Sequences of Functions

As promised, we're going to apply some of our topological technology to metric spaces of functions. The first question to ask is: which metric? Equivalently, we can ask: what should it mean for a sequence of functions to converge?

Ex: For each \( n \in \mathbb{N} \), define \( f_n : [0,1] \to \mathbb{R} \) by

\[
x \mapsto x^n
\]

Each function is continuous, and indeed uniformly continuous. Let's try and take a limit of \( (f_n)_{n \in \mathbb{N}} \).

Fix \( x \in [0,1] \). What is \( \lim_{n \to \infty} f_n(x) \)?

- **Case 1:** \( x = 1 \). Then \( f_n(x) = 1 \). So \( \lim_{n \to \infty} f_n(x) = 1 \)
- **Case 2:** \( x < 1 \). Then \( f_n(x) = x^n \), so \( \lim_{n \to \infty} f_n(x) = 0 \)

So \( \lim_{n \to \infty} f_n(x) = \begin{cases} 1 & \text{if } x = 1 \\ 0 & \text{if } x < 1 \end{cases} \)

Def.: Given a sequence of functions \( f_n : A \to (\mathbb{N}, \rho) \), if for each \( x \in A \), we have \( \lim_{n \to \infty} f_n(x) = f(x) \), we say \( (f_n)_{n \in \mathbb{N}} \) converges pointwise to \( f \).

Write \( f_n \to f \) on \( A \).

Our example shows that pointwise convergence isn't all that.

Pointwise limits of sequences of continuous functions need not be continuous!

"[0,1] is not closed under pointwise limits."

We can understand the problem with this family as a problem of interchange of limits.

\[
\lim_{n \to \infty} \lim_{x \to 1} x^n = \lim_{x \to 1} (\lim_{n \to \infty} x^n) = 0
\]

The homework has several similar examples.
Let's unpack the quantities in the definition of pointwise convergence.

\( f_n \rightarrow f \text{ on } A \) if

\[ \forall x \in A \forall \varepsilon > 0 \exists N \in \mathbb{N} \ni \forall n \geq N, \rho(f_n(x), f(x)) < \varepsilon. \]

If we swap a \( \forall \) and a \( \exists \), we get:

\[ \forall \varepsilon > 0 \exists N \in \mathbb{N} \ni \forall x \in A, \forall n \geq N, \rho(f_n(x), f(x)) < \varepsilon. \]

**Definition:** Given \( f_n : A \rightarrow (\mathbb{N}, \rho) \), we say the sequence \((f_n)_{n \in \mathbb{N}}\) converges uniformly to \( f \) on \( A \) if for any \( \varepsilon > 0 \), there is \( N \in \mathbb{N} \) so that \( n \geq N \) guarantees \( \rho(f_n(x), f(x)) < \varepsilon \). We write \( f_n \rightarrow f \text{ on } A \) or \( f_n \rightarrow f \text{ (unif.) on } A \).

**Example:** The functions \( f_n : x \mapsto x^n \) do not converge uniformly on \([0, 1]\).

For each \( n \in \mathbb{N} \), let \( y_n = \sqrt[n]{\frac{1}{2}} \). Then \( f_n(y_n) = \frac{1}{2} \).

So each \( f_n \) takes the value \( \frac{1}{2} \) which is \( \frac{1}{2} \) away from both 0 and 1 (which are the only candidates for the values of a uniform limit).

**Proposition (Uniform Convergence Implies Pointwise Convergence):**

\( f_n \rightarrow f \text{ on } A \Rightarrow f_n \rightarrow f \text{ on } A \)

**Proposition (Uniform Cauchy Criterion):**

Suppose \((N, \rho)\) is complete. Then \( f_n : A \rightarrow (N, \rho) \) have

\( f_n \rightarrow f \text{ if and only if } \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n, m \geq N, \forall x \in A, \rho(f_n(x), f_m(x)) < \varepsilon. \)

**Theorem:** Suppose \( f_n \rightarrow f \) and \( f \) are all continuous. Then \( f \) is continuous.

This will follow from the following theorem about limits:
Theorem. Suppose \( f_n : A \rightarrow (N, p) \) are a sequence of maps

- \( \lim_{a \to x_0} f_n(a) = L_n \)
- \( f_n \rightarrow f \) on \( A \)
- \( \lim_n L_n = L \)

Then \( \lim_{a \to x_0} f(a) = L \).

