f_n(0) = 0 = f(0)$. If $x > 0$, for sufficiently large n $f_n(x) = 1 = f(x)$ so $f_n(x) \to f(x)$.

Example 1.3 (Every f_n integrable but f not)

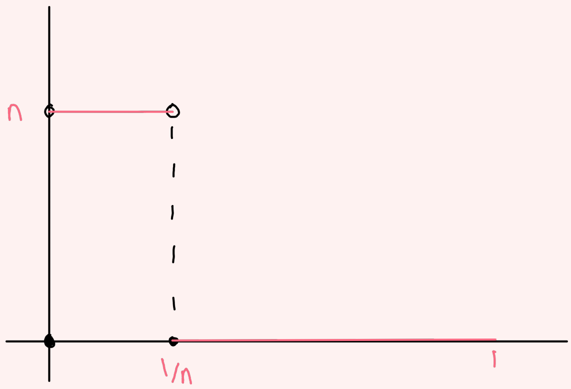
$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}.$$

This is a non integrable[a] function so now we want to find f_n such that they converge pointwise to this. Enumerate the rationals in $[0, 1]$ as $q_1, q_2, \ldots$ For $n \geq 1$, set $f_n(x) = \mathbb{1}_{q_1,\ldots,q_n}$. f_n integrable as it is nonzero at finitely many points.

[a] N.B. As in IA 'integrable' means 'Riemann integrable'

Example 1.4 (Every f_n and f integrable but $\int_0^1 f_n \not\to \int_0^1 f$)

Let $f(x) = 0$ for all x, so $\int_0^1 f = 0$. Define f_n s.t. $\int_0^1 f_n = 1$ for all n.



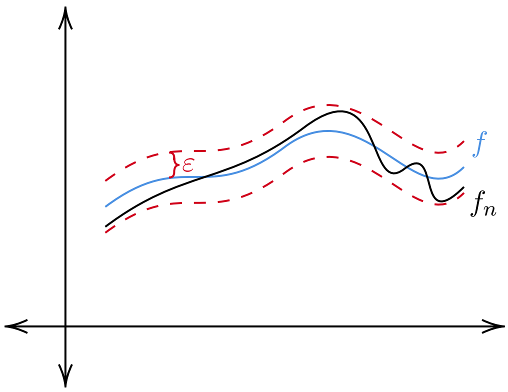
$$f_n(x) = \begin{cases} n & 0 < x < \frac{1}{n} \\ 0 & \text{otherwise} \end{cases}.$$

Better definition:

Definition 1.4 (Uniform convergence)

Let $X \subset \mathbb{R}$, $f_n : X \to \mathbb{R}$ $(n \geq 1)$, $f : X \to \mathbb{R}$. We say (f_n) **converges uniformly** to f and write $f_n \to f$ uniformly if $\forall\, \epsilon > 0\ \exists\, N\ \forall\, x \in X\ \forall\, n \geq N\ |f_n(x) - f(x)| < \epsilon$.

cf $f_n \to f$ pointwise: $\forall\, \epsilon > 0\ \forall\, x \in X\ \exists\, N\ \forall\, n \geq N\ |f_n(x) - f(x)| < \epsilon$. (We have swapped the $\forall\, x \in x$ and $\exists\, N$). Pointwise convergence allows for N to be a function of x whilst uniform convergence requires N to work for all x even the worst case. In particular, $f_n \to f$ uniformly $\implies f_n \to f$ pointwise.



Equivalently, $f_n \to f$ uniformly if for sufficiently large n $f_n - f$ is bounded and $\sup_{x \in X} |f_n - f| \to 0$.

Theorem 1.4 (A uniform limit of cts functions is cts)

Let $X \subset \mathbb{R}$, let $f_n : X \to \mathbb{R}$ be continuous $(n \geq 1)$ and let $f_n \to f : X \to \mathbb{R}$ uniformly. Then f is cts.

Proof. Let $x \in X$. Let $\epsilon > 0$. As $f_n \to f$ uniformly, we can find N s.t. $\forall\, n \geq N\ \forall\, y \in X\ |f_n(y) - f(y)| < \epsilon$. In particular, $\forall\, y \in X\ |f_N(y) - f(y)| < \epsilon$. As f_N is cts, we can find $\delta > 0$ s.t. $\forall\, y \in X$, $|y - x| < \delta \implies |f_N(y) - f_N(x)| < \epsilon$. Now let $y \in X$ with $|y - x| < \delta$. Then

$$|f(y) - f(x)| \leq |f(y) - f_N(y)| + |f_N(y) - f_N(x)| + |f_N(x) - f(x)|^{a}$$

$$< \epsilon + \epsilon + \epsilon = 3\epsilon.$$

Hence f is cts. $\square$

[a]The core of this proof is this inequality.

Remark 2. This is often called a '3ϵ proof' (or an $\frac{\epsilon}{3}$ proof).

Theorem 1.5

Let $f_n : [a, b] \to \mathbb{R}$ $(n \geq 1)$ be integrable and let $f_n \to f : [a, b] \to \mathbb{R}$ uniformly. Then f is integrable and $\int_a^b f_n \to \int_a^b f$ as $n \to \infty$.

Proof. As $f_n \to f$ uniformly, we can pick n suff. large s.t. $f_n - f$ is bounded. Also f_n is bounded (as integrable). So by triangle inequality, $f = (f - f_n) + f_n$ is bounded. Let $\epsilon > 0$. As $f_n \to f$ uniformly there is some N s.t. $\forall\, n \geq N$ $\forall\, x \in [a, b]$ we have $|f_n(x) - f(x)| < \epsilon$.
In particular, $\forall\, x \in [a, b]$ $|f_N(x) - f(x)| < \epsilon$.

By Riemann's criterion, there is some dissection $\mathcal{D}$ of $[a, b]$ for which $S(f_n, \mathcal{D}) - s(f_n, \mathcal{D}) < \epsilon$. Let $\mathcal{D} = \{x_0, x_1, x_2, \dots, x_k\}$ where $a = x_0 < x_1 < \cdots < x_k = b$. Now

$$\begin{aligned}
S(f, \mathcal{D}) &= \sum_{i=1}^{k} (x_i - x_{i-1}) \sup_{x \in [x_{i-1}, x_i]} f(x) \\
&\leq \sum_{i=1}^{k} (x_i - x_{i-1}) \sup_{x \in [x_{i-1}, x_i]} (f_N(x) + \epsilon) \\
&= \sum_{i=1}^{k} (x_i - x_{i-1}) \left(\left(\sup_{x \in [x_{i-1}, x_i]} f_N(x) \right) + \epsilon \right) \\
&= \sum_{i=1}^{k} (x_i - x_{i-1}) \sup_{x \in [x_{i-1}, x_i]} f_N(x) + \sum_{i=1}^{k} (x_i - x_{i-1}) \epsilon \\
&= S(f_N, \mathcal{D}) + (b - a)\epsilon.
\end{aligned}$$

That is $S(f, \mathcal{D}) \leq S(f_N, \mathcal{D}) + (b - a)\epsilon$. Similarly $s(f, \mathcal{D}) \geq s(f_N, \mathcal{D}) - (b - a)\epsilon$. Hence

$$\begin{aligned}
S(f, \mathcal{D}) - s(f, \mathcal{D}) &\leq S(f_N, \mathcal{D}) - s(f_N, \mathcal{D}) + 2(b - a)\epsilon \\
&< (2(b - a) + 1)\epsilon
\end{aligned}$$

But $2(b - a) + 1$ is a constant so $(2(b - a) + 1)\epsilon$ can be made arbitrarily small. Hence by Riemann's criterion, f is integrable over $[a, b]$.

Now, for any n suff. large that $f_n - f$ is bounded,

$$\begin{aligned}\left|\int_a^b f_n - \int_a^b f\right| &= \left|\int_a^b (f_n - f)\right| \\ &\le \int_a^b |f_n - f| \\ &\le (b-a) \sup_{x\in[a,b]} |f_n - f| \\ &\to 0 \text{ as } n \to \infty \text{ since } f_n \to f \text{ uniformly.}^{a}\end{aligned}$$

□

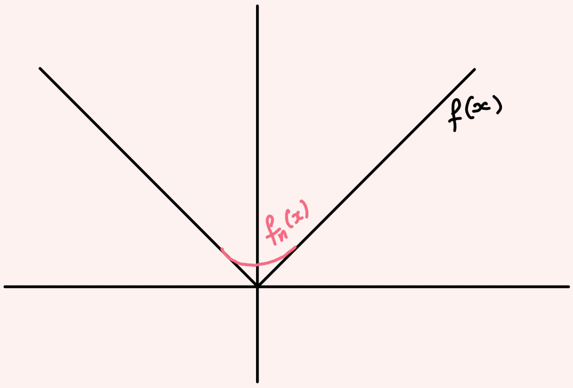
[a] Note we said that $f_n \to f$ uniformly if $\sup |f_n - f| \to 0$.

What about differentiation? Here even uniform convergence isn't enough.

Example 1.5

$f_n : (-1,1) \to \mathbb{R}$, each f_n differentiable, $f_n \to f$ uniformly, f not diff.

Let $f(x) = |x|$ which is not differentiable at 0.



$$f_n = \begin{cases} |x| & |x| \ge \frac{1}{n} \\ ax^2 + bx + c & |x| < \frac{1}{n} \end{cases}.$$

We need $a(\frac{1}{n})^2 + \frac{b}{n} + c = \frac{1}{n}$ for continuity. Thus $b = 0$ and $c = \frac{1}{n} - \frac{a}{n^2}$.

Also need $2a\frac{1}{n} + b = 1$ and $2a(-\frac{1}{n}) = -1$ for differentiability so take $a = \frac{n}{2}$, $c = \frac{1}{n} - \frac{1}{2n} = \frac{1}{2n}$.

