Complex numbers

(For those that took my class with my book where complex numbers were not treated. In Rudin a number is generally complex unless stated otherwise).

A complex number is just a pair $(x, y) \in \mathbb{R}^2$

We call the set of complex numbers $\mathbb{C}$. We identify $x \in \mathbb{R}$ with $(x, 0) \in \mathbb{C}$. The x-axis is then called the *real axis* and the y-axis is called the *imaginary axis*. The set $\mathbb{C}$ is sometimes called the *complex plane*.

Note to self: *DRAW complex plane*

Define

$$(x, y) + (s, t) := (x + s, y + t)$$
$$(x, y)(s, t) := (xs - yt, xt + ys)$$

We have a field. LTS to **check the algebraic properties**.

Generally, we write complex number (x, y) as $x + iy$, where we define

$$i := (0, 1)$$

We check that $i^2 = -1$. So we have a solution to the polynomial equation

$$z^2 + 1 = 0.$$

(Note that engineers use j instead of i)

We will generally use x, y, r, s, t for real values and z, w, ξ, ζ for complex values.

For $z = x + iy$, define

$$\operatorname{Re} z := x,$$
$$\operatorname{Im} z := y,$$

the real and imaginary parts of z.

No ordering like the real numbers!

If $z = x + iy$, write

$$\bar{z} := x - iy$$

for the *complex conjugate*. That is, a reflection across the real axis. Real numbers are characterized by the equation

$$z = \bar{z}.$$

When $z = x + iy$, then define *modulus* as

$$|z| = \sqrt{x^2 + y^2}.$$

Note that if we let d be the standard metric on $\mathbb{R}^2$ and hence $\mathbb{C}$, then

$$|z| = d(z, 0).$$

In fact,

$$|z - w| = d(z, w),$$

which looks precisely like the metric on $\mathbb{R}$. Hence we immediately have that:

Proposition: (Triangle inequality)

$$|z + w| \leq |z| + |w|$$
$$\big|\, |z| - |w| \,\big| \leq |z - w|$$

We also note that

$$|z|^2 = z\bar{z} = (x + iy)(x - iy) = x^2 + y^2$$

Most things that we proved about real numbers that did not require the ordering carries over. Also, the standard topology on $\mathbb{C}$ is just the standard topology on $\mathbb{R}^2$, using the modulus for our distance as above.

Hence we can carry over all that we know about the metric space $\mathbb{R}^2$ to $\mathbb{C}$. In particular we know what limits of sequences mean, we know about complex-valued functions and their continuity, we know that $\mathbb{C}$ is a **complete metric space**, etc...

It is also not hard to show that the algebraic operations are continuous. This is because convergence in $\mathbb{R}^2$ is the same as convergence for each component. So for example: let $z_n = x_n + iy_n$ and $w_n = s_n + it_n$ and suppose that $\lim z_n = z = x + iy$ and $\lim w_n = w = s + it$. Let us show that

$$\lim_{n\to\infty} z_n w_n = zw$$

Note that as topology on $\mathbb{C}$ is the same as on $\mathbb{R}^2$, then $x_n \to x$, $y_n \to y$, $s_n \to s$, and $t_n \to t$. Then

$$z_n w_n = (x_n s_n - y_n t_n) + i(x_n t_n + y_n s_n)$$

now $\lim(x_n s_n - y_n t_n) = xs - yt$ and $\lim(x_n t_n + y_n s_n) = xt + ys$ and $(xs - yt) + i(xt + ys) = zw$ so

$$\lim_{n\to\infty} z_n w_n = zw$$

The rest is left to student.

Similarly the modulus, and complex conjugate are continuous functions.

Seq. and ser. of functions

Let us get back to swapping on limits. Let $\{f_n\}$ be a sequence of functions $f_n \colon X \to Y$ for a set X and a metric space Y. Let $f \colon X \to Y$ be a function and for every $x \in X$ suppose that

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

We say the sequence $\{f_n\}$ *converges pointwise* to f.