That is, \( \lim_n \lim_{a \to x_0} f(a) = \lim_{a \to x_0} \lim_n f_n(a) \).

Proof. Recall the sequential characterization of the limit of a function:

Proposition ("Two-Poler Test")

\[
\lim_{a \to x_0} f(a) = b \quad \text{iff} \quad \text{for any sequence } (x_n)_{n \in \mathbb{N}} \text{ in } A \setminus \{x_0\}
\]

with \( x_n \to x_0 \) we have \( f(x_n) \to b \).

We want to show that \( \lim_{a \to x_0} f(a) = L \), so we'll let \( (x_k)_{k \in \mathbb{N}} \) be a sequence in \( A \) with \( x_k \to x_0 \), and show that \( f(x_k) \to L \).

We need to estimate \( p(f(x_k), L) \). The triangle (quadriple?) inequality says

\[
p(f(x_k), L) \leq p(f(x_k), f_n(x_k)) + p(f_n(x_k), L_n) + p(L_n, L)
\]

(as shown with black arrows in the figure)
More niceties of uniform convergence

Theorem. Suppose \( f : [a,b] \to \mathbb{R} \) are integrable and \( f_n \to f \) on \([a,b]\). Then \( f \) is integrable and
\[
\lim_{n \to \infty} \int_a^b f_n(x) \, dx = \int_a^b f(x) \, dx.
\]

Proof. We're asked to show \( f \) is integrable. So given \( \varepsilon > 0 \), \( \exists N \in \mathbb{N} \) guarantees, \( \forall x \in [a,b] \),
\[
|f_n(x) - f(x)| < \frac{\varepsilon}{b-a}.
\]
So for such \( n \), \( \sup_{x \in [a,b]} f_n(x) + \frac{\varepsilon}{b-a} \) on any interval \([k, \ell] \subseteq [a,b] \). So for any partition \( P \), \( n \geq N \) guarantees
\[
U(f,P) - L(f,P) \leq U(f_n,P) + \frac{\varepsilon}{b-a} \quad \text{and sim.}
\]
\[
L(f,P) \leq L(f_n,P) - \frac{\varepsilon}{b-a}.
\]
So
\[
U(f,P) - L(f,P) \leq U(f_n,P) - L(f_n,P) + \frac{\varepsilon}{b-a}.
\]
Now fixing a particular \( n \), there is some \( P(n, \varepsilon) \) so that
\[
U(f_n, P(n, \varepsilon)) - L(f_n, P(n, \varepsilon)) \leq \frac{\varepsilon}{b-a}.
\]
Thus
\[
U(f, P(n, \varepsilon)) - L(f, P(n, \varepsilon)) \leq \varepsilon.
\]
So \( f \) is integrable.

Now consider the function \( f(x) - f_n(x) \), which is integrable. If \( n \geq N \), we have

\[
\left| \int_a^b f_n(x) \, dx - \int_a^b f(x) \, dx \right| = \left| \int_a^b (f_n(x) - f(x)) \, dx \right|
\]

\[
\leq \int_a^b |f_n(x) - f(x)| \, dx \leq \int_a^b \frac{\varepsilon}{\epsilon(b-a)} \, dx
\]

\[
= \frac{\varepsilon}{\epsilon(b-a)} < \varepsilon.
\]

Thus \( \lim_{n \to \infty} \int_a^b f_n(x) \, dx = \int_a^b f(x) \, dx. \)

**Note:** Uniform convergence is necessary here.

**Example:** Let \( f_n(x) = \begin{cases} \frac{n^2}{n^2} & x \in [n^{-1}, n] \\ 1 & x \in (n^{-1}, n) \end{cases} \)

Then each \( f_n \) is continuous and \( \int_0^1 f_n(x) \, dx = 1. \)

