IB Analysis & Topology

Ishan Nath, Michaelmas 2022

Based on Lectures by Dr. Paul Russell

October 25, 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 for all $\varepsilon > 0$, there exists N such that for all $n \geq N$, $|x_n - x| < \varepsilon$.

In $\mathbb{R}$, one useful fact is the **triangle inequality** – $|a + b| \leq |a| + |b|$. We also have two key theorems:

Theorem 1.1 (Bolzano-Weierstrass Theorem)**.**

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

Recall that a sequence (x_n) in $\mathbb{R}$ is **Cauchy** if for all $\varepsilon > 0$, there exists N, such that for all $m, n \geq N$, $|x_m - x_n| < \varepsilon$. It is easy to show every convergent sequence is Cauchy. We also have the following:

Theorem 1.2 (General Principle of Convergence)**.**

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

This can be proven by Bolzano-Weierstrass theorem.

1.2 Convergence in $\mathbb{R}^2$

Let (z_n) be a sequence in $\mathbb{R}^2$, and $z \in \mathbb{R}^2$. We wish to define $(z_n) \to z$.

In $\mathbb{R}$, we used the norm $|x|$. In $\mathbb{R}^2$, if we have $z = (x, y)$, then we can say $\|z\| = \sqrt{x^2 + y^2}$. This also satisfies the triangle inequality – $\|a + b\| \leq \|a\| + \|b\|$.

Definition 1.1. Let (z_n) be a sequence in $\mathbb{R}^2$, and $z \in \mathbb{R}^2$. We say that (z_n) **converges** to z, and write $z_n \to z$, if for all $\varepsilon > 0$, there exists N such that for all $n \geq N$, $\|z_n - z\| < \varepsilon$.

Equivalently, $z_n \to z$ if and only if $\|z_n - z\| \to 0$.

Lemma 1.1. *If (z_n), (w_n) are sequences in $\mathbb{R}^2$ with $z_n \to z$, $w_n \to w$. Then $z_n + w_n \to z + w$.*

Proof:

$$\|(z_n + w_n) - (z + w)\| \leq \|z_n - z\| + \|w_n - w\| \to 0 + 0 = 0.$$

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 $z_n \to z$ if and only if $x_n \to x$ and $y_n \to y$.*

Proof:

First, note $|x_n - x|, |y_n - y| \leq \|z_n - z\|$, so $\|z_n - z\| \to 0$ implies $|x_n - x|, |y_n - y| \to 0$.

Now, if $|x_n - x|, |y_n - y| \to 0$, then $\|z_n - z\| = \sqrt{|x_n - x|^2 + |y_n - y|^2} \to 0$.

Definition 1.2. A sequence (z_n) in $\mathbb{R}^2$ is **bounded** if there exists $M \in \mathbb{R}$ such that for all n, $\|z_n\| \leq M$.

Theorem 1.3 (Bolzano-Weierstrass in $\mathbb{R}^2$)**.**

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

Proof: Let (z_n) be a bounded subsequence in $\mathbb{R}^2$. Write $z_n = (x_n, y_n)$. Now $|x_n|, |y_n| \leq \|z_n\|$, so x_n, y_n are bounded in $\mathbb{R}$.

By Bolzano-Weierstrass, x_n has a convergent subsequence, say $x_{n_j} \to x \in \mathbb{R}$. Similarly (y_{n_j}) is bounded, so it has a convergent subsequence $y_{n_{j_k}} \to y$. Since we know $x_{n_{j_k}} \to x$, $z_{n_{j_k}} \to z = (x, y)$.

Definition 1.3. A sequence $(z_n) \in \mathbb{R}^2$ is **Cauchy** if for all $\varepsilon > 0$, there exists N such that for all $m, n \geq N$, $\|z_m - z_n\| < \varepsilon$.

It is easy to show a convergent sequence in $\mathbb{R}^2$ is Cauchy.

Theorem 1.4 (General Principle of Convergence 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 $\mathbb{R}$, thus it converges in $\mathbb{R}$. Similarly, (y_n) converges in $\mathbb{R}$, so (z_n) converges.

1.3 Convergence of Functions

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

Definition 1.4. Say (f_n) **converges pointwise** to f, and we write $f_n \to f$ pointwise, if for all $x \in X$, $f_n(x) \to f(x)$ as $n \to \infty$.

Although this is simple and easy to check, it doesn't preserve some 'nice' properties that we want.

Example 1.1. In all three examples, $X = [0, 1]$, and $f_n \to f$ pointwise.

1. We will construct f_n continuous, but f not. Take

$$f_n(x) = \begin{cases} nx & x \leq \frac{1}{n}, \\ 1 & x \geq \frac{1}{n}. \end{cases}, f = \begin{cases} 0 & x = 0, \\ 1 & x > 0. \end{cases}$$

   Then $(f_n) \to f$ pointwise, but f is not continuous.