If $|x| \geq \frac{1}{n}$ then $|f_n(x) - f(x)| = 0$. If $|x| < \frac{1}{n}$:

$$\begin{aligned} |f_n(x) - f(x)| &= \left| \frac{n}{2}x^2 + \frac{1}{2n} - |x| \right| \\ &\leq \frac{n}{2}x^2 + \frac{1}{2n} + |x| \\ &\leq \frac{n}{2}(\frac{1}{n})^2 + \frac{1}{2n} + \frac{1}{n} \\ &= \frac{1}{2n} + \frac{1}{2n} + \frac{1}{n} \\ &= \frac{2}{n} \end{aligned}$$

So $\sup_{x \in (-1,1)} |f_n(x) - f(x)| \leq \frac{2}{n} \to 0$ as $n \to \infty$. So $f_n \to f$ uniformly.

If fact we need uniform convergence of the derivatives.

Theorem 1.6

Let $f_n : (u, v) \to \mathbb{R}$ $(n \geq 1)$ with $f_n \to f : (u, v) \to \mathbb{R}$ pointwise. Suppose further each f_n is continuously differentiable and that $f_n' \to g : (u, v) \to \mathbb{R}$ uniformly. Then f is differentiable with $f' = g$.

Proof. Fix $a \in (u, v)$. Let $x \in (u, v)$, by FTC we have each f_n' is integrable over $[a, x]$ and $\int_a^x f_n' = f_n(x) - f_n(a)$. But $f_n' \to g$ uniformly so by theorem 1.5 g is integrable over $[a, x]$ and $\int_a^x g = \lim_{n\to\infty} \int_a^x f_n' = f(x) - f(a)$. So we have shown that for all $x \in (u, v)$

$$f(x) = f(a) + \int_a^x g.$$

By theorem 1.4, g is cts so by FTC, f is differentiable with $f' = g$. $\square$

Remark 3. It would have sufficed to assume $f_n(x) \to f(x)$ for a single value of x rather than $f_n \to f$ pointwise.

GPC?

Definition 1.5 (Uniform Cauchy)

Let $X \subset \mathbb{R}$ and let $f_n : X \to \mathbb{R}$ for each $n \geq 1$. We say (f_n) is **uniformly Cauchy** if $\forall\, \epsilon > 0\ \exists\, N\ \forall\, m, n \geq N\ \forall\, x \in X\ |f_m(x) - f_n(x)| < \epsilon$

Exercise 1.7. Show uniformly convergence $\implies$ uniformly Cauchy.

Theorem 1.8 (General principle of Uniform Convergence (GPUC))

Let (f_n) be a uniformly Cauchy sequence of functions $X \to \mathbb{R}$ $(X \subset \mathbb{R})$. Then (f_n) is uniformly convergent.

Proof. Let $x \in X$. Let $\epsilon > 0$. Then $\exists\, N\ \forall\, m, n \geq N\ \forall\, y \in X\ |f_m(y) - f_n(y)| < \epsilon$. In particular, $\forall\, m, n \geq N\ |f_m(x) - f_n(x)| < \epsilon$. So $(f_n(x))$ is a Cauchy sequence in $\mathbb{R}$ so by GPC it converges, say $f_n(x) \to f(x)$ as $n \to \infty$.

We have now constructed $f : X \to \mathbb{R}$ s.t. $f_n \to f$ pointwise.
Let $\epsilon > 0$. Then we can find a N s.t. $\forall\, m, n \geq N\ \forall\, y \in X\ |f_m(y) - f_n(y)| < \epsilon$. Fix $y \in X$, keep $m \geq N$ fixed and let $n \to \infty$: $|f_m(y) - f(y)| \leq \epsilon$. So we have shown that $\forall\, m \geq N$, $|f_m(y) - f(y)| < \epsilon$.

But y was arbitrary so $\forall\, x \in X\ \forall\, m \geq N\ |f_m(x) - f(x)| \leq \epsilon$. That is $f_n \to f$ uniformly. $\square$

BW?

Definition 1.6 (Pointwise bounded)

Let $X \subset \mathbb{R}$ and let $f_n : X \to \mathbb{R}$ for each $n \geq 1$. We say (f_n) is **pointwise bounded** if $\forall\, x\ \exists\, M\ \forall\, n\ |f_n(x)| \leq M$.

Definition 1.7 (Uniformly bounded)

Let $X \subset \mathbb{R}$ and let $f_n : X \to \mathbb{R}$ for each $n \geq 1$. We say (f_n) is **uniformly bounded** if $\exists\, M\ \forall\, x\ \forall\, n\ |f_n(x)| \leq M$.[a]

[a] Again we have just swapped $\forall\, x\ \exists\, M$ as in convergence.

What would uniform BW say? 'If (f_n) is a uniformly bounded sequence of functions that it has a uniformly convergent subsequence'. But this is not true.

Example 1.6 (Counterexample of BW)

$$f_n : \mathbb{R} \to \mathbb{R}$$
$$f_n(x) = \begin{cases} 1 & x = n \\ 0 & x \neq n. \end{cases}$$

Obviously (f_n) uniformly bounded (by 1). However, if $m \neq n$ then $f_m(m) = 1$ and $f_n(m) = 0$ so $|f_m(m) - f_n(m)| = 1$ so no subsequence can be uniformly Cauchy so no subsequence can be uniformly convergent.

Application to power series Recall that if $\sum a_n x^n$ is a real power series with r.o.c $R > 0$

then we can differentiate/ integrate it term-by-term within $(-R, R)$.

Definition 1.8

Let $f_n : X \to \mathbb{R}$ ($X \subset \mathbb{R}$) for each $n \geq 0$. We say the series $\sum_{n=0}^{\infty} f_n$ **uniformly converges** if the sequence of partial sums (F_n) does, where $F_n = \sum_{m=0}^{n} f_m$.

We can apply theorems 1.4 to 1.6 to get e.g. if conditions hold with f_n cts diff and uniform convergence then $\sum f_n$ has derivative $\sum f_n'$.

Hope Prove $\sum a_n x^n$ converges uniformly on $(-R, R)$ then hit it with earlier theorems.

Not quite true:

Example 1.7

$\sum_{n=0}^{\infty} x^n$ r.o.c 1. This does not converge uniformly on $(-1,1)$. Let $f(x) = \sum_{n=0}^{\infty} x^n$ and $F_n(x) = \sum_{m=0}^{n} x^m$. Note $f(x) = \frac{1}{1-x} \to \infty$ as $x \to 1$. However, $\forall\, x \in (-1,1)$ $|F_n(x)| \leq n+1$.

Fix any n. We can find a point $x \in (-1,1)$ where $f(x) \geq n+2$ and so $|f(x) - F_n(x)| \geq 1$. So we don't have uniform convergence.

Back-up plan: It does work if we look at a smaller interval.

New plan: show if $0 < r < R$ then we do have uniform convergence on $(-r, r)$.
Given $x \in (-R, R)$ there's some r with $|x| < r < R$: use uniform convergence on $(-r, r)$ to check everything is nice at x. 'Local uniform convergence of power series'.

Aside

In $\mathbb{R}$ $x_n \to 0$ if

1. $\forall\, \epsilon > 0\ \exists\, N\ \forall\, n \geq N\ |x_n| < \epsilon$.
2. Equivalently: $\forall\, \epsilon > 0\ \exists\, N\ \forall\, n \geq N\ |x_n| \leq \epsilon$.

Proof. i $\implies$ ii: obvious
ii $\implies$ ii: Let $\epsilon > 0$. Pick N s.t. $\forall\, n \geq N\ |x_n| \leq \frac{1}{2}\epsilon$. Then $\forall\, n \geq N\ |x_n| < \epsilon$. $\square$

Also: $f_n, f : X \to \mathbb{R}$, $f_n \to f$ uniformly.

1. $\forall\, \epsilon > 0\ \exists\, N\ \forall\, x \in X\ \forall\, n \geq N\ |f_n(x) - f(x)| < \epsilon$.
2. For n suff large $f_n - f$ is bounded and $\forall\, \epsilon > 0\ \exists\, N\ \forall\, n \geq N\ \sup_{x \in X} |f_n(x) - f(x)| < \epsilon$.

Proof. ii $\implies$ i: obvious
i $\implies$ ii: if i holds then $\sup_{x\in X} |f_n(x) - f(x)| \leq \epsilon$. But OK by same argument as previously. $\square$

Lemma 1.1

Let $\sum a_n x^n$ be a real power series with r.o.c $R > 0$. Let $0 < r < R$. Then $\sum a_n x^n$ converges uniformly on $(-r, r)$.

Proof. Define $f, f_n : (-r, r) \to \mathbb{R}$ by $f(x) = \sum_{n=0}^{\infty} a_n x^n$ and $f_m(x) = \sum_{n=0}^{m} a_n x^n$. Recall that $\sum a_n x^n$ converges absolutely for all x with $|x| < R$.

Let $x \in (-r, r)$. Then f

$$\begin{aligned} |f(x) - f_m(x)| &= \left| \sum_{n=m+1}^{\infty} a_n x^n \right| \\ &\leq \sum_{n=m+1}^{\infty} |a_n||x|^n \\ &\leq \sum_{n=m+1}^{\infty} |a_n| r^n \end{aligned}$$

which converges by absolute convergence at r. Hence if m suff large, $f - f_m$ is bounded and

$$\sup_{x\in(-r,r)} |f(x) - f_m(x)| \leq \sum_{n=m+1}^{\infty} |a_n| r^n \to 0$$

as $m \to \infty$ by absolute convergence of r. $\square$

Theorem 1.9

Let $\sum a_n x^n$ be a real power series with r.o.c $R > 0$. Define $f : (-R, R) \to \mathbb{R}$ by $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Then

1. f is continuous;
2. for any $x \in (-R, R)$ f is integrable over $[0, x]$ with

$$\int_0^x f = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}.$$

Proof. Let $x \in (-R, R)$. Pick r s.t. $|x| < r < R$. By lemma 1.1, $\sum a_n y^n$ converges uniformly on $(-r, r)$. But the partial sum functions $y \mapsto \sum_{n=0}^m a_n y^n$ $(m \geq 0)$ are all cts functions on $(-r, r)$ (as they are polynomials). Hence by theorem 1.4, $f\,|_{(-r,r)}$[a] is cts. Hence f is cts at x, but x was arbitrary so f is a cts fcn on $(-R, R)$.