Moreover, \( f_n(0) = 0 \) and if \( x > 0 \) there is \( n \) large enough that \( \frac{x}{n} < x \), so \( f_n(x) = 0. \)

Thus \( f_n \to 0 \) pointwise.

\[
\lim_{n \to \infty} \int_a^b f_n(x) \, dx = \int_a^b \lim_{n \to \infty} f_n(x) \, dx = 1 - 0 = 1
\]

\[
\int_0^1 \lim_{n \to \infty} f_n(x) \, dx = 0
eq 1!
\]

**Theorem:** Suppose \( f_n : [a,b] \to \mathbb{R} \) are differentiable and \( f_n \to f \).

Suppose also \( f_n \) are continuous and \( f_n \to g \). Then \( f \) is differentiable and \( f' = g. \)

**Proof:** Since \( f_n \) are continuous, we have, for any particular \( x_0 \),

\[
f_n(x) = f_n(x_0) + \int_{x_0}^x f_n'(t) \, dt
\]

Since \( f_n \to f \) pointwise, we have
\[
f(x) = \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \left( f_n(x) + \int_{x_0}^{x} f_n'(t) \, dt \right) \\
= \lim_{n \to \infty} f_n(x) + \lim_{n \to \infty} \int_{x_0}^{x} f_n'(t) \, dt \\
= f(x) + \int_{x_0}^{x} g(t) \, dt 
\]

Since each \( f_n' \) is \( \text{cts} \), \( g \) is \( \text{cts} \). So \( f \) is \( \text{diff} \), and \( f'(x) = g(x) \)

Note: It is possible to have \( f_n \) \( \text{diffe} \), \( f_n \to f \), but \( f \) not \( \text{diffe} \).

E.g. Set \( f_n(x) = \sqrt{x^2 + \frac{1}{n}} \), \( f(x) = |x| \).

Then \( f_n(x) = \sqrt{x^2 + \frac{1}{n}} \leq \sqrt{x^2 + \frac{1}{n}} = |x| + \frac{1}{n} \)

So \( |f_n(x) - f(x)| = f_n(x) - f(x) \leq \frac{1}{n} \).

Given \( \varepsilon \), choose \( N = \frac{1}{\varepsilon^2} + 1 \). Then \( n > N \) guarantees \( |f_n(x) - f(x)| < \varepsilon \). Thus \( f_n \to f \) on \( \mathbb{R} \).

Each \( f_n \) is \( \text{diff} \). (In fact, \( f_n'(x) = \frac{x}{\sqrt{x^2 + \frac{1}{n}}} \) are uniformly bounded.)

But \( f \) is not \( \text{diff} \).

---

**Proposition (Weierstrass M-test for sequences)**

Given \( f_n : A \to \mathbb{R} \) with \( f_n \to f \) pointwise on \( A \),

Set \( M_n = \sup_{x \in A} |f_n(x) - f(x)| \).

Then \( f_n \to f \) iff \( M_n \to 0 \).

What does \( \sup_{x \in A} |f_n(x) - f(x)| \) measure?

**Stone-Weierstrass Theorem (baby version)**

Let \( f : [a,b] \to \mathbb{R} \) be a \( \text{cts} \) function.

Then there is a sequence \( (P_n)_{n=1}^{\infty} \) of polynomials with \( P_n \to f \) on \( [a,b] \).

**Proof.** We'll just prove it for \( c = 0, b = 1 \).

WLOG we may assume \( f(1) = f(0) = 0 \).

Extend \( f \) to \( \mathbb{R} \) by setting \( f = 0 \) outside \([0,1] \).

Then \( f : \mathbb{R} \to \mathbb{R} \) is \( \text{unif} \) \( \text{cts} \).

Now define \( Q_n(x) = C_n (1-x)^n \), where \( C_n = \frac{1}{\int_{-1}^{1} (1-x)^n \, dx} \), so that \( \int_{-1}^{1} Q_n(x) \, dx = 1 \).
Now it's not hard to compute that \( c_n < \sqrt{n} \).