2. We will construct f_n Riemann integrable, but f not. Take the function

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

   Enumerate the rationals in $[0, 1]$ as $q_1, q_2, \ldots$. For $n \geq 1$, set

$$f_n(x) = \begin{cases} 1 & x = q_1, \ldots, q_n, \\ 0 & \text{otherwise}. \end{cases}$$

3. We will construct f_n Riemann integrable, f Riemann integrable, but the integrals do not converge. Take $f(x) = 0$ for all x. We construct f_n with integral 1, such as

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

We consider another definition of convergence.

Definition 1.5 (Uniform Convergence)**.** Let $X \subset \mathbb{R}$, $f_n : X \to \mathbb{R}$, $f : X \to \mathbb{R}$. We say (f_n) **converges uniformly** to f, and write $f_n \to f$ uniformly, if for all $\varepsilon > 0$, there exists N, such that for all $x \in X$ and all $n \geq N$, $|f_n(x) - f(x)| < \varepsilon$.

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

$$\sup_{x \in X} |f_n(x) - f(x)| \to 0.$$

Theorem 1.5. *Let $X \subset \mathbb{R}$, $f_n : X \to \mathbb{R}$ be continuous, and let $f_n \to f : X \to \mathbb{R}$ uniformly. Then f is continuous.*

Proof: Let $x \in X$, and pick $\varepsilon > 0$. As $f_n \to f$ uniformly, we can find N such that for all $n \geq N$ and $\in X$,

$$|f_n(y) - f(y)| < \varepsilon.$$

In particular, we may take $n = N$. As f_N is continuous, we can find $\delta > 0$ such that for all $y \in X$,

$$|y - x| < \delta \implies |f_N(y) - f_N(x)| < \varepsilon.$$

Now let $y \in X$ with $|y - x| < \delta$. Then

$$|f(y) - f(x)| \leq |f(y) - f_N(y)| + |f_N(y) - f_N(x)| + |f_N(x) - f(x)| < \varepsilon + \varepsilon + \varepsilon = 3\varepsilon.$$

But 3ε can be made arbitrarily small, so f is continuous.

Remark. This is often called a '3ε proof' (or a '$\varepsilon/3$ proof').

Theorem 1.6. *Let $f_n : [a, b] \to \mathbb{R}$ 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 sufficiently large such that $f_n - f$ is bounded. Also f_n is bounded, so by the triangle inequality $f = (f - f_n) + f_n$ is bounded.

Let $\varepsilon > 0$. As $f_n \to f$ uniformly, there is some N such that for all $n \geq N$ and $x \in [a, b]$, we have $|f_n(x) - f(x)| < \varepsilon$. By Riemann's criterion, there is some dissection $\mathcal{D}$ of $[a, b]$ for which

$$S(f_N, \mathcal{D}) - s(f_N, \mathcal{D}) < \varepsilon.$$

Let $\mathcal{D} = \{x_0, x_1, \ldots, x_k\}$, where $a = x_0 < x_1 < \cdots < x_k = b$. Now,

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

Similarly, $s(f, \mathcal{D}) \geq s(f_N, \mathcal{D} - (b - a)\varepsilon$, so

$$S(f, \mathcal{D}) - s(f, \mathcal{D}) \leq S(f_N, \mathcal{D}) - s(f_N, \mathcal{D}) + 2(b - a)\varepsilon < (2(b - a) + 1)\varepsilon..$$

But this can be made arbitrarily small, so by Riemann's criterion, f is integrable over $[a, b]$.

Now for any n sufficiently large such that $f_n - f$ is bounded,

$$\begin{aligned}
\left| \int_a^b f_n - \int_a^b f \right| &= \left| \int_a^b (f_n - f) \right| \\
&\leq \int_a^b |f_n - f| \\
&\leq (b - a) \sup_{x \in [a, b]} |f_n(x) - f(x)| \to 0
\end{aligned}$$

as $n \to \infty$ since $f_n \to f$ uniformly.

Unfortunately, uniform convergence cannot preserve all properties.

Example 1.2. Take $f_n : [-1, 1] \to \mathbb{R}$, where each f_n is differentiable and $f_n \to f$ uniformly, but f is not differentiable. Take

$$f_n = \sqrt{\left(\frac{1}{n} + x^2\right)}.$$

Then f_n is differentiable, and also uniformly converges to $f(x) = |x|$. But f is not differentiable.

In fact, we need uniform convergence of the **derivatives**.

Theorem 1.7. *Let $f_n : (u, v) \mapsto \mathbb{R}$ with $f_n \to f : (u, v) \to \mathbb{R}$ pointwise. Suppose further that 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 5, g is integrable over $[a, x]$ and

$$\int_a^x g = \lim_{n\to\infty} \int_a^x f_n'(x) = f(x) - f(a).$$

So we have shown that for all $x \in (u, v)$,

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

By theorem 4, g is continuous so by FTC, f is differentiable with $f' = g$.

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

Definition 1.6. 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 for all $\varepsilon > 0$, there exists N such that for all $m, n \geq N$ and for all $x \in X$,

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

It is easy to show that a uniformly convergent sequence is uniformly Cauchy.