Similarly if $Y = \mathbb{C}$, we can have a pointwise convergence of a series, for every $x \in X$ we can have

$$g(x) = \lim_{n\to\infty} \sum_{k=1}^{n} f_k(x) = \sum_{k=1}^{\infty} f_k(x).$$

Q: If f_n are all continuous, is f continuous? Differentiable? Integrable? What are the derivatives or integrals of f?

For example for continuity we are asking if

$$\lim_{x\to x_0} \lim_{n\to\infty} f_n(x) \stackrel{?}{=} \lim_{n\to\infty} \lim_{x\to x_0} f_n(x) \tag{1}$$

We don't even know a priory if both sides exist, let alone equal each other.

Example:

$$f_n(x) = \frac{1}{1 + nx^2}$$

converges pointwise to

$$f(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{else} \end{cases}$$

which is not continuous of course.

We've seen continuity is preserved if we require a stronger convergence, that is, we require uniform convergence.

Let $f_n \colon X \to Y$ be functions. Then $\{f_n\}$ *converges uniformly* to f if for every $\epsilon > 0$, there exists an M such that for all $n \geq M$ and all $x \in X$ we have

$$d(f_n(x), f(x)) < \epsilon$$

If we are dealing with complex-valued functions then

$$|f_n(x) - f(x)| < \epsilon$$

Similarly a series of functions converges uniformly if the sequence of partial sums converges uniformly, that is for every $\epsilon > 0$ ... we have

$$\left| \left(\sum_{k=1}^{n} f_k(x) \right) - f(x) \right| < \epsilon$$

Again recall from last semester that this is stronger than pointwise convergence. Pointwise convergence can be stated as: $\{f_n\}$ *converges pointwise* to f if for every $x \in X$ and every $\epsilon > 0$, there exists an M such that for all $n \geq M$ we have

$$d(f_n(x), f(x)) < \epsilon$$

Note that for uniform convergence M does not depend on x. We need to have one M that works for all M.

Theorem 7.8: Let $f_n \colon X \to \mathbb{C}$ be functions. Then $\{f_n\}$ converges uniformly if and only if for every $\epsilon > 0$, there is an M such that for all $n, m \geq M$, and all $x \in X$ we have

$$|f_n(x) - f_m(x)| < \epsilon$$

Proof. Suppose that $\{f_n\}$ converges uniformly to some $f \colon X \to \mathbb{C}$. Then find M such that for all $n \geq M$ we have

$$|f_n(x) - f(x)| < \frac{\epsilon}{2} \qquad \text{for all } x \in X$$

Then for all $m, n \geq M$ we have

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

done.

For the other direction, first fix x. The sequence of complex numbers $\{f_n(x)\}$ is Cauchy and so has a limit $f(x)$.

Given $\epsilon > 0$ find an M such that for all $n, m \geq M$, and all $x \in X$ we have

$$|f_n(x) - f_m(x)| < \epsilon$$

Now we know that $\lim f_m(x) = f(x)$, so as compositions of continuous functions are continuous and algebraic operations and the modulus are continuous we have

$$|f_n(x) - f(x)| \leq \epsilon$$

(the nonstrict inequality is no trouble) $\square$

Sometimes we write for $f \colon X \to \mathbb{C}$

$$\|f\|_u = \sup_{x \in X} |f(x)| .$$

This is the *supremum norm* or *uniform norm*. Then we have that $f_n \colon X \to \mathbb{C}$ converge to f if and only if

$$\lim_{n \to \infty} \|f_n - f\|_u = 0$$

That is if

$$\lim_{n \to \infty} \left(\sup_{x \in X} |f_n(x) - f(x)| \right) = 0$$

(That was Theorem 7.9)

If $f \colon X \to \mathbb{C}$ is continuous and X is compact then $\|f\|_u < \infty$. With the metric $d(f,g) = \|f + g\|_u$ the set of continuous complex-valued functions $C(X, \mathbb{C})$ is a metric space (we have seen this for real-valued functions which is almost the same). Let us check the triangle inequality.

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f(x)\|_u + \|g(x)\|_u .$$

Now take a supremum on the left to get

$$\|f(x) + g(x)\|_u \leq \|f(x)\|_u + \|g(x)\|_u .$$

So we have seen above that convergence in the metric space $C(X, \mathbb{C})$ is uniform convergence.