Set \( P_n(x) = \int_{-1}^{1} f(x+t) Q_n(t) \, dt \).

Problem: this isn't a polynomial in \( x \) (at least, not obviously so).

\[
\int_{-1}^{1} f(x+t) Q_n(t) \, dt = \int_{-x}^{x} f(s) Q_n(s-x) \, ds
\]

which is obviously a polynomial in \( x \).

Now we want to show \( P_n \rightarrow f \). Given \( \varepsilon > 0 \), choose \( \delta > 0 \) so that \( |x-y| < \delta \Rightarrow |f(x) - f(y)| < \frac{\varepsilon}{3} \).

Set \( M = \sup |f| \).

\[
|P_n(x) - f(x)| = |\int_{-1}^{1} f(x+t) Q_n(t) \, dt - \int_{-1}^{1} f(x) Q_n(t) \, dt|
\]

\[
= |\int_{-1}^{1} (f(x+t) - f(x)) Q_n(t) \, dt| \leq \int_{-1}^{1} |f(x+t) - f(x)| Q_n(t) \, dt
\]

\[
= \int_{-\delta}^{\delta} |f(x+t) - f(x)| Q_n(t) \, dt + \int_{-\delta}^{-1} |f(x+t) - f(x)| Q_n(t) \, dt
\]

Now when \( |t| < \delta \), we have \( |f(x+t) - f(x)| < \frac{\varepsilon}{3} \).

So
\[
\int_{-\delta}^{\delta} |f(x+t) - f(x)| Q_n(t) \, dt \leq \frac{\varepsilon}{3} \int_{-\delta}^{\delta} Q_n(t) \, dt \leq \frac{\varepsilon}{3}
\]

The other terms can be estimated since \( |f| \leq M \) and \( x^2 \geq \delta^2 \):

\[
\int_{-\delta}^{-1} |f(x+t) - f(x)| Q_n(t) \, dt \leq 2M \sup_{-\delta^2} (-t^2)^n \leq 2M \delta \delta^n
\]

\[
\int_{-1}^{1} |f(x+t) - f(x)| Q_n(t) \, dt \leq 2M \sup_{-\delta^2} (1-t^2)^n \leq 2M \delta \delta^n
\]

Lemma. For \( \delta > 0 \), \( \sqrt{n} (1-\delta^n) \rightarrow 0 \) as \( n \rightarrow \infty \).

So we can take \( n \) large enough that

\[
\frac{\varepsilon}{4M \sqrt{n} (1-\delta^n)} < \frac{\varepsilon}{2}.
\]

Then \( |P_n(x) - f(x)| < \varepsilon \).
The Space of Continuous Maps

Definition: Given $A \subset (\mathbb{R}, d)$, set $B(A) = \{ f: A \to \mathbb{R} \mid f \text{ is bounded}\}$.

We proved (in the special case $A = [0,1]$) in homework that $B(A)$ has a norm:

$$\|f\|_\infty = \sup_{x \in A} |f(x)|$$

(We also showed for HW that this norm doesn't come from an inner product.)

Given a normed space $(N, \|\cdot\|)$, we can define $B(A; N) = \{ f: A \to N \mid f \text{ is bounded} \}$ and equip this space with $\|f\| = \sup_{x \in A} \|f(x)\|$ to get a normed space.

Theorem: Let $(f_n)_{n \in \mathbb{N}}$ be a sequence of bounded functions $A \to N$. Then $f_n \to f$ in $B(A; N)$ iff $f_n \to f$ on $A$.

Proof: It's just the M-test (which you proved in HW.)

$$f_n \to f \iff M_n \to 0$$

$$\iff \forall \epsilon > 0 \exists N \ni n \geq N \Rightarrow M_n < \epsilon$$

$$\iff f_n \to f \text{ in } B(A; N). \quad \square$$

Definition: $C_0(A; N) = \{ f: A \to N \mid f \text{ is cont. and bounded} \}$.