Theorem 1.8 (General Principle of Uniform Convergence)**.** *Let (f_n) be a uniformly Cauchy sequence of functions $X \to \mathbb{R}$. Then (f_n) is uniformly convergent.*

Proof: Let $x \in X$, and $\varepsilon > 0$. Then there exists N, such that for all $m, n \geq N$ and for all $y \in X$, $|f_m(y) - f_n(y)| < \varepsilon$. In particular, $|f_m(x) - f_n(x)| < \varepsilon$, so $(f_n(x))$ is a Cauchy sequence in $\mathbb{R}$, so by GPC, it converges pointwise, say $f_n(x) \to f(x)$ as $n \to \infty$.

Let $\varepsilon > 0$. Then we can find an N such that for all $m, n \geq N$ and all $y \in X$, $|f_m(y) - f_n(y)| < \varepsilon$. Fixing y, m and letting $n \to \infty$, $|f_m(y) - f(y)| \leq \varepsilon$. But since y is arbitrary, this shows $f_n \to f$ uniformly.

We will also try to take Bolzano-Weierstrass over to the space of functions.

Definition 1.7. 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 for all x, there exists M such that for all n, $|f_n(x)| \leq M$.

We say (f_n) is **uniformly bounded** if there exists M, such that for all x and n, $|f_n(x)| \leq M$.

We would like a uniform Bolzano-Weierstrass, saying if (f_n) is a uniformly bounded sequence of functions, then it has a uniformly convergent subsequence. But this is not true.

Example 1.3. Take $f_n : \mathbb{R} \to \mathbb{R}$,

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

Then (f_n) is uniformly bounded, but if $m \neq n$, then $f_m(m) = 1$ and $f_n(m) = 0$, so $|f_m(m) - f_n(m)| = 1$, hence (f_n) are not uniformly Cauchy, so cannot be uniformly convergent.

1.4 Application to Power Series

Recall that if $\sum a_n x^n$ is a real power of series with radius of convergence $R > 0$, then we can differentiate and integrate it term-by-term within $(-R, R)$.

Definition 1.8. Let $f_n : X \to \mathbb{R}$ for each $n \geq 0$. We say that the series

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

converges uniformly if the sequence of partial sums (F_n) does, where $F_n = f_0 + f_1 + \cdots + f_n$.

If we can prove that $\sum a_n x^n$ is uniformly convergent, then we can apply earlier theorems to show differentiability. However this is not quite true, for example take

$$\sum_{n=0}^{\infty} x^n.$$

However, we do have another approach. We can show that if $0 < r < R$, then we do have uniform convergence on $(-r, r)$, and then given $x \in (-R, R)$, we can choose $|x| < r < R$ and use the above to show all the properties we want. This is known as the **local uniform convergence of power series**.

Lemma 1.2. *Let $\sum a_n x^n$ be a real power series with radius of convergence $R > 0$. Let $0 < r < R$. Then $\sum a_n x^n$ converges uniformly on $(-r, r)$.*

Proof: Define $f, f_m : (-r, r) \to \mathbb{R}$ by

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \quad f_m(x) = \sum_{n=0}^{m} a_n x^n.$$

Recall that $\sum a_n x^n$ converges absolutely for all x with $|x| < R$. Let $x \in (-r, r)$. Then

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

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

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

as $m \to \infty$.

Theorem 1.9. *Let $\sum a_n x^n$ be a real power series with radius of convergence $R > 0$.*

Define $f : (-R, R) \to \mathbb{R}$ *by*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n.$$

Then,

(i) f *is continuous;*

(ii) *For any* $x \in (-R, R)$, f *is integrable over* $[0, x]$ *with*

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

Proof: Let $x \in (-R, R)$. Pick r such that $|x| < r < R$. By the above lemma, $\sum a_n y^n$ converges uniformly on $(-r, r)$. But the partial sum functions are all continuous on $(-r, r)$, hence $f|_{(-r,r)}$ is continuous. Thus f is a continuous function 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 on $[0, x]$ is a polynomial, so can be integrated with

$$\int_0^x \sum_{n=0}^{m} a_n y^n \, \mathrm{d}y = \sum_{n=0}^{m} \int_0^x a_n y^n \, \mathrm{d}y = \sum_{n=0}^{m} \frac{a_n}{n+1} x^{n+1}.$$

Thus, f is integrable over $[0, x]$ with

$$\int_0^x f = \lim_{m\to\infty} \int_0^x \sum_{n=0}^{m} a_n y^n \, \mathrm{d}y = \lim_{m\to\infty} \sum_{n=0}^{m} \frac{a_n}{n+1} x^{n+1} = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1}.$$

For differentiation, we need the following lemma:

Lemma 1.3. *Let* $\sum a_n x^n$ *be a real power series with radius of convergence* $R > 0$. *Then the power series* $\sum n a_n x^{n-1}$ *has radius of convergence at least* R.