Moreover, $[0, x] \subset (-r, r)$ so we also have $\sum a_n y^n$ converges uniformly on $[0, x]$. Each partial sum function on $[0, x]$ is a poly so can be integrated with $\int_0^x \sum_{n=0}^m a_n y^n \, dy = \sum_{n=0}^m \int_0^x a_n y^n \, dy = \sum_{n=0}^m \frac{a_n}{n+1} x^{n+1}$. Hence by theorem 1.5, f is integrable over $[0, x]$ with

$$\begin{aligned} \int_0^x f &= \lim_{m \to \infty} \int_0^x \sum_{n=0}^m a_n y^n \, dy \\ &= \lim_{m \to \infty} \sum_{n=0}^m \frac{a_n}{n+1} x^{n+1} \\ &= \sum_{n=0}^\infty \frac{a_n}{n+1} x^{n+1}. \end{aligned}$$

□

[a] f restricted to domain $(-r, r)$

For differentiation, need technical lemma:

Lemma 1.2

Let $\sum a_n x^n$ be a real power series with r.o.c $R > 0$. Then the power series $\sum_{n \geq 1} n a_n x^{n-1}$ has r.o.c at least R.

Proof. Let $x \in \mathbb{R}$ with $0 < x < R$. Pick w with $x < w < R$. Then $\sum a_n w^n$ is absolutely convergent, so $a_n w^n \to 0$ (terms of a convergent series) so $\exists$ M s.t. $\forall$ n, $|a_n w^n| \leq M$.

For each n,

$$|n a_n x^{n-1}| = |a_n w^n| \left|\frac{x}{w}\right|^n \frac{1}{|x|} n.$$

Fix n. Let $\alpha = |\frac{x}{w}| < 1$. Let $c = \frac{M}{|x|}$, a constant. Then $|n a_n x^{n-1}| \leq c n \alpha^n$. By comparison test, ETS (enough to show) $\sum n \alpha^n$ converges.
Note $\left|\frac{(n+1)\alpha^{n+1}}{n \alpha^n}\right| = (1 + \frac{1}{n})\alpha \to \alpha < 1$ as $n \to \infty$ so done by ratio test. □

Theorem 1.10

Let $\sum a_n x^n$ be a real power series with r.o.c. $R > 0$. Let $f : (-R, R) \to \mathbb{R}$ be defined by $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Then f is differentiable and $\forall\ x \in (-R, R)$ $f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$.

Proof. Let $x \in (-R, R)$. Pick r with $|x| < r < R$. Then $\sum a_n y^n$ converges uniformly on $(-r, r)$. Moreover, the power series $\sum_{n \geq 1} n a_n y^{n-1}$ has r.o.c at least R and so also converges uniformly on $(-r, r)$.

The partial sum functions $f_m(y) = \sum_{n=0}^{m} a_n y^n$ are polys so differentiable with $f_m'(y) = \sum_{n=1}^{m} n a_n y^{n-1}$. We now have f_m' converging uniformly on $(-r, r)$ to the function $g(y) = \sum_{n=1}^{\infty} n a_n y^{n-1}$.

Hence by theorem 1.6, $f\mid_{(-r,r)}$ is differentiable and $\forall\ y \in (-r, r)$ $f'(y) = g(y)$.

In particular, f is differentiable at x with $f'(x) = g(x)$. Hence f is a differentiable function on $(-R, R)$ with derivative g as desired. □

§1.4 Uniform Continuity

Let $X \subset \mathbb{R}$. Let $f : X \to \mathbb{R}$. (May as well think of $X = \mathbb{R}$ or $X = (a, b)$).

Definition 1.9 (Continuous function)

f is **continuous** if

$$\forall\ \epsilon > 0\ \forall\ x \in X\ \exists\ \delta > 0\ \forall\ y \in X\ |x - y| < \delta \implies |f(x) - f(y)| < \epsilon.$$

Definition 1.10 (Unifomly Continuous function)

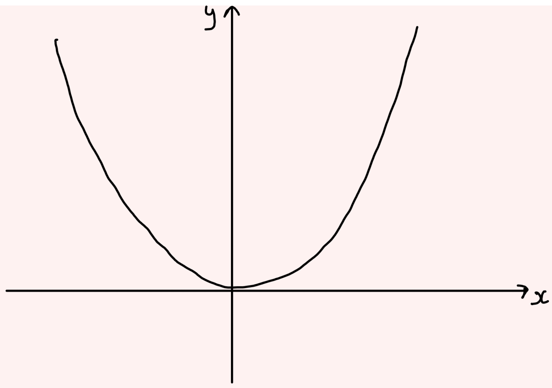
f is **uniformly continuous** if

$$\forall\ \epsilon > 0\ \exists\ \delta > 0\ \forall\ x \in X\ \forall\ y \in X\ |x - y| < \delta \implies |f(x) - f(y)| < \epsilon.$$

Remark 4. Clearly if f is uniformly cts then f is cts. We would suspect that f cts doesn't imply f uniformly cts.

Example 1.8

A function $f : \mathbb{R} \to \mathbb{R}$ that is cts but not uniformly cts.



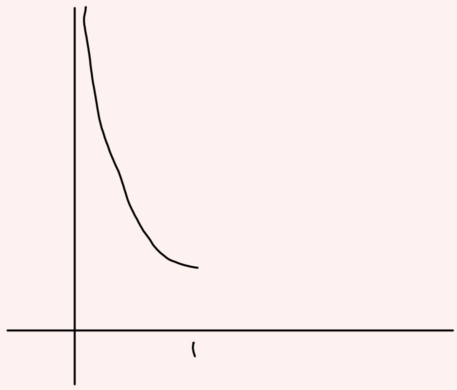
We want some function that looks like this, a continuous function which gets steeper as we go to infinity. So $f(x) = x^2$ ought to work. We know f is cts (as it's a poly). Suppose $\delta > 0$. Then

$$\begin{aligned} f(x+\delta) - f(x) &= (x+\delta)^2 - x^2 \\ &= 2\delta x + \delta^2 \to \infty \text{ as } x \to \infty. \end{aligned}$$

So in particular, $\forall\ \delta > 0\ \exists\ x, y \in \mathbb{R}$ s.t. $|x - y| < \delta$ but $|f(x) - f(y)| \geq 1$. So conditions for uniform cty fails for $\epsilon = 1$. So f not uniform cty.

Example 1.9

Make domain bounded. We can still fail, e.g. $f : (0, 1) \to \mathbb{R}$ cts but not uniform cts.



Let $f(x) = \frac{1}{x}$, clearly cts. Proof that its not uniform continuity is left as an exercise to the reader.

Theorem 1.11

A continuous real-valued function on a closed bounded interval is uniformly continuous.

Proof. Let $f : [a, b] \to \mathbb{R}$ and suppose f is cts but not uniformly cts. Then we can find $\epsilon > 0$ st. $\forall\, \delta > 0\ \exists\, x, y \in [a, b]$ with $|x - y| < \delta$ but $|f(x) - f(y)| \geq \epsilon$.

In particular, taking $\delta = \frac{1}{n}$ we can find sequences $(x_n), (y_n) \in [a, b]$ with, for each n, $|x_n - y_n| < \frac{1}{n}$ but $|f(x_n) - f(y_n)| \geq \epsilon$. The sequence (x_n) is bounded so by BW[a] it has a convergent subsequence $x_{n_j} \to x$. And $[a, b]$ is a closed interval so $x \in [a, b]$. Then $x_{n_j} - y_{n_j} \to 0$ so $y_{n_j} \to x$.

But f is cts at x so $\exists\, \delta > 0$ s.t. $\forall\, y \in [a, b]\ |y - x| < \delta \implies |f(y) - f(x)| < \frac{\epsilon}{2}$. Take such a δ. As $x_{n_j} \to x$ we can find J_1 s.t. $j \geq J_1 \implies |x_{n_j} - x| < \delta$. Similarly we can find J_2 s.t. $j \geq J_2 \implies |y_{n_j} - x| < \delta$. Now let $j = \max(J_1, J_2)$ then $|x_{n_j} - x|, |y_{n_j} - x| < \delta$ so we have $|f(x_{n_j}) - f(x)|, |f(y_{n_j}) - f(x)| < \epsilon/2$. Then $|f(x_{n_j}) - f(y_{n_j})| \leq |f(x_{n_j}) - f(x)| + |f(y_{n_j}) - f(x)| < \epsilon$ ↯. □

[a]Bolzano Weierstrass

Corollary 1.1

A continuous real-valued function on a closed bounded interval is bounded.

Proof. Let $f : [a, b] \to \mathbb{R}$ be a continuous function, and so uniformly continuous by theorem 1.11. Then we can find $\delta > 0$ s.t. $\forall\, x, y \in [a, b]\ |x - y| < \delta \implies |f(x) - f(y)| < 1$.

Let $M = \lceil \frac{b-a}{\delta} \rceil$. Let $x \in [a, b]$. We can find $a = x_0 \leq x_1 \leq \cdots \leq x_M = x$ with $|x_i - x_{i-1}| < \delta$ for each i. Hence

$$\begin{aligned} |f(x)| &= \left| f(a) + \sum_{i=1}^{M} f(x_i) - f(x_{i-1}) \right| \\ &\leq |f(a)| + \sum_{i=1}^{M} |f(x_i) - f(x_{i-1})| \\ &< |f(a)| + \sum_{i=1}^{M} 1 \\ &= |f(a)| + M. \end{aligned}$$

□

Remark 5. Referring back to example 1.9, starting at $x = 1$ and going towards $x = 0$ we can that δ gets smaller and smaller s.t. you require an infinite number of steps to get 0. So $M = \infty$ essentially.

Corollary 1.2

A continuous real-valued function on a closed bounded interval is integrable.