Theorem 7.10 (Weierstrass M-test) Suppose that $f_n : X \to \mathbb{C}$ are functions,

$$|f_n(x)| \leq M_n$$

and

$$\sum_{n=1}^{\infty} M_n$$

converges. Then

$$\sum_{n=1}^{\infty} f_n(x)$$

converges uniformly. (converse not true by the way)

Proof. Suppose that $\sum M_n$ converges. Given $\epsilon > 0$, we have that the partial sums of $\sum M_n$ are Cauchy so for there is an N such that for all $m, n \geq N$ with $m \geq n$ we have

$$\sum_{k=n+1}^{m} M_k < \epsilon$$

Now let us look at a Cauchy difference of the partial sums of the functions

$$\left| \sum_{k=n+1}^{m} f_k(x) \right| \leq \sum_{k=n+1}^{m} |f_k(x)| \leq \sum_{k=n+1}^{m} M_k < \epsilon.$$

And we are done by Theorem 7.8. □

Examples:

$$\sum_{n=1}^{\infty} \frac{\sin(nx)}{n^2}$$

converges uniformly on $\mathbb{R}$. (Note: this is a Fourier series, we'll see more of these later). That is because

$$\left| \frac{\sin(nx)}{n^2} \right| \leq \frac{1}{n^2}$$

and

$$\sum_{n=1}^{\infty} \frac{1}{n^2}$$

converges.

Another example.

$$\sum_{n=0}^{\infty} \frac{1}{n!} x^n$$

converges uniformly on any bounded interval. For example take the interval $[-r, r] \subset \mathbb{R}$ (any bounded interval is in such an interval)

$$\left| \frac{1}{n!} x^n \right| \leq \frac{r^n}{n!}$$

and

$$\sum_{n=1}^{\infty} \frac{r^n}{n!}$$

converges by the quotient rule

$$\frac{r^{n+1} \ / \ (n+1)!}{r^n \ / \ n!} = \frac{r}{n+1} \to 0 \quad \text{as} \quad n \to \infty$$

Now we would love to say something about the limit. For example, is it continuous?

The following theorem would work with an arbitrary complete metric space rather than just the complex numbers. We use complex numbers for simplicity.

Theorem 7.11: Let X be a metric space and $f_n \colon X \to \mathbb{C}$ be functions. Suppose that $\{f_n\}$ converges uniformly to $f \colon X \to \mathbb{C}$. Let $\{x_k\}$ be a sequence in X and $x = \lim x_k$. Suppose that

$$a_n = \lim_{k\to\infty} f_n(x_k)$$

exists for all n. Then $\{a_n\}$ converges and

$$\lim_{k\to\infty} f(x_k) = \lim_{n\to\infty} a_n$$

In other words

$$\lim_{k\to\infty} \lim_{n\to\infty} f_n(x_k) = \lim_{n\to\infty} \lim_{k\to\infty} f_n(x_k)$$

Proof. First we have to show that $\{a_n\}$ converges. We know that $\{f_n\}$ is uniformly Cauchy (Theorem 7.8). Let $\epsilon > 0$ be given. There is an M such that for all $m, n \geq M$ we have

$$|f_n(x_k) - f_m(x_k)| < \epsilon \qquad \text{for all } k.$$

Letting $k \to \infty$ we obtain that

$$|a_n - a_m| \leq \epsilon.$$

Hence $\{a_n\}$ is Cauchy and hence converges. Let us write

$$a = \lim_{n\to\infty} a_n.$$

Now find a $k \in \mathbb{N}$ such that

$$|f_k(y) - f(y)| < \epsilon/3$$

for all $y \in X$. We can assume that k is large enough so that

$$|a_k - a| < \epsilon/3.$$

We find an $N \in \mathbb{N}$ such that for $m \geq N$ we have

$$|f_k(x_m) - a_k| < \epsilon/3.$$

Then for $m \geq N$ we have

$$|f(x_m) - a| \leq |f(x_m) - f_k(x_m)| + |f_k(x_m) - a_k| + |a_k - a| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.$$

$\square$

Theorem 7.12: Let $f_n \colon X \to \mathbb{C}$ be continuous functions such that $\{f_n\}$ converges uniformly to $f \colon X \to \mathbb{C}$. Then f is continuous.