Proof: Let $x \in \mathbb{R}$, $0 < |x| < R$. Pick w with $|x| < w < R$. Then $\sum a_n w^n$ is absolutely convergent, so $a_n w^n \to 0$. Therefore, there exists M such that $|a_n w^n| \leq M$ for all n. 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 $\alpha = |x/w| < 1$, and let $c = M/|x|$, a constant. Then $|na_n x^{n-1}| \leq cn\alpha^n$. By comparison test, it suffices to show $\sum n\alpha^n$ converges. Note

$$\left| \frac{(n+1)\alpha^{n+1}}{n\alpha^n} \right| = \left(1 + \frac{1}{n}\right) \alpha \to \alpha < 1$$

as $n \to \infty$, so this converges by the ratio test.

Theorem 1.10. *Let $\sum a_n x^n$ be a real power series with radius of convergence $R > 0$. 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 for all $x \in (-R, R)$,

$$f'(x) = \sum_{n=1}^{\infty} na_n x^{n-1}.$$

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 $\sum na_n y^{n-1}$ had radius of convergence at least R, and so also converges uniformly on $(-r, r)$.

The partial sum functions $f_m(y)$ are polynomials, so are differentiable with

$$f'_m(y) = \sum_{n=1}^{m} na_n y^{n-1}.$$

we now have f'_m converging uniformly on $(-r, r)$ to the function

$$g(y) = \sum_{n=1}^{\infty} na_n y^{n-1}.$$

Hence, $f|_{(-r,r)}$ is differentiable and for all $y \in (-r, r)$, $f'(y) = g(y)$. In particular, f is differentiable at x with $f'(x) = g(x)$. This gives f is a differentiable function on $(-R, R)$ with derivative g as desired.

1.5 Uniform Continuity

Let $X \subset \mathbb{R}$. Let $f : X \mapsto \mathbb{R}$. Recall that f is **continuous** if for all $\varepsilon > 0$ and for all $x \in X$, there exists $\delta > 0$, such that for all $y \in X$ with $|x - y| < \delta$, we have

$|f(x) - f(y)| < \varepsilon$.

Definition 1.9. We say f is **uniformly continuous** if for all $\varepsilon > 0$, there exists $\delta > 0$, such that for all $x, y \in X$ with $|x - y| < \delta$, we have $|f(x) - f(y)| < \varepsilon$.

Remark. Clearly if f is uniformly continuous, then f is continuous. The converse is not true.

Example 1.4. Consider $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$. Then f is continuous as it is a polynomial. Suppose $\delta > 0$. Then,

$$f(x + \delta) - f(x) = (x + \delta)^2 - x^2 = 2\delta x + \delta^2 \to \infty$$

as $x \to \infty$. So the condition fails for $\varepsilon = 1$.

Even on the bounded interval $(0, 1)$, take $f(x) = 1/x$. This is clearly continuous, but cannot be uniformly continuous as it approaches infinity as x approaches 0.

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 continuous but not uniformly continuous. Then we can find an $\varepsilon > 0$ such that, for all $\delta > 0$, there exist $x, y \in [a, b]$ with $|x-y| < \delta$ but $|f(x)-f(y)| \geq \varepsilon$. In particular, take $\delta = 1/n$.

Thus, we can find sequence $(x_n), (y_n)$ in $[a, b]$ with $|x_n - y_n| < 1/n$ but $|f(x_n)-f(y_n)| \geq \varepsilon$. The sequence (x_n) is bounded, so by Bolzano-Weierstrass it has a convergent subsequence $x_{n_j} \to x$. Since $[a, b]$ is closed, $x \in [a, b]$.

Then $x_{n_j} - y_{n_j} \to 0$, so also $y_{n_j} \to x$. But f is continuous at x, so there exists $\delta > 0$ such that for all $y \in [a, b]$, $|y - x| < \delta$ implies $|f(y) - f(x)| < \varepsilon/2$. Take such a δ. As $x_{n_j} \to x$, we can find J_1 such that $j \geq J_1$ implies $|x_{n_j} - x| < \delta$. Similarly, we can find J_2 such that for $j \geq J_2$, $|y_{n_j} - x| < \delta$. Let $j = \max\{J_1, J_2\}$. Then we have $|f(x_{n_j}) - f(x)|, |f(y_{n_j}) - f(x)| < \varepsilon/2$. But by triangle inequality,

$$|f(x_{n_j}) - f(y_{n_j})| \leq |f(x_{n_j}) - f(x)| + |f(x) - f(y_{n_j})| < \varepsilon,$$

a contradiction.