Proof. Let $f : [a, b] \to \mathbb{R}$ be a continuous function, and so uniformly continuous by theorem 1.11. Let $\epsilon > 0$. Then we can find $\delta > 0$ s.t. $\forall\, x, y \in [a, b]$ $|x - y| < \delta \implies |f(x) - f(y)| < \epsilon$. Let $\mathcal{D} = \{x_0 < x_1 < \cdots < x_n\}$ be a dissection s.t. for each i we have $x_i - x_{i-1} < \delta$.

Let $i \in \{1, \ldots, n\}$. Then for any $u, v \in [x_{i-1}, x_i]$ we have $|u-v| < \delta$ so $|f(u)-f(v)| < \epsilon$. Hence

$$\sup_{x\in[x_{i-1},x_i]} f(x) - \inf_{x\in[x_{i-1},x_i]} f(x) \leq \epsilon.$$

Hence:

$$\begin{aligned}
S(f, \mathcal{D}) - s(f, \mathcal{D}) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \left(\sup_{x\in[x_{i-1},x_i]} f(x) - \inf_{x\in[x_{i-1},x_i]} f(x) \right) \\
&\leq \sum_{i=1}^{n} (x_i - x_{i-1})\epsilon \\
&= \epsilon \sum_{i=1}^{n} (x_i - x_{i-1}) \\
&= \epsilon(b - a).
\end{aligned}$$

But $\epsilon(b - a)$ can be made arbitrarily small by taking ϵ small. So by Riemann's criterion f is integrable over $[a, b]$. □

§2 Metric Spaces

§2.1 Definitions and Examples

Question

Can we think about convergence in a more general setting? Convergence seemed similar in our 3 settings.

What do we really need?

Answer

We need a notion of distance.

In $\mathbb{R}$: distance x to y is $|x-y|$.
In $\mathbb{R}^2$: its $\|x-y\|$.
For functions: distance f to g is $\sup_{x\in X} |f(x)-g(x)|$ (where this exists, i.e. if $f-g$ bounded).

The triangle inequality was often important (see the proof of uniqueness of limits).

Definition 2.1 (Metric)

A **metric** d is a function $d : X^2 \to \mathbb{R}$ satisfying:

- $d(x,y) \geq 0$ for all $x, y \in X$ with equality iff $x = y$;
- $d(x,y) = d(y,x)$ for all $x, y \in X$.
- $d(x,z) \leq d(x,y) + d(y,z)$ for all $x, y, z \in X$.

Definition 2.2 (Metric Space)

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

We could also define a metric space as an ordered pair (X, d). If it is obvious what d is, we sometimes write 'The metric space X ...'.

Example 2.1

$X = \mathbb{R}$, $d(x,y) = |x-y|$ 'The <u>usual metric</u> on $\mathbb{R}$'.

Example 2.2

$X = \mathbb{R}^n$ with the Euclidean metric, $d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$.

Uniform convergence of functions doesn't quite work: we want $d(f, g) = \sup |f - g|$ but this might not exist if $f - g$ is unbounded. However, we can do something with appropriate sets of functions.

Example 2.3

Let $Y \subset \mathbb{R}$. Take $X = B(Y) = \{f : Y \to \mathbb{R} \mid f \text{ bounded}\}$ with the uniform metric $d(f, g) = \sup_{x \in Y} |f - g|$.

Checking triangle inequality:

Proof. Let $f, g, h \in B(Y)$. Let $x \in Y$. Then

$$\begin{aligned} |f(x) - h(x)| &\leq |f(x) - g(x)| + |g(x) - h(x)| \\ &\leq d(f, g) + d(g, h) \end{aligned}$$

Taking sup over all $x \in Y$

$$d(f, h) \leq d(f, g) + d(g, h).$$

□

Definition 2.3 (Subspace)

Suppose (X, d) a metric space and $Y \subset X$. Then $d\,|_{Y^2}$ is a metric on Y. We say Y with this metric is a **subspace** of X.

Example 2.4

Subspaces of $\mathbb{R}$: any of $\mathbb{Q}, \mathbb{Z}, \mathbb{N}, [0, 1], \ldots$ with the usual metric $d(x, y) = |x - y|$.

Example 2.5

Recall that a cts function on a closed bounded interval is bounded. Define $C([a, b]) = \{f : [a, b] \to \mathbb{R} \mid f \text{ cts}\}$. This is a subspace of $B([a, b])$, example 2.3. That is $C([a, b])$ is a metric space with the uniform metric $\mathcal{L}(f, g) = \sup_{x \in [a,b]} |f(x) - g(x)|$

Example 2.6

The empty metric space $X = \emptyset$ with the empty metric.

Could maybe define different metrics on the same set:

Example 2.7

The ℓ_1 metric on $\mathbb{R}^n$: $d(x,y) = \sum_{i=1}^n |x_i - y_i|$.

Example 2.8

The ℓ_∞ metric on $\mathbb{R}^n$: $d(x,y) = \max_i |x_i - y_i|$.[a]

[a] Proof of triangle inequality similar to example 2.3

Example 2.9

On $C([a,b])$ we can define the L_1 metric: $d(f,g) = \int_a^b |f-g|$.

Example 2.10

$X = \mathbb{C}$ with

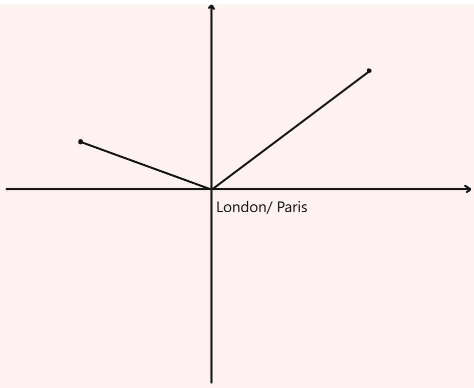
$$d(z,w) = \begin{cases} 0 & z = w \\ |z| + |w| & z \neq w. \end{cases}.$$

First two conditions of a metric hold obviously, for triangle inequality we need $d(u,w) \leq d(u,v) + d(v,w)$.

1. If $u = w$, LHS $= 0$ ✓
2. If $u = v$ or $v = w$ then LHS = RHS ✓
3. If u, v, w distinct:

$$\begin{aligned} LHS &= |u| + |w| \\ RHS &= |u| + |w| + 2|v| \checkmark \end{aligned}$$

This metric is often called the British Rail metric or SNCF metric, you can think of it as for distinct points you have to travel through the origin.



Example 2.11 (Discrete metric)

Let X be any set. Define a metric d on X by

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

Easy to check this works. This is called the <u>discrete metric</u> on X.

Example 2.12 (p-adic metric)

Let $\mathbb{X} = \mathbb{Z}$. Let p be a prime. The **p-adic metric** on $\mathbb{Z}$ is the metric d defined by:

$$d(x, y) = \begin{cases} 0 & x = y \\ p^{-a} & \text{if } x \neq y \text{ and } x - y = p^a m \text{ with } p \nmid m. \end{cases}$$

'Two numbers are close if difference is divisible by a large power of p'.

Only thing we need to check is triangle inequality

Proof. STP: $d(x,z) \le d(x,y) + d(y,z)$

1. If $x = z$, LHS $= 0$ ✓
2. If $x = y$ or $y = z$ then LHS $=$ RHS ✓

So easy if any two of x, y, z the same so assume x, y, z all distinct. Let $x - y = p^a m$ and $y - z = p^b n$ where $p \nmid m, p \nmid n$ and wlog $a \le b$. So $d(x,y) = p^{-a}$ and $d(y,z) = p^{-b}$.

Now:

$$\begin{aligned} x - z &= (x - y) = (y - z) \\ &= p^a m + p^b n \\ &= p^a \underbrace{(m + p^{b-a} n)}_{\text{integer}} \text{ as } a \le b. \end{aligned}$$

So $p^a \mid x - z$ so $d(x,z) \le p^{-a}$. But $d(x,y) + d(y,z) \ge d(x,y) = p^{-a}$. □

[a] p^a is the largest a s.t. $p^a \mid x - y$

Definition 2.4 (Convergence)

Let (X,d) be a metric space, let (x_n) be a sequence in X and let $x \in X$. We say (x_n) **converges** to x and write '$x_n \to x$' or '$x_n \to x$ as $n \to \infty$' if

$$\forall \epsilon > 0 \; \exists N \; \forall n \ge N \; d(x_n, x) < \epsilon.$$

Equivalently $x_n \to x$ iff $d(x_n, x) \to 0$ in $\mathbb{R}$.

Proposition 2.1

Limits are unique. That is, if (X,d) is a metric space, (x_n) a sequence in X, $x, y \in X$ with $x_n \to x$ and $x_n \to y$ then $x = y$.

Proof. For each n,

$$\begin{aligned} d(x,y) &\le d(x,x_n) + d(x_n,y) \text{ by triangle inequality} \\ &\le d(x_n,x) + d(x_n,y) \text{ by symmetry} \\ &\to 0 + 0 = 0 \text{ as } d(x_n,x), d(x_n,y) \to 0 \end{aligned}$$

So $d(x,y) \to 0$ as $n \to \infty$. But $d(x,y)$ is constant so $d(x,y) = 0$ so $x = y$. □

Remark 6. This justifies talking about <u>the</u> limit of a convergent sequence in a metric space, and writing $x = \lim_{n\to\infty} x_n$ if $x_n \to x$.

Remark 7 (Remarks on definition of convergence in a metric space).

1. Constant sequences obviously converge. More over, eventually constant sequences converge.
2. Suppose (X, d) is a metric space and Y is a subspace of X. Suppose (x_n) is a sequence in Y which converges in Y to x. Then also (x_n) converges in X to x.

   However, converse is false: e.g. in $\mathbb{R}$ with the usual metric then $\frac{1}{n} \to 0$ as $n \to \infty$. Consider the subspace $\mathbb{R} \setminus \{0\}$. Then $(\frac{1}{n})$ is a sequence in $\mathbb{R} \setminus \{0\}$ but it doesn't converge in $\mathbb{R} \setminus \{0\}$. (Why? Suppose $\frac{1}{n} \to x$ in $\mathbb{R} \setminus \{0\}$. Then also $\frac{1}{n} \to x$ in $\mathbb{R}$. But $\frac{1}{n} \to 0$ in $\mathbb{R}$ so by uniqueness of limits $x = 0$. But $x \in \mathbb{R} \setminus \{0\}$ and $0 \notin \mathbb{R} \setminus \{0\}$ $\lightning$.)