Then $C_0(A; N)$ is a vector subspace of $B(A; N)$.

(If $A$ is compact, $C(A; N) = C_0(A; N)$. Why?)

From here, we'll assume that $C(A; N)$ is equipped with the metric $d_0(f, g) = \|f - g\|_{\infty}$, the uniform metric.

Some rephrasings

Stone-Weierstrass: Polynomials are dense in $C([a,b]; \mathbb{R})$

Another Theorem: $C_0(A; N)$ is a closed subset of $B(A; N)$.

Another Theorem: $I: C([a,b]; \mathbb{R}) \to \mathbb{R}$ is continuous.

$$f \mapsto \int_a^b f$$
Consider \( B = \{ f \in C([0,1]; \mathbb{R}) \mid f(x) > 1 \text{ for all } x \in [0,1] \} \).

**Claim** \( B \) is open in \( C([0,1]; \mathbb{R}) \).

**Proof.** Let \( f \in B \). By Max-Min, \( \exists x_0 \in [0,1] \) with \( f(x) > f(x_0) \) for all \( x \in [0,1] \). Also notice \( f(x) > 1 \).

Set \( \epsilon = \frac{1}{2} (f(x_0) - 1) \). Then \( \|f - g\|_\infty < \epsilon \)

\[ \Rightarrow \forall x, |f(x) - g(x)| < \epsilon \]

\[ \Rightarrow \forall x, g(x) > f(x) - \epsilon = f(x) - \frac{1}{2} (f(x) - 1) \]

\[ = \frac{1}{2} (f(x) + 1) > 1 \]

So \( \|f - g\|_\infty < \epsilon \Rightarrow g \in B \). Thus \( B \) is open!

**Example.** The unit sphere in \( C([0,1]; \mathbb{R}) \) is not compact.

**Proof.** It suffices to exhibit a sequence of functions \( \{f_n\}_{n \in \mathbb{N}} \), with \( \|f_n\|_\infty = 1 \) but no subsequence of \( \{f_n\}_{n \in \mathbb{N}} \) converges uniformly.

Set \( f_n(x) = \begin{cases} 1 & x \in [0, \frac{1}{n}] \\ 2 - nx & x \in [\frac{1}{n}, \frac{2}{n}] \\ 0 & x \in [\frac{2}{n}, 1] \end{cases} \)

Then any subsequential pointwise limit is discontinuous.

Note that the unit sphere is by definition bounded, and since any norm is continuous in the metric it induces, the unit sphere is closed. Thus:

Hence-Bolz fails for \( C([0,1]; \mathbb{R}) \)!

How can we tell if a subset of \( C([0,1]; \mathbb{R}) \) is compact?

**Defn.** A collection ("family") of functions \( F \) is called equicontinuous if for any \( \epsilon > 0 \) there is \( \delta > 0 \) so that for any \( x, y \in A_j \) and any \( f \in F \), we have \( d(x,y) < \delta \Rightarrow \rho(f(x), f(y)) < \epsilon \)

**Immediate consequence.**

If \( F \) is an equicontinuous family, then each \( f \in F \) is uniformly continuous.

**Example.** Let \( L = \{ f : [0,1] \rightarrow \mathbb{R} \mid f \text{ is differentiable and } \sup|f'| \leq 2 \} \).

For each \( f \in L \), MVT gives \( |f(x) - f(y)| \leq 2 |x - y| \)

So we can take \( \delta = \frac{\epsilon}{2} \), which depends neither on \( x \) nor on \( f \). Thus \( L \) is equicontinuous.
A family $F$ of functions $A \to N$ is **compact** at $x \in A$ if the image set $F_x = \{ f(x) | f \in F \} \subseteq N$ is compact. $F$ is **pointwise compact** if $F$ is compact at $x$ for all $x \in A$.

**Example:**

- Unit sphere of $C([a,b];\mathbb{R})$?
- $\{ f \in C([a,b];\mathbb{R}) | f(x) > 1 \text{ for all } x \in [0,1] \}$?
- The set of all power functions $[0,1] \to \mathbb{R}$?
- The set of polynomials with constant term 0?