Corollary 1.1. *A continuous real-valued function on a closed bounded interval is bounded.*

Proof: Let $f : [a, b] \mapsto \mathbb{R}$ be continuous, and so uniformly continuous. Then we can find $\delta > 0$ such that for all $x, y \in [a, b]$, $|x - y| < \delta$ implies $|f(x) - f(y)| < 1$. Let $M = \lceil (b - a)/\delta \rceil$. Then for any $x \in [a, b]$, we can find $a = x_0 \leq x_1 \leq \ldots \leq x_M = x$, with $|x_i - x_{i-1} < \delta$. Then we have

$$|f(x)| \leq |f(a)| + \sum_{i=1}^{M} |f(x_i) - f(x_{i-1})| < |f(a)| + M.$$

Corollary 1.2. *A continuous real-valued function on a closed bounded interval is integrable.*

Proof: Let $f : [a, b] \to \mathbb{R}$ be continuous, and so uniformly continuous. Let $\varepsilon > 0$. Then we can find $\delta > 0$ such that for all $x, y \in [a, b]$, $|x - y| < \delta$, we have $|f(x) - f(y)| < \varepsilon$. Let $\mathcal{D} = \{x_0 < x_1 < \ldots < x_n\}$ be a dissection such that $x_i - x_{i-1} < \delta$, and $i \in \{1, \ldots, n\}$. Then for any $u, v \in [x_{i-1}, x_i]$, we have $|u - v| < \delta$, so $|f(u) - f(v)| < \varepsilon$. Hence

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

This gives

$$S(f, \mathcal{D}) - s(f, \mathcal{D}) \leq \sum_{i=1}^{n} (x_i - x_{i-1})\varepsilon = \varepsilon(b - a).$$

But this can be made arbitrarily small, so by Riemann's criterion, f is integrable over $[a, b]$.

2 Metric Spaces

2.1 Definitions and Examples

Our goal is to generalize the idea of convergence. This requires a notion of distance.

We have seen in $\mathbb{R}$, we have a norm $|x-y|$, in $\mathbb{R}^2$ we have $\|x-y\|$, and in function space, we can take

$$\sup_{x\in X}|f(x)-g(x)|.$$

We have seen that the triangle inequality is very useful, so we wish to preserve this property.

Definition 2.1. A **metric space** is a set X endowed with a **metric** d, i.e. a function $d : X^2 \to \mathbb{R}$, satisfying:

(i) $d(x,y) \geq 0$ for all $x,y \in X$, with equality if and only if $x=y$;

(ii) $d(x,y) = d(y,x)$ for all $x,y \in X$;

(iii) $d(x,z) \leq d(x,y) + d(y,z)$ for all $x,yz \in X$.

We could define a metric space as an ordered pair (X,d), but usually it is obvious what d is, so we often refer to the metric space as the set X.

Example 2.1.

(i) If $X = \mathbb{R}$, we have the usual metric $d(x,y) = |x-y|$.

(ii) If $X = \mathbb{R}^n$, we can take the Euclidean metric

$$d(x,y) = \|x-y\| = \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}.$$

(iii) Uniform convergence might not work. We wish to take $d(f,g) = \sup|f-g|$, but this might not exist if $f-g$ is unbounded. However, with the appropriate subspace of functions, we can take this metric. Let $Y \subset \mathbb{R}$, and take

$$X = B(Y) = \{f : Y \to \mathbb{R} \mid f \text{ bounded}\},$$

with the **uniform metric**

$$d(f,g) = \sup_{x\in Y}|f(x)-g(x)|.$$

We can check the triangle inequality: if $f, g, h \in B(Y)$, then for all $x \in Y$,

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

Taking the sup over all $x \in Y$, we get

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

Remark. Suppose (X, d) is a metric space and $Y \subset X$. Then $d|_Y$ is a metric on Y. We say Y with this metric is a **subspace** of X.

Example 2.2.

(i) We can take $\mathbb{Q}, \mathbb{Z}, \mathbb{N}, [0, 1]$ as subspaces of $\mathbb{R}$.

(ii) A continuous function on a bounded interval is bounded, so $\mathcal{C}([a, b])$ is a subspace of $B([a, b])$, with the uniform metric.

(iii) We can take the empty metric space $X = \emptyset$ with the empty metric.

Moreover, we can define different metrics on the same set.

Example 2.3.

(i) We can take the l_1 metric on $\mathbb{R}^n$:

$$d(x, y) = \sum_{i=1}^{n} |x_i - y_i|.$$

(ii) We can also take the l_∞ metric on $\mathbb{R}^n$:

$$d(x, y) = \max_i |x_i - y_i|.$$

(iii) On $\mathcal{C}([a, b])$, we can define the L_1 metric

$$d(f, g) = \int_a^b |f - g|.$$

(iv) If $X = \mathbb{C}$, we can define a metric

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

We can check that the triangle inequalitys. This is known as the British Rail metric or SNCF metric.

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

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

This is called the discrete metric on X.

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

$$d(x, y) = \begin{cases} 0 & x = y, \\ p^{-a} & p^a \| x - y. \end{cases}$$

We show the triangle quality holds. If any of x, y, z are the same, this is easy, so assume all of x, y, z are distinct. Let $x - y = p^a m$, $y - z = p^b n$. Then if $a \leq b$, we have

$$x - z = (x - y) + (y - z) = p^a(m + p^{b-a}n).$$

Hence $p^a \mid x - z$, so $d(x, z) \leq p^{-a}$.