Example 2.13

Let d be the Euclidean metric on $\mathbb{R}^n$. Exactly as in $\mathbb{R}^2$, we have $x_n \to x$ iff the sequence converges in each coordinate in the usual way in $\mathbb{R}$.

What about other metrics on $\mathbb{R}^n$? E.g. let d_∞ be the uniform metric: $d_\infty(x, y) = \max_i |x_i - y_i|$. Which sequences converge in $(\mathbb{R}^n, d_\infty)$? $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \leq \sqrt{\sum_{i=1}^n d_\infty(x, y)^2}$ so $d(x, y) \leq \sqrt{n} d_\infty(x, y)$. But also $d_\infty(x, y) \leq d(x, y)$ as one of the terms in $d(x, y)$ is d_∞^2.

Now suppose (x_n) is a sequence in $\mathbb{R}^n$. Then $d(x_n, x) \to 0 \iff d(x_n, x) \to 0$. So exactly same sequences converge in $(\mathbb{R}^n, d)$ and $(\mathbb{R}^n, d_\infty)$

What about ℓ_1 metric d_1? $d_1(x, y) = \sum_{i=1}^n |x_i - y_i|$. Similarly, $d_\infty(x, y) \leq d_1(x, y) \leq n d_\infty(x, y)$. So again, exactly the same sequences converge in $(\mathbb{R}^n, d_1)$.

Example 2.14

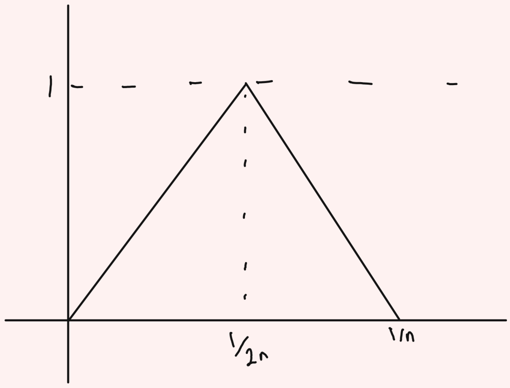
Let $X = C([0, 1]) = \{f : [0, 1] \to \mathbb{R} \mid f \text{ continuous}\}$. Let d_∞ be the uniform metric on X: $d_\infty(f, g) = \sup_{x \in [0,1]} |f(x) - g(x)|$.

$$\begin{aligned} f_n \to f \text{ in } (X, d_\infty) &\iff d_\infty(f_n, f) \to 0 \\ &\iff \sup_{x \in [0,1]} |f_n(x) - f(x)| \to 0 \\ &\iff f_n \to f \text{ uniformly.} \end{aligned}$$

We also have L_1-metric d_1 on X: $d_1(f, g) = \int_0^1 |f - g|$. Now $d_1(f, g) = \int_0^1 |f - g| \leq \int_0^1 d_\infty(f, g) = d_\infty(f, g)$. So similarly to previous example,

$$f_n \to f \text{ in } (X, d_\infty) \implies f_n \to f \text{ in } (X, d_1).$$

But converse does not hold, i.e. we can find a sequence (f_n) in X s.t. $f_n \to 0$ in d_1-metric but f_n doesn't converge in d_∞-metric, i.e. $\int_0^1 |f_n| \to 0$ as $n \to \infty$ but (f_n) does not converge uniformly.



$$f_n(x) = \begin{cases} 2nx & x \leq \frac{1}{2n} \\ 2n(\frac{1}{n} - x) & \frac{1}{2n} < x \leq \frac{1}{n} \\ 0 & x > \frac{1}{n}. \end{cases}$$

Then $d_1(f_n, 0) = \frac{1}{2} \times \frac{1}{n} \times 1 = \frac{1}{2n} \to 0$. So in (X, d_1) we have $f_n \to 0$. But f_n does not converge uniformly: indeed, $f_n \to 0$ pointwise; if we have uniform convergence then uniform limit is the same as pointwise limit; but $\forall\ n\ f_n(\frac{1}{2n}) = 1$ so $f_n \not\to 0$ uniformly.

Example 2.15

Let (X, d) be a discrete metric space; $d(x, y) = \begin{cases} 0 & x = y \\ 1 & x \neq y \end{cases}$. When do we have $x_n \to x$ if (X, d)?

Suppose $x_n \to x$, i.e. $\forall\ \epsilon > 0\ \exists\ N\ \forall\ n \geq N\ d(x_n, x) < \epsilon$. Setting $\epsilon = 1$ in this, we can find N s.t. $\forall\ n \geq N\ d(x_n, x) < 1$, i.e. $\forall\ n \geq N\ d(x_n, x) = 0$ i.e. $\forall\ n \geq N\ x_n = x$. Thus (x_n) is eventually constant.
But we know in any metric space, eventually constant sequences converge.

So in this space, (x_n) converges iff (x_n) eventually constant.

Definition 2.5 (Continuity)

Let (X, d) and (Y, e) be metric spaces and let $f : X \to Y$.

1. Let $a \in X$ and $b \in Y$. We say $f(x) \to b$ as $x \to a$ if $\forall\, \epsilon > 0\ \exists\, \delta > 0\ \forall\, x \in X\ 0 < d(x,a) < \delta \implies e(f(x), b) < \epsilon$.
2. Let $a \in X$. We say f is **continuous** at a if $f(x) \to f(a)$ as $x \to a$. That is: $\forall\, \epsilon > 0\ \exists\, \delta > 0\ \forall\, x \in X\ d(x,a) < \delta \implies e(f(x), f(a)) < \epsilon$.
3. If $\forall\, a \in X$ f is continuous at a we say f is a **continuous** function or simply f is **continuous**.
4. We say f is **uniformly continuous** if $\forall\, \epsilon > 0\ \exists\, \delta > 0\ \forall\, x, y \in X\ d(x,y) < \delta \implies e(f(x), f(y)) < \epsilon$
5. Suppose $W \subset X$. We say f is **continuous on** W (respectively **uniformly continuous on** W) if the function $f\mid_W$ is continuous (resp. uniformly continuous), as a function from $W \to Y$ where we are now thinking of W as a subspace of X.

Remark 8.

1. Don't have a nice rephrasing of item 1 in terms of similar concepts in the reals. We would want to write '$e(f(x), b) \to 0$ as $d(x,a) \to 0$'. But this is meaningless, we haven't defined such a concept in the reals.
2. Item 1 says nothing about what happens at the point a itself. E.g. let $f : \mathbb{R} \to \mathbb{R}, f(x) = \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases}$. Then $f(x) \to 0$ as $x \to 0$ (but $f(0) \neq 0$ so f is not continuous at 0).
   If we have f cts then $d(x,a) = 0 \implies x = a \implies f(x) = f(a) \implies e(f(x), f(a)) = 0$. So we can drop the '$0 <$' from definition of continuity.
3. We can rewrite item 5: f is continuous on W iff $f\mid_W$ is a continuous function $f\mid_W : W \to Y$ thinking of W as a subspace of X. That is: $\forall\, a \in W\ \forall\, \epsilon > 0\ \exists\, \delta > 0\ \forall\, x \in X\ d(x,a) < \delta \implies e(f(x), f(a)) < \epsilon$. In particular, note the subtlety that this <u>only</u> mentions points of W. So under this definition, e.g. $f : \mathbb{R} \to \mathbb{R}, f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & x \notin [0,1] \end{cases}$ then $f\mid_{[0,1]}$ is cts. But f is not cts at points $0, 1$.

Proposition 2.2

Let $(X,d), (Y,e)$ be metric spaces, $f : X \to Y$ and $a \in X$. Then f is continuous at a iff whenever (x_n) is a sequence in X with $x_n \to a$ then $f(x_n) \to f(a)$.

Proof. ($\implies$): Suppose f is cts at a. Let (x_n) be a sequence in X with $x_n \to a$. Let $\epsilon > 0$. As f cts at a we can find $\delta > 0$ s.t. $\forall\, x \in X$ s.t. $d(x,a) < \delta \implies e(f(x), f(a)) < \epsilon$. As $x_n \to x$ we can find N s.t. $n \geq N \implies d(x_n, a) < \delta$. Let $n \geq N$ then $d(x_n, a) < \delta$ so $e(f(x_n), f(a)) < \epsilon$. Hence $f(x_n) \to f(a)$.

($\Longleftarrow$): Suppose f is not cts at a. Then there is some $\epsilon > 0$ s.t. $\forall\, \delta > 0\ \exists\, x \in X$ with $d(x, a) < \delta$ but $e(f(x), f(a)) \geq \epsilon$. Now take $\delta = \frac{1}{n}$ we obtain a sequence (x_n) with, for each n $d(x_n, a) < \frac{1}{n}$ but $e(f(x_n), f(a)) > \epsilon$. Hence $x_n \to a$ but $f(x_n) \not\to f(a)$. □

Proposition 2.3

Let $(W, c), (X, d), (Y, e)$ be metric spaces, left $f : W \to X$, let $g : X \to Y$ and let $a \in W$. Suppose f is cts at a and g is cts at $f(a)$. Then $g \circ f$ is cts at a.

Proof. Let (x_n) be a sequence in W with $x_n \to a$. Then by proposition 2.2, $f(x_n) \to f(a)$ and so also $g(f(x_n)) \to g(f(a))$. So by proposition 2.2 $g \circ f$ cts at a.

What?airsntaeirsnt □

Example 2.16

In $\mathbb{R} \to \mathbb{R}$ with the usual metric, this is the same definition as when we defined continuity directly for $\mathbb{R}$ only. So we already have lots of cts fcns $\mathbb{R} \to \mathbb{R}$: polynomials, sin, e^x, ...