Proof is immediate application of 7.11. Note that the theorem also holds for an arbitrary target metric space rather than just the complex numbers, although that would have to be proved directly.

Converse is not true. Just because the limit is continuous doesn't mean that the convergence is uniform. For example: $f_n \colon (0,1) \to \mathbb{R}$ defined by $f_n(x) = x^n$ converge to the zero function, but not uniformly.

A combination of 7.8 and 7.12 shows that for a compact X, $C(X, \mathbb{C})$ is a complete metric space. By 7.8 we have that a Cauchy sequence in $C(X, \mathbb{C})$ converges uniformly to a function that is continuous by 7.12. This is Theorem 7.15 in Rudin.

We have seen that the example Fourier series

$$\sum_{n=1}^{\infty} \frac{\sin(nx)}{n^2}$$

converges uniformly and hence is continuous by 7.12. We note that this can be extended to show that if there is a constant C such that if $\alpha > 1$

$$|a_n| \leq \frac{C}{|n|^\alpha}$$

for nonzero $n \in \mathbb{Z}$, then the general Fourier series

$$\sum_{n=-\infty}^{\infty} a_n e^{inx}$$

is a continuous function on $\mathbb{R}$. (Note that $e^{ix} = \cos(x) + i\sin(x)$ to write the series in terms of sines and cosines. The condition on those coefficients is the same). Also note that the standard way to sum a series with two infinite limits is

$$\sum_{n=-\infty}^{\infty} c_n = \left(\lim_{N\to\infty} \sum_{n=1}^{N} c_{-n} \right) + \left(\lim_{N\to\infty} \sum_{n=0}^{N} c_n \right),$$

although for Fourier series we often sum it as

$$\lim_{N\to\infty} \sum_{n=-N}^{N} a_n e^{inx}.$$

When the series converges absolutely it does not matter how we sum it.

Let us see that we can have extra conditions when a converse of 7.12 works.

Theorem 7.13: (Dini's theorem) Suppose that X is compact and $f_n \colon X \to \mathbb{R}$ is a sequence of continuous functions converging pointwise to a continuous $f \colon X \to \mathbb{R}$ and such that

$$f_n(x) \geq f_{n+1}(x).$$

Then $\{f_n\}$ converges to f uniformly.

Proof. Let $g_n = f_n - f$. The g_n are continuous, go to 0 pointwise, and $g_n(x) \geq g_{n+1}(x) \geq 0$. If we show that $\{g_n\}$ converges uniformly to 0, then $\{f_n\}$ converges uniformly.

Let $\epsilon > 0$ be given. Take the set

$$U_n = \{x \in X : g_n(x) < \epsilon\} = g_n^{-1}\big((-\infty, \epsilon)\big). \tag{2}$$

U_n are open (inverse image of open sets by a continuous function). Now for every $x \in X$, since $\{g_n\}$ converges pointwise to 0, there must be some n such that $g_n(x) < \epsilon$ or in other words $x \in U_n$. Therefore, $\{U_n\}$ are an open cover, so there is a finite subcover.

$$X = U_{n_1} \cup U_{n_2} \cup \cdots \cup U_{n_k}$$

for some $n_1 < n_2 < \cdots < n_k$. As $\{g_n\}$ is decreasing we get that $U_n \subset U_{n+1}$ so

$$X = U_{n_1} \cup U_{n_2} \cup \cdots \cup U_{n_k} = U_{n_k}.$$

Write $N = n_k$. Hence $g_N(x) < \epsilon$ for all x. As $\{g_n(x)\}$ is always decreasing we have that for all $n \geq N$ we have for all $x \in X$

$$|g_n(x) - 0| = g_n(x) \leq g_N(x) < \epsilon.$$

So $\{g_n\}$ goes to 0 uniformly. $\square$

Compactness is necessary. For example,

$$f_n(x) = \left| \frac{x}{n} \right|$$

are all continuous on $\mathbb{R}$, monotonically converge to 0 as above, but the convergence is of course not uniform.