Definition 2.2. 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 $x_n \to x$, if for all $\varepsilon > 0$, there exists N such that for all $n \geq N$,

$$d(x_n, x) < \varepsilon.$$

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

Proposition 2.1. *Limits are unique. That is, if (X, d) is a metric space, (x_n) is 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,

$$\begin{aligned} d(x,y) &\le d(x,x_n) + d(x_n,y) \le d(x_n,x) + d(x_n,y) \\ &\to 0 + 0 = 0. \end{aligned}$$

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.

1. This justifies talking about the limit of a convergent sequence in a metric space, and writing $x = \lim x_n$.
2. Constant sequences and eventually constant sequences converge.
3. 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 (x_n) also converges in X to x.

   However the converse is false: take the reals, then $1/n \to 0$. But if we consider the subspace $\mathbb{R}\setminus\{0\}$, then $(1/n)$ is a sequence, but does not converge in $\mathbb{R}\setminus\{0\}$.

Example 2.4. Let d be the Euclidean metric on $\mathbb{R}^n$. Then we have $x_n \to x$ if and only if the sequence converges in each coordinate in the usual way in $\mathbb{R}$. Let's consider other metrics, such as the uniform metric

$$d_\infty(x,y) = \max_i |x_i - y_i|, \text{ then}$$

$$d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \le \sqrt{\sum_{i=1}^n d_\infty(x,y)^2}.$$

So $d(x,y) \le \sqrt{n} d_\infty(x,y)$. But also $d_\infty(x,y) \le d(x,y)$. So for (x_n) in $\mathbb{R}^n$,

$$d(x_n,x) \to 0 \iff d_\infty(x_n,x) \to 0.$$

So the same sequences converge in $(\mathbb{R}^n, d)$ and $(\mathbb{R}^n, d_\infty)$. Similarly, for

$$d_1(x,y) = \sum_{i=1}^n |x_i - y_i|, \text{ then}$$

$$d_\infty(x,y) \le d_1(x,y) \le n d_\infty(x,y).$$

Consider $X = \mathcal{C}([0,1])$. Let d_∞ be the uniform metric on X, so

$$d_\infty(f,g) = \sup_{x\in[0,1]} |f(x) - g(x)|, \text{ so}$$

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

Similarly we can take the L_1 metric

$$d_1(f,g) = \int_0^1 |f-g|, \text{ then}$$

$$d_1(f,g) = \int_0^1 |f-g| \le \int_0^1 d_\infty(f,g) = d_\infty(f,g).$$

So we can prove $f_n \to f$ in (X, d_∞) implies $f_n \to f$ in (X, d_1). But the converse is not true, from our previous examples on uniform convergence.

We can also take (X,d) a discrete metric. Consider a convergence sequence $x_n \to x$. Then letting $\varepsilon = 1$, the definition of convergence says for all $n \ge N$, $d(x_n, x) < 1$, so $x_n = x$. Thus (x_n) is eventually constant. So in a discrete metric, (x_n) converges if and only if (x_n) is eventually constant.

Definition 2.3. Let (X,d) and (Y,e) be metric spaces, and let $f : X \to Y$.

(i) Let $a \in X$ and $b \in Y$. We say $f(x) \to b$ as $x \to a$ if for all $\varepsilon > 0$, there exists $\delta > 0$, such that for all $x \in X$,

$$0 < d(x,a) < \delta \implies e(f(x), b) < \varepsilon.$$

(ii) Let $a \in X$. We say f is **continuous** if $f(x) \to f(a)$ as $x \to a$.

(iii) If for all $a \in X$, f is continuous, then we say f is a continuous function.

(iv) We say f is uniformly continuous if for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x, y \in X$,

$$d(x,y) < \delta \implies e(f(x), f(y)) < \varepsilon.$$

(v) Suppose $W \subset X$. We say f is continuous on W (resp. uniformly continuous on W) if the function $f|_W$ is continuous (resp. uniformly continuous), as a function $W \to Y$.

Remark.

1. We don't have a nice rephrasing of (i) in terms of concepts in the reals: we want something like

$$e(f(x), b) \to 0 \text{ as } d(x, a) \to 0,$$

   but this is meaningless.V

2. (i) says nothing about what happens at the point a itself. For example, $f : \mathbb{R} \to \mathbb{R}$ as $f(x) = \delta_{0x}$ tends to 0 as $x \to 0$. If we have f continuous, then $d(x, a) = 0$ implies $e(f(x), f(a)) = 0$, so we may take $0 \leq d(x, a) < \delta$.

3. We can rewrite (v): f is continuous on W if and only if $f|_W$ is a continuous function if and only if for all $a \in W$ and $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x \in W$,

$$d(x, a) < \delta \implies e(f(x), f(a)) < \varepsilon.$$

   In particular, note that this only considers points in W. For example,

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

   is continuous on $[0, 1]$, but not continuous at 0 or 1.