Example 2.17

Constant functions are continuous. Also if X is any metric space and $f : X \to X$ by $f(x) = x$ for all $x \in X$ (the indentity function) then that is continuous.

Example 2.18 (Projection Maps)

Consider $\mathbb{R}^n$ with the usual metric and $\mathbb{R}$ with the usual metric. The **projection maps** $\pi_i : \mathbb{R}^n \to \mathbb{R}$ given by $\pi_i(x) = x_i$ are continuous.

(Why? We've seen convergence in $\mathbb{R}^n$ of sequences is the same as convergence in each coordinate. Let's denote a sequence in $\mathbb{R}^n$ by $(x^{(m)})_{m \geq 1}$. So e.g. $x_5^{(3)}$ is the 5th coord of the 3rd term. We know $x^{(m)} \to x$ iff for each i $x_i^{(m)} \to x_i$, i.e. for each i $\pi_i(x^{(m)}) \to \pi_i(x)$. Then we can use proposition 2.2)

Similarly, suppose $f_1, \ldots, f_n : \mathbb{R} \to \mathbb{R}$ Let $f : \mathbb{R} \to \mathbb{R}^n$ defined by $f(x) = (f_1(x), \ldots, f_n(x))$. Then f is cts at a point iff all of $f_1, \ldots, f_n$ are. Using these facts example 2.16 and proposition 2.3, we have many cts fcns $\mathbb{R}^n \to \mathbb{R}^m$. E.g. $f : \mathbb{R}^3 \to \mathbb{R}^2, f(x, y, z) = (e^{-x} \sin y, 2x \cos z)$ is cts. (Why? write $w = (x, y, z) \in \mathbb{R}^3$, we have $f_1(w) = e^{-\pi_1(w)} \sin \pi_2(w)$ and $f_2(w) = 2\pi_1(w) \cos \pi_3(w)$. So f_1, f_2 cts so f cts.)

Example 2.19

Recall that if we have the Euclidean metric, the l_1 or l_∞ metric on $\mathbb{R}^n$ then the convergent sequences are the same. So by proposition 2.2, the ctf fcns $X \to \mathbb{R}^n$ or from $\mathbb{R}^n \to Y$ are the same with each of these three metrics.

Example 2.20

Let (x, d) be the discrete metric space, example 2.11, and let (Y, e) be any metric space. Which functions $f : X \to Y$ are cts? Suppose $a \in X$ and (x_n) a sequence in X with $x_n \to a$. Then (x_n) is eventually constant, i.e. for sufficiently large n $x_n = a$ and so $f(x_n) = f(a)$. So $f(x_n) \to f(a)$.

Hence every function on a discrete metric space is cts.

§2.2 Completeness

Question

In section 1 we saw a version of GPC held in each of the three examples we considered. Does GPC hold in a general metric space?

Definition 2.6 (Cauchy Sequences)

Let (X, d) be a metric space and let (x_n) be a sequence in X. We say (x_n) is **Cauchy** if $\forall\ \epsilon > 0\ \exists\ N\ \forall\ m, n \geq N\ d(x_m, x_n) < \epsilon$.

Theorem 2.1

(x_n) convergent $\implies$ (x_n) Cauchy.

Proof. Left as an exercise. □

But converse is not true in general.

Example 2.21

Let $X = \mathbb{R} \setminus \{0\}$ with the usual metric and $x_n = \frac{1}{n}$. We say previously that (x_n) does not converge.
Note that X is a subspace of $\mathbb{R}$. In $\mathbb{R}$ (x_n) is convergent $(x_n \to 0)$ so (x_n) is Cauchy in $\mathbb{R}$ so (x_n) is Cauchy in X.

Example 2.22

$\mathbb{Q}$ with the usual metric. Let x_n be $\sqrt{2}$ to n decimal places. This converges in $\mathbb{R}$ so is Cauchy in $\mathbb{Q}$ but clearly doesn't converge in $\mathbb{Q}$.

Definition 2.7 (Completeness)

Let (X, d) be a metric space. We say X is **complete** if every Cauchy sequence in X converges.

Example 2.23

Example 2.21 says $\mathbb{R} \setminus \{0\}$ with the usual metric is not complete.
Similarly Q with usual metric is not complete.

Example 2.24

GPC says $\mathbb{R}$ with the usual metric is complete.

Example 2.25

GPC for $\mathbb{R}^n$ says $\mathbb{R}^n$ with Euclidean metric is complete.

Example 2.26

GPUC, theorem 1.8, (almost) says if $X \subset \mathbb{R}$ and $B(X) = \{f : X \to \mathbb{R} \mid f \text{ is bounded}\}$ with the uniform norm then $B(X)$ is complete.

Proof. Let (f_n) be a Cauchy sequence in $B(X)$. Then (f_n) is uniformly Cauchy so by GPUC is uniformly convergent. That is $f_n \to f$ uniformly for some $f : X \to \mathbb{R}$. As $f_n \to f$ uniformly we know $f_n - f$ is bounded for n suff. large. Take such an n, then $f_n - f$ and f_n are bounded so $f = f_n - (f_n - f)$ is bounded. That is, $f \in B(X)$. Finally, $f_n \to f$ uniformly and $d(f_n, f) \to 0$, i.e. $f_n \to f$ in $(B(X), d)$. $\square$

Remark 9. In many ways, this is typical of a proof that a given space (X, d) is complete:

1. Take (x_n) Cauchy in X;
2. Construct/ find a putative limit object x where it seems (x_n) converges to x in some sense;
3. Show $x \in X$,
4. Show $x_n \to x$ in metric space (X, d) i.e. that $d(x_n, x) \to 0$.

This is often tricky/ fiddly/ annoying/ repetitive/ boring. But we need to take care as for example, it's tempting to talk about $d(x_n, x)$ while doing (ii) or (iii); but makes no sense to write '$d(x_n, x)$' until we have completed (iii) as d is only defined on X^2 (if $x \notin X$ then can't use d).

Example 2.27

If $[a, b]$ is a closed interval then $C([a, b])$ with uniform norm d is complete.

Proof. (i): Let (f_n) be a Cauchy sequence in $C([a, b])$.
(ii): We know $C([a, b])$ is a subspace of $B([a, b])$ with uniform metric. We know $B([a, b])$ is complete by example 2.26 and (f_n) is a Cauchy sequence in $B([a, b])$ so in $B([a, b])$, $f_n \to f$ for some f.
(iii) Each f_n is cts and $f_n \to f$ uniformly so f is cts, i.e. $f \in C([a, b])$.
(iv) Finally, each $f_n \in C([a, b])$, $f \in C([a, b])$ and $f_n \to f$ uniformly so $d(f_n, f) \to 0$. □

This generalises:

Definition 2.8 (Closed Metric Space)

Let (X, d) be a metric space and $Y \subset X$. We say Y is **closed** if whenever (x_n) a sequence in Y with $x_n \to x \in X$ then $x \in Y$.

Proposition 2.4

A closed subset of a complete metric space is complete.

Remark 10. This <u>does</u> make sense: if $Y \subset X$ then Y is itself a metric space or a subspace of X so we can say e.g. 'Y is complete' to mean the metric space Y (as a subspace of X) is complete.

We could do exactly the same with any other properties of metric spaces we define.

Proof. Let (X, d) be a metric space and $Y \subset X$ with X complete and Y closed. (i): Let (x_n) be a Cauchy sequence in Y.
(ii): Now (x_n) is a Cauchy sequence in X so by completeness $x_n \to x$ in X for some $x \in X$.
(iii) $Y \subset X$ is closed so $x \in Y$.
(iv) Finally we now have each $x_n \in Y, x \in Y$ and $x_n \to x$ in X, so $d(x_n, x) \to 0$ so $x_n \to x$ in Y. □

Example 2.28

Define $\ell_1 = \{(x_n)_{n\geq 1} \in \mathbb{R}^{\mathbb{N}} \mid \sum_{n=1}^{\infty} |x_n| \text{ converges}\}$. Define a metric d on ℓ_1 by $d((x_n),(y_n)) = \sum_{n=1}^{\infty} |x_n - y_n|$.

Note we have $\sum |x_n|, \sum |y_n|$ converge as we are in ℓ_1. For each n $|x_n - y_n| \leq |x_n| + |y_n|$ so by comparison test $\sum |x_n - y_n|$ converges. So d is well-defined. Easy to check d is a metric on ℓ_1. Then (ℓ_1, d) is complete.

Proof. (i): Let $(x^{(n)})_{n\geq 1}$ be a Cauchy sequence in ℓ_1, so for each n $(x_i^{(n)})_{i\geq 1}$ is a sequence in $\mathbb{R}$ with $\sum_{i=1}^{\infty} |x_i^{(n)}|$ convergent.
(ii) For each i, $(x_i^{(n)})_{n\geq 1}$ is a Cauchy sequence in $\mathbb{R}$, since if $y, z \in \ell_1$ then $|y_i - z_i| \leq d(y,z)$. But $\mathbb{R}$ is complete, so for each i we can find $x_i \in \mathbb{R}$ s.t. $x_i^{(n)} \to x_i$ as $n \to \infty$. Let $x = (x_1, x_2, \dots) \in \mathbb{R}^{\mathbb{N}}$.
(iii) We next show $x \in \ell_1$, i.e. that $\sum_{i=1}^{\infty} |x_i|$ converges.

Given $y \in \ell_1$, define $\sigma(y) = \sum_{i=1}^{\infty} |y_i|$, i.e. $\sigma(y) = d(y,z)$ where z is the constant zero sequence.
We now have, for any m, n

$$\begin{aligned}
\sigma(x^{(m)}) &= d(x^{(m)}, z) \\
&\leq d(x^{(m)}, x^{(n)}) + d(x^{(n)}, z) \\
&= d(x^{(m)}, x^{(n)}) + \sigma(x^{(n)})
\end{aligned}$$

So $\sigma(x^{(m)}) - \sigma(x^{(n)}) \leq d(x^{(m)}, x^{(n)})$. Similarly, for any m, n $\sigma(x^{(n)}) - \sigma(x^{(m)}) \leq d(x^{(m)}, x^{(n)})$ and so $|\sigma(x^{(m)}) - \sigma(x^{(n)})| \leq d(x^{(m)}, x^{(n)})$. Hence $(\sigma(x^{(m)}))_{m\geq 1}$ is a Cauchy sequence in $\mathbb{R}$, and so by GPC converges, say $\sigma(x^{(m)}) \to K$ as $m \to \infty$.