---

**Arzelà-Ascoli**

The Arzelà-Ascoli theorem gives a way to tell if sets of functions are compact with respect to $\sup$.

**Arzelà-Ascoli Theorem**

Suppose $A \subseteq (M,d)$ is compact, $(N,\mid \cdot \mid)$ complete. Then $F \subseteq C(A;N)$ is compact iff

- $F$ is uniformly closed
- $F$ is pointwise compact
- $F$ is equicontinuous

**Proof:** We'll show ($\Rightarrow$). Use Bolzano-Weierstrass and show such $F$ is sequentially compact.

First recall the following lemma from the proof of Stone-Weierstrass:

**Lemma:** If $A$ is compact, for each $\delta > 0$ there is a finite set $\{ y_1, \ldots, y_k \} \subseteq A$ so that $A \subseteq \bigcup_{i=1}^{k} B(y_i, \delta)$.

Set $C_A = \{ y_1, \ldots, y_k \}$ to be the set given by the lemma with $\delta^{-1}$. Let $C = \bigcup C_A$. Note that $C$ is a countable set.

Denote its elements $C = \{ x_1, x_2, x_3, \ldots \}$. Given a sequence $(f_n(x))_{n \in \mathbb{N}}$ of points of $F$. For each $x \in C$, note that $(f_n(x))_{n \in \mathbb{N}}$ is a sequence in $F$, which is...
compact. So some subsequence \( f_{k_n}(x) \) converges for each \( k \).

Apply this first to \( x \) to get a subsequence \( f_{k_n}(x) \to z_1 \).

From \( f_{k_n} \) select a subsequence \( f_{k_{2n}} \) with \( f_{k_{2n}}(x_2) \to z_2 \).

Since \( f_{k_{2n}} \) is a subsequence of \( f_{k_n} \), \( f_{k_{2n}}(x_1) \to z_1 \) as well.

In this manner, set
\[
\begin{align*}
f_{k_1}(x_1) & \to z_1, \\
f_{k_2}(x_2) & \to z_2, \\
f_{k_3}(x_3) & \to z_3, \\
& \vdots
\end{align*}
\]

where each \( f_{k_n} \) is a subsequence of \( f_{k_n} \) and each \( f_{k_n}(x_k) \to z_k \) for all \( k \leq p \).

Now consider the diagonal sequence \( g_n = f_{k_n} \).

Claim: \((g_n)_{n \in \mathbb{N}}\) is uniformly Cauchy.

Proof: Given \( \varepsilon > 0 \), let \( \delta > 0 \) be such that \( d(x,y) < \delta \Rightarrow \rho(f(x),y) < \frac{\varepsilon}{3} \). Choose \( p \) with \( \frac{1}{p} < \delta \). Then there are \( y_1, \ldots, y_p \) so that every point \( x \in A \) is at most \( \frac{1}{p} \) from one of the \( y_i \). Since each \((g_n(x))_{n \in \mathbb{N}}\) is convergent in \( N \), it is Cauchy in \( N \). So there are \( N_1, \ldots, N_k \) with \( m, n \geq N_i \)

\[
\rho(g_n(x), g_m(x)) \leq \rho(g_n(x), g_m(y_i)) + \rho(g_m(y_i), g_m(y_j)) + \rho(g_m(y_j), g_m(x))
\]

\( \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \)

Since \( N_1 \) depends only on \( \varepsilon \), and not on \( x \), we have shown \((g_n)_{n \in \mathbb{N}}\) is uniformly Cauchy.

Since \((N, ||\cdot||)\) is complete, \((g_n)_{n \in \mathbb{N}}\) is uniformly convergent.

Since \( F \) is uniformly closed, \( g_n \to g \) implies \( g \in F \).

So we have shown every sequence in \( F \) has a convergent subsequence, whose limit lies in \( F \), i.e. \( F \) is sequentially compact.