Proposition 2.2. *Let (X, d), (Y, e) be metric spaces. Let $f : X \to Y$ and $a \in X$. Then f is continuous at a if and only if whenever (x_n) is a sequence in X with $x_n \to a$, then $f(x_n) \to f(a)$.*

Proof: Suppose f is continuous at a. Let (x_n) be a sequence in X with $x_n \to a$. Let $\varepsilon > 0$. As f is continuous at a we can find $\delta > 0$ such that for all $x \in X$, $d(x, a) < \delta$ implies $e(f(x), f(a)) < \varepsilon$.

As $x_n \to x$, we can find N such that for all $n \geq N$, $d(x_n, a) < \delta$. Hence, for $n \geq N$, $e(f(x), f(a)) < \varepsilon$. This gives $f(x_n) \to f(a)$.

Now suppose f is not continuous at a. Then there is some $\varepsilon > 0$ such that for all $\delta > 0$, there exists $x \in X$ with $d(x, a) < \delta$ but $e(f(x), f(a0) \geq \varepsilon$. Take $\delta_n = 1/n$, to obtain a sequence (x_n) with $d(x_n, a) < 1/n$ but $(f(x_n), f(a)) \geq \varepsilon$. 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. Let $f : W \to X$, $g : X \to Y$ and let $a \in W$. Suppose f is continuous at a and g is continuous at $f(a)$. Then $g \circ f$ is continuous at a.*

Proof: Let (x_n) be a sequence in W with $x_n \to a$. Then, $f(x_n) \to f(a)$, and so also $g(f(x_n)) \to g(f(a))$. So $g \circ f$ is continuous at a.

Example 2.5.

1. Consider $\mathbb{R} \to \mathbb{R}$ with the usual metric. This is the same metric as defined for $\mathbb{R}$ only. We know many continuous function on $f : \mathbb{R} \to \mathbb{R}$, such as polynomials, trigonometric functions, exponential functions, etc.

2. 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 identity function) is continuous.

3. Consider $\mathbb{R}^n$ with the EUclidean 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. Let's denote a sequence in $\mathbb{R}^n$ by $(x^{(m)})_{m \geq 1}$. We known that $x^{(m)} \to x$ if and only if for each i, $x_i^{(m)} \to x_i$. Hence π_i is continuous.

   Similarly, suppose $f_1, \ldots, f_n : \mathbb{R} \to \mathbb{R}$, and let $f : \mathbb{R} \to \mathbb{R}^n$ by $f(x) = (f_1(x), \ldots, f_n(x))$. Then f is continuous at a point if and only if all of $f_1, \ldots, f_n$ are.

   Using these facts, we can show that $f(x, y, z) = (e^{-x} \sin y, 2x \cos z)$ is continuous.

4. Recall that if we have the Euclidean metric, the l_1 metric or the l_∞ metric on $\mathbb{R}^n$, then convergent sequences are the same in each case. So continuous functions $X \to \mathbb{R}^n$ or $\mathbb{R}^n \to Y$ are the same with each of these metrics.

5. Let (X, d) be a discrete metric space and let (Y, e) be any metric space. Then all functions $f : X \to Y$ are continuous. Suppose $a \in X$ and (x_n) is a sequence in X with $x_n \to a$. Then (x_n) is eventually constant, so $f(x_n) \to f(a)$.

2.2 Completeness

We saw a version of the general principle of convergence held in each of the three examples we considered. We try to extend this to all metric spaces:

Definition 2.4. Let (X, d) be a metric space and let (x_n) be a sequence in X. We say (x_n) is **Cauchy** if for all $\varepsilon > 0$, then there exists N such that for all $m, n \geq N$, $d(x_m, d_n) < \varepsilon$.

It is easy to show that if (x_n) is convergent, then (x_n) is Cauchy, but the converse is not true in general.

For example, let $X = \mathbb{R} \setminus \{0\}$, and let $x_n = 1/n$. Then the (x_n) do not converge, but are Cauchy as they are Cauchy in $\mathbb{R}$.

Similarly, we can consider $\mathbb{Q}$, then this does not satisfy the general principle of convergence.

Definition 2.5. Let (X, d) be a metric space. We say X is **complete** if every Cauchy sequence in X converges.

Example 2.6.

1. $\mathbb{R} \setminus \{0\}$ and $\mathbb{Q}$ are not complete.
2. $\mathbb{R}$ with the usual metric is complete.
3. The general principle of convergence for $\mathbb{R}^n$ say $\mathbb{R}^n$ with the Euclidean metric is complete.
4. 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. Indeed, let (f_n) be a Cauchy sequence in $B(X)$. Then (f_n) is uniformly Cauchy, so by the general principle of uniform convergence, it is uniformly convergent, so $f_n \to f$ uniformly for some $f : X \to \mathbb{R}$.