Claim 2.1

For any $I \in \mathbb{N}$, $\sum_{i=1}^{I} |x_i| \leq K + 2$.

Proof. As $\sigma(x^{(n)}) \to K$ as $n \to \infty$ we can find N_1 s.t. $n \geq N_1 \implies \sum_{i=1}^{\infty} |x_i^{(n)}| \leq K + 1$. Also, $n \geq N_1 \implies \sum_{i=1}^{I} |x_i^{(n)}| \leq K + 1$ (as each term non-negative).

Next, for each $i \in \{1, 2, \dots, I\}$ we have $x_i^{(n)} \to x_i$ as $n \to \infty$. So we can find N_2 s.t. $n \geq N_2 \implies \forall\, i \in \{1, \dots, I\}$ $|x_i^{(n)} - x_i| < I^{-1}$.

Now let $n = \max(N_1, N_2)$ then $\sum_{i=1}^{I} |x_i| \leq \sum_{i=1}^{I} |x_i^{(n)}| + \sum_{i=1}^{I} |x_i^{(n)} - x_i| \leq K + 1 + I(I^{-1}) = K + 2$. $\square$

Now the partial sums of $\sum |x_i|$ are increasing and bounded above so $\sum |x_i|$ converges. That is $x \in \ell_1$.

(iv) Finally, need to check $x^{(n)} \to x$ as $n \to \infty$ in ℓ_1, i.e. that $d(x^{(n)}, x) \to 0$ as $n \to \infty$.
We have, for all n, I:

$$\begin{aligned}
d(x^{(n)}, x) &= \sum_{i=1}^{\infty} |x_i^{(n)} - x_i| \\
&\le \sum_{i=1}^{I} |x_i^{(n)} - x_i| + \sum_{i=I+1}^{\infty} |x_i^{(n)}| + \sum_{i=I+1}^{\infty} |x_i|.
\end{aligned}$$

Let $\epsilon > 0$. We know $\sum |x_i|$ convergent (as $x \in \ell_1$) so we can pick I_1 s.t. $\sum_{i=I_1+1}^{\infty} |x_i| < \epsilon$.
As $(x^{(n)})$ is Cauchy, we can find N_1 s.t. $m, n \ge N_1 \implies d(x^{(m)}, x^{(n)}) < \epsilon$. As $\sum_i |x_i^{(N_1)}|$ converges, we can find I_2 s.t. $\sum_{i=I_2+1}^{\infty} |x_i^{(N_1)}| < \epsilon$. Then

$$\begin{aligned}
n \ge N_1 \implies \sum_{i=I_2+1}^{\infty} |x_i^{(n)}| &\le \sum_{i=I_2+1}^{\infty} |x_i^{(N_1)}| + \sum_{i=I_2+1}^{\infty} |x_i^{(n)} - x_i^{(N_1)}| \\
&< \epsilon + d(x^{(n)}, x^{(N_1)}) \\
&< 2\epsilon.
\end{aligned}$$

Let $I = \max(I_1, I_2)$. For each $i = 1, 2, \dots, I$ we have $|x_i^{(n)} - x_i| \to 0$ as $n \to \infty$, so $\sum_{i=1}^{I} |x_i^{(n)} - x_i| \to 0$ as $n \to \infty$. Hence we can find N_2 s.t $n \ge N_2 \implies \sum_{i=1}^{I} |x_i^{(n)} - x_i| < \epsilon$. Let $N = \max(N_1, N_2)$ and let $n \ge N$. Then

$$\begin{aligned}
d(x^{(n)}, x) &\le \sum_{i=1}^{I} |x_i^{(n)} - x_i| + \sum_{i=I+1}^{\infty} |x_i^{(n)}| + \sum_{i=I+1}^{\infty} |x_i| \\
&\le \sum_{i=1}^{I} |x_i^{(n)} - x_i| + \sum_{i=I_2+1}^{\infty} |x_i^{(n)}| + \sum_{i=I_1+1}^{\infty} |x_i| \\
&< \epsilon + 2\epsilon + \epsilon = 4\epsilon
\end{aligned}$$

Hence $d(x^{(n)}, x) \to 0$ as $n \to \infty$, i.e. $x^{(n)} \to x$ in ℓ_1.
Hence ℓ_1 is complete. □

Now we will move on to main theorem of completeness.

Definition 2.9 (Contraction mapping)

Let (X, d) be a metric space and $f : X \to X$. We say f is a **contraction** if $\exists\, \lambda \in [0, 1)$ s.t. $\forall\, x, y \in X$ $d(f(x), f(y)) \le \lambda d(x, y)$.

Theorem 2.2 (The Contraction Mapping Theorem)

Let (X, d) be a complete, non-empty metric space and $f : X \to X$ a contraction. Then f has a unique fixed point.

Proof. Let $\lambda \in [0, 1)$ satisfy $\forall\, x, y \in X\ d(f(x), f(y)) \leq \lambda d(x, y)$.
Let $x_0 \in X$. Recursively define $x_n = f(x_{n-1})$ for $n \geq 1$. Let $\Delta = d(x_0, x_1)$. Then, by induction $d(x_n, x_{n+1}) \leq \lambda^n \Delta$ for all n.
Now suppose $N \leq m < n$. Then

$$\begin{aligned} d(x_m, x_n) &\leq \sum_{i=m}^{n-1} d(x_i, x_{i+1}) \\ &\leq \sum_{i=m}^{n-1} \lambda^i \Delta \\ &\leq \sum_{i=N}^{\infty} \lambda^i \Delta \\ &= \frac{\lambda^N \Delta}{1 - \lambda} \to 0 \text{ as } N \to \infty. \end{aligned}$$

So $\forall\ \epsilon > 0\ \exists\ N\ \forall\ m, n \geq N\ d(x_m, x_n) < \epsilon$ (i.e. we take N s.t. $\frac{\lambda^N \Delta}{1-\lambda} < \epsilon$). Thus (x_n) is Cauchy, so by completeness converges, say $x_n \to x \in X$. But also $x_n = f(x_{n-1}) \to f(x)$ as f continuous[a]. So by uniqueness of limits, $f(x) = x$.

Suppose also $f(y) = y$ for some $y \in X$. Then $d(x, y) = d(f(x), f(y)) \leq \lambda d(x, y)$ with $\lambda < 1$. So $d(x, y) = 0$ so $x = y$. □

[a] follows immediately from definition of contraction mapping (e.g. let $\delta = \epsilon$ in definition of continuity).

Remark 11.

1. Why is f cts? We have, for all $x, y \in X$ $d(f(x), f(y)) \leq d(x, y)$. So $\forall\ \epsilon > 0$, $d(x, y) < \epsilon \implies d(f(x), f(y)) < \epsilon$. (Indeed, this shows f is uniformly continuous.)

2. We have proved more than claimed. Not only does f have a unique fixed point, but starting from any point of the space and repeatedly apply f then the resulting sequence converges to the fixed point. In fact, the speed of convergence is exponential.

Example 2.29 (Application)

Suppose we want to numerically approximate the solution to $\cos x = x$. Any root must lie in $[-1, 1]$. Consider metric space $X = [-1, 1]$ with usual metric. X is a closed subset of a complete space $\mathbb{R}$ so is complete. Obviously X is non-empty.

Think of $\cos : [-1,1] \to [-1,1]$. Suppose $x, y \in [-1,1]$.

$$\begin{aligned}|\cos x - \cos y| &= |x-y||\cos' z| \text{ for some } z \in [-1,1] \text{ by MVT} \\ &|x-y||-\sin' z| \\ &\leq |x-y| \sin 1\end{aligned}$$

But $0 \leq \sin 1 < 1$ so cos is a contraction of $[-1,1]$. So by The Contraction Mapping Theorem, cos has a unique fixed point in $[-1,1]$. That is $\cos x = x$ has a unique solution.

How do we find it numerically? Use remark 2, we will have rapid convergence to the root.

We will see two major applications of CMT (The Contraction Mapping Theorem) later.

§2.3 Sequential Compactness

Recall BW for $\mathbb{R}^n$ says a bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.

Definition 2.10 (Bounded)
Let (X, d) be a metric space. We say X is **bounded** if

$$\exists\, M \in \mathbb{R}\; \forall\, x, y \in X\;\; d(x,y) \leq M.$$

Remark 12. Easy to check by triangle inequality that X bounded $\iff$ ($X = \emptyset$ or $\exists\, M \in \mathbb{R}\;\, \epsilon x \in X$ s.t. $\forall\; y \in X\;\; d(x,y) \leq M$). So definition agrees with earlier definition for subsets of $\mathbb{R}^n$.

Definition 2.11 (Closed subspace)
Let (X, d) be a metric space and $Y \subset X$. We say Y is **closed** in X if whenever (x_n) is a sequence in Y with, in X, $x_n \to x \in X$ then $x \in Y$.

Definition 2.12 (Sequentially Compact)
A metric space is **sequentially compact** if every sequence has a convergent subsequence.

BW for $\mathbb{R}^n$ is essentially the following:

Theorem 2.3
Let $X \subset \mathbb{R}^n$ with the Euclidean metric. Then X is sequentially compact iff X is

closed and bounded.

Proof. ($\Longleftarrow$) Suppose X is closed and bounded. Let (x_n) be a sequence in X. Then (x_n) is a bounded sequence in $\mathbb{R}^n$ so by BW, in $\mathbb{R}^n$, $x_{n_j} \to x$ for some $x \in \mathbb{R}^n$ and some subsequence (x_{n_j}) of (x_n).
As X is closed, $x \in X$. Hence the subsequence (x_{n_j}) converges in X. So X is sequentially compact.