   This gives $f_n - f$ is bounded for n sufficiently large. Take such an n, and then since f_n bounded, $f = f_n - (f_n - f)$ implies f is bounded, so $f \in B(X)$. Finally, $f_n \to f$ uniformly implies $d(f_n, f) \to 0$, so $f_n \to f$ in $(B(X), d)$.

Remark. This is a typical example of a proof that a given space (X, d) is complete:

(i) Take (x_n) Cauchy in X.

(ii) Construct a limit object where it seems $(x_n) \to x$.

(iii) Show $x \in X$.

(iv) Show $x_n \to x$ in the metric space (X, d).

It is important to do things in this order, as we cannot talk about $d(x_n, x)$ until we known $x \in X$.

5. If $[a, b]$ is a closed interval, then $\mathcal{C}([a, b])$ with the uniform norm d is complete. Indeed, let (f_n) be Cauchy in $\mathcal{C}([a, b])$. Then since $\mathcal{C}([a, b]) \subset B([a, b])$, and $B([a, b])$ is complete, then $f_n \to f$ for some $f \in B([a, b])$.

   Each function is continuous, and $f_n \to f$ uniformly, so f is continuous, giving $f \in \mathcal{C}([a, b])$. Finally, $f_n \to f$ uniformly gives $d(f_n, f) \to 0$.

Definition 2.6. Let (X, d) be a metric space and $Y \subset X$. We say Y is **closed** if whenever (x_n) is 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. This makes sense: if $Y \subset X$, then Y itself is a metric space as a subspace of X, so we can say Y is complete.

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, $x_n \to x$ for some $x \in X$.

(iii) $Y \subset X$ is closed, so $x \in Y$.

(iv) Finally, we have for each $x_n \in Y$, $x \in Y$, and $x_n \to x$ in X, so $d(x_n, x) \to 0$, giving $x_n \to x$ in Y.

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

Then since $\sum |x_n|$, $\sum |y_n|$ converge and $|x_n - y_n| \leq |x_n| + |y_n|$, by comparison test $\sum |x_n - y_n|$ converges, so d is well-defined. It is easy to check that d is a metric. Then (l_1, d) is complete. Indeed, let $(x^{(n)})$ be a Cauchy sequence in l_1, so $(x_i^{(n)})$ is a sequence in $\mathbb{R}$.

Then for each i, $(x_i^{(n)})$ is a Cauchy sequence in $\mathbb{R}$, since if $y, z \in l_1$, then $|y_i - z_i| \leq d(y, z)$. But $\mathbb{R}$ is complete, so we can find $x_i \in \mathbb{R}$ with $x_i^{(n)} \to x_i$ as $n \to \infty$. Let $x = (x_1, x_2, x_3, \dots,) \in \mathbb{R}^{\mathbb{N}}$.

We next show that $x \in l_1$. Given $y \in l_1$, define

$$\sigma(y) \sum_{i=1}^{\infty} |y_i|,$$

i.e. $\sigma(y) = d(y, \bar{0})$, where $\bar{0}$ is the constant zero sequence. Now we have, for any m, n,

$$\sigma(x^{(m)}) = d(x^{(m)}, z) \leq d(x^{(m)}, x^{(n)}) + d(x^{(n)}, \bar{0}) = d(x^{(m)}, x^{(n)}) + \sigma(x^{(n)}).$$

This gives $\sigma(x^{(m)}) - \sigma(x^{(n)}) \leq d(x^{(m)}, x^{(n)})$. Similarly, we can swap around m, n, to give $|\sigma(x^{(m)}) - \sigma(x^{(n)})| \leq d(x^{(m)}, x^{(n)})$, so $(\sigma(x^{(m)}))$ is a Cauchy sequence in $\mathbb{R}$, and so converges to K. Now we claim for any $I \in \mathbb{N}$,

$$\sum_{i=1}^{I} |x_i| \leq K + 2.$$

Indeed, as $\sigma(x^{(n)}) \to K$ as $n \to \infty$, we can find N_1 such that for all $n \geq N_1$,

$$\sum_{i=1}^{I} |x_i^{(n)}| \leq \sum_{i=1}^{\infty} |x_i^{(n)}| \leq K + 1.$$

For all $i \in \{1, 2, \dots, I\}$, we have $x_i^{(n)} \to x$, so we can find N_2 such that for all $n \geq N_2$ and $i \in \{1, 2, \dots, I\}$ we have $|x_i^{(n)} - x_i| < 1/I$. Letting $n = \max N_1, N_2$, we have

$$\sum_{i=1}^{I} |x_i| \leq \sum_{i=1}^{I} |x_i^{(n)}| + \sum_{i=1}^{I} |x_i^{(n)} - x_i| \leq K + 1 + 1 = K + 2.$$

Since the partial sums $\sum |x_i|$ are increasing and bounded above, they converge.

Finally, we need to check $x^{(n)} \to x$ as $n \to \infty$ in l_1, i.e. $d(x^{(n)}, x) \to 0$.

Index