($\implies$) Suppose X is not closed. Then we can find a sequence (x_n) in X s.t. in $\mathbb{R}^n$ $x_n \to x \in \mathbb{R}^n$ with $x \notin X$. Now any subsequence $(x_{n_j}) \to x$ in $\mathbb{R}^n$. But $x \notin X$ so by uniqueness of limits (x_{n_j}) does not converge in X. So X is not sequentially compact.

Suppose instead X is not bounded. Then we can find a sequence (x_n) in X with $\forall\ n\ \|x_n\| \geq n$, i.e. $\|x_n\| \to \infty$ as $n \to \infty$. Suppose we have a subsequence $x_{n_j} \to x \in X$. Then $\left\|x_{n_j}\right\| \to \|x\|$ but $\left\|x_{n_j}\right\| \to \infty$ ↯. So, again, X is not sequentially compact. ☐

Remark 13. Does this hold in a general metric space? Obviously not: e.g. in $\mathbb{R} \setminus \{0\}$ with the usual metric, the set $[-1,0) \cup (0,1]$ is closed and bounded but the sequence $(\frac{1}{n})_{n\geq 1}$ has no convergent subsequence.

<u>Problem</u>: Space is not complete. Maybe complete and bounded $\implies$ sequentially compact?

Even this doesn't work. Recall example from section 1: Let $X = \{f \in B(\mathbb{R}) : \sup_{x\in\mathbb{R}} |f(x)| \leq 1\}$ with uniform metric. Then X is complete (closed subset of a complete space $B(\mathbb{R})$) and bounded (if $f, g \in X$ then $d(f,g) \leq 2$). But consider $f_n(x) = \begin{cases} 1 & x = n \\ 0 & x \neq n \end{cases}$. Then f_n is a sequence in X but $\forall\ m,n\ m \neq n \implies d(f_m, f_n) = 1$. So (f_n) cannot have a convergent subsequence.

<u>Problem</u>: X is 'too big'.

We need a stronger concept of boundedness.

Definition 2.13 (Totally Bounded)

Let (X,d) be a metric space. We say X is **totally bounded** if $\forall\ \delta > 0$ we can find a <u>finite</u> subset $A \subset X$ s.t. $\forall\ x \in X\ \exists\ a \in A\ d(x,a) < \delta$.

Theorem 2.4

A metric space is sequentially compact iff it is complete and totally bounded.

Proof. ($\Longleftarrow$): Suppose the metric space (X, d) is complete and totally bounded. Let $(x_n)_{n\geq 1}$ be a sequence in X.
As X is totally bounded, we can find finite $A_1 \subset X$ s.t. $\forall\, x \in X\ \exists\, a \in A_1\ d(x, a) < 1$.
In particular, there is an infinite set $N_1 \subset \mathbb{N}$ and a point $a_1 \in A_1$ s.t. $\forall\, n \in N_1\ d(x_n, a_1) < 1$. Hence $\forall\, m, n \in N_1\ d(x_m, x_n) < 2$.
Similarly, we can find finite $A_2 \subset X$ s.t. $\forall\, x \in X\ \exists\, a \in A_2\ d(x, a) < \frac{1}{2}$. In particular, there is an infinite $N_2 \subset N_1$ s.t. $\forall\, n \in N_2\ d(x_n, a_2) < \frac{1}{2}$ and thus $\forall\, m, n \in N_2\ d(x_m, x_n) < 1$.
Keep going. We get a sequence $N_1 \supset N_2 \supset N_3 \dots$ of infinite subsets of $\mathbb{N}$ s.t. $\forall\, i\ \forall\, m, n \in N_i \implies d(x_m, x_n) < \frac{2}{i}$.
Now pick $n_1 \in N_1$. Then pick $n_2 \in N_2$ with $n_2 > n_1$. Then pick $n_i \in N_i$ with $n_i > n_{i-1}$.
We obtain a subsequence (x_{n_j}) of (x_n) s.t. $\forall\, j\ x_{n_j} \in N_j$. Thus if $i \leq j$ then $x_{n_i}, x_{n_j} \in N_i$ and so $d(x_{n_i}, x_{n_j}) < \frac{2}{i}$. Hence (x_{n_j}) is a Cauchy sequence and hence by completeness, converges. Thus X is sequentially compact.

($\implies$) Suppose X is not complete. Then X has a Cauchy sequence (x_n) which doesn't converge. Suppose we have a convergent subsequence, say $x_{n_j} \to x$. Then $x_n \to x$ (left as an exercise, same as in gpc for $\mathbb{R}$) $\lightning$.

Suppose instead X not totally bounded. Then there is some $\delta > 0$ s.t. whenever $A \subset X$ is finite $\exists\, x \in X\ \forall\, a \in A\ d(x, a) < \delta$. So pick $x_1 \in X$, pick $x_2 \in X$ s.t. $d(x_1, x_2) \geq \delta$, pick $x_3 \in X$ s.t. $d(x_1, x_3) \geq \delta$ and $d(x_2, x_3) \geq \delta$ ...We get a sequence $(x_n) \in X$ s.t. $\forall\, i, j\ i \neq j \implies d(x_i, x_j) \geq \delta$. Hence (x_n) has no convergent subsequence. $\square$

Exercise 2.5. A cts fcn on a sequentially compact metric space is uniformly compact. If the fcn is real-valued then it's bounded and attains its bounds.

§2.4 The Topology of Metric Spaces

Theme of section 2: to generalise convergence/ continuity, all we need is a distance.

But: e.g. in $\mathbb{R}^n$ we have three very different concepts of distance given by the Euclidean, ℓ_1 and ℓ_∞ metrics. But all give same concept of convergence and continuity.

Definition 2.14 (Homeomorphism)

Let (X, d) and (Y, e) be metric spaces. Let $f : X \to Y$. We say f is a **homeomorphism** and that X, Y are **homeomorphic** if f is a cts bijection with a cts inverse.

Remark 14. Homeomorphism is an equivalence 'relation' (it satisfies symmetry, transitivity and reflexivity but not an actual relation).

Example 2.30

If $x, y \in \mathbb{R}^n$: $d_\infty(x,y) \le d_1(x,y) \le n d_\infty(x,y)$. ($d_1, d_\infty$ ℓ_1 and ℓ_∞ metrics respectively). So identity map $\mathbb{R}^n \to \mathbb{R}^n$ is continuous as map $(\mathbb{R}^n, d_1) \to (\mathbb{R}^n, d_\infty)$ and as a map $(\mathbb{R}^n, d_\infty) \to (\mathbb{R}^n, d_1)$. So it's a homeomorphism.
Similarly, $\mathbb{R}^n$ with the Euclidean metric is homeomorphic to both of these spaces.

Example 2.31

Same argument would show: If $(X,d), (Y,e)$ metric spaces and $f : X \to Y$ is a bijection satisfying

1. $\exists\, A\ \forall\, x, y \in X\ e(f(x), f(y)) \le A d(x,y)$
2. $\exists\, B\ \forall\, x, y \in X\ d(x,y) \le B e(f(x), f(y))$ then f, f^{-1} are cts so X, Y homeomorphic.

Example 2.32

Define $f : (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ by $f(x) = \tan x$. Then f is a homeomorphism (usual metric in each case). But there is no constant A s.t. $\forall\, x, y \in (-\frac{\pi}{2}, \frac{\pi}{2})$ $|\tan x - \tan y| \le A|x - y|$.

Proposition 2.5

Let $(V,b), (W,c), (X,d), (Y,e)$ be metric spaces and $f : X \to V$, $g : Y \to W$ be homeomorphisms.

1. In X, $x_n \to x$ iff in V $f(x_n) \to f(x)$;
2. A function $h : X \to Y$ is cts at $a \in X$ iff $g \circ h \circ f^{-1}$ is cts at $f(a) \in V$.

Proof. (i) $x_n \to x \implies f(x_n) \to f(x)$ as f cts; $f(x_n) \to f(x) \implies x_n = f^{-1}(f(x_n)) \to f^{-1}(f(x)) = x$ as f^{-1} cts.

(ii) h cts $\implies g \circ h \circ f^{-1}$ cts (composition of cts fcns). And $g \circ h \circ f^{-1}$ cts $\implies h = g^{-1} \circ (g \circ h \circ f^{-1}) \circ f$ is cts (similarly). □

We now have examples of metric spaces that look very different but behave identically w.r.t convergence/ continuity.

Thought: Could we dispense with distance altogether?

Another way to think about continuity.

Definition 2.15 (Open ball)
Let (X, d) be a metric space, let $a \in X$ and let $\epsilon > 0$. The **open ball of radius ϵ about a** is the set $B_\epsilon(a) = \{x \in X : d(x, a) < \epsilon\}$.

Remark 15. Suppose $f : X \to Y$, $a \in X$. d metric on X, e metric on Y.

$$\begin{aligned} f \text{ cts at } a &\iff \forall\, \epsilon > 0 \; \exists\, \delta > 0 \; d(x, a) < \delta \implies d(f(x), f(a)) < \epsilon \\ &\iff \forall\, \epsilon > 0 \; \delta > 0 \; x \in B_\delta(a) \implies f(x) \in B_\epsilon(f(a)) \\ &\iff \forall\, \epsilon > 0 \; \exists\, \delta > 0 \; f(B_\delta(a)) \subset B_\epsilon(f(a)) \\ &\iff \forall\, \epsilon > 0 \; \exists\, \delta > 0 \; B_\delta(a) \subset f^{-1}(B_\epsilon(f(a))). \end{aligned}$$

So we have redefined continuity in terms of open balls. But open balls have radii so still mentioning distance.

Definition 2.16 (Open, Neighbourhood)
Let X be a metric space. A subset $G \subset X$ is **open** if $\forall\, x \in G \; \exists\, \epsilon > 0 \; B_\epsilon(x) < G$. A subset $N \subset X$ is a **neighbourhood** (nbd) of a point $a \in X$ if there exists an open set $G \subset X$ s.t. $a \in G \subset N$.

Part II
Generalizing differentiation