If f_n's are not continuous the theorem doesn't hold either. For example, if $f_n \colon [0,1] \to \mathbb{R}$

$$f_n(x) = \begin{cases} x^2 & \text{if } x < 1 \\ 0 & \text{else} \end{cases}$$

then $\{f_n\}$ goes monotonically pointwise to 0, and the domain is compact, but the convergence is not uniform (Exercise: see where the proof breaks).

Finally,

$$f_n(x) = \frac{nx}{1+n^2x^2}$$

are continuous, they go to zero pointwise (but not monotonically). If we take $X = [0,1]$ then the domain is compact, yet the convergence is not uniform since

$$f_n(1/n) = 1/2 \qquad \text{for all } n.$$

Uniform convergence and integration.

Proposition: If $f_n \colon X \to \mathbb{C}$ are bounded functions and converge uniformly to $f \colon X \to \mathbb{C}$, then f is bounded.

Proof. There must exist an n such that

$$|f_n(x) - f(x)| < 1$$

for all x. Now find M such that $|f_n(x)| \leq M$ for all x. By reverse triangle inequality

$$|f(x)| < 1 + |f_n(x)| \leq 1 + M.$$

$\square$

We have seen Riemann integrals of real-valued functions. For complex-valued function $f \colon [a,b] \to \mathbb{C}$, we write $f(x) = u(x) + iv(x)$ where u and v are real-valued. We say f is Riemann integrable if both u and v are, in which case

$$\int_a^b f = \int_a^b u + i \int_a^b v$$

So most statements about real-valued Riemann integrable functions just carry over immediately to complex-valued functions.

We will not do Riemann-Stieltjes integration like Rudin, just Riemann.

Theorem 7.16: Suppose that $f_n \colon [a,b] \to \mathbb{C}$ are Riemann integrable and suppose that $\{f_n\}$ converges uniformly to $f \colon [a,b] \to \mathbb{C}$. Then f is Riemann integrable and

$$\int_a^b f = \lim_{n\to\infty} \int_a^b f_n$$

Proof. Without loss of generality suppose that f_n, and hence f, are all real-valued. As f_n are all bounded, we have that f is bounded by above proposition.

Let $\epsilon > 0$ be given. As f_n goes to f uniformly, we find an $M \in \mathbb{N}$ such that for all $n \geq M$ we have $|f_n(x) - f(x)| < \frac{\epsilon}{2(b-a)}$ for all $x \in [a,b]$. Note that f_n is integrable and compute

$$\begin{aligned}
\overline{\int_a^b} f - \underline{\int_a^b} f &= \overline{\int_a^b} \big(f(x) - f_n(x) + f_n(x)\big)\,dx - \underline{\int_a^b} \big(f(x) - f_n(x) + f_n(x)\big)\,dx \\
&= \overline{\int_a^b} \big(f(x) - f_n(x)\big)\,dx + \overline{\int_a^b} f_n(x)\,dx - \underline{\int_a^b} \big(f(x) - f_n(x)\big)\,dx - \underline{\int_a^b} f_n(x)\,dx \\
&= \overline{\int_a^b} \big(f(x) - f_n(x)\big)\,dx + \int_a^b f_n(x)\,dx - \underline{\int_a^b} \big(f(x) - f_n(x)\big)\,dx - \int_a^b f_n(x)\,dx \\
&= \overline{\int_a^b} \big(f(x) - f_n(x)\big)\,dx - \underline{\int_a^b} \big(f(x) - f_n(x)\big)\,dx \\
&\leq \frac{\epsilon}{2(b-a)}(b-a) + \frac{\epsilon}{2(b-a)}(b-a) = \epsilon.
\end{aligned}$$

The inequality follows from the fact that for all $x \in [a,b]$ we have $\frac{-\epsilon}{2(b-a)} < f(x) - f_n(x) < \frac{\epsilon}{2(b-a)}$. As $\epsilon > 0$ was arbitrary, f is Riemann integrable.

Finally we compute $\int_a^b f$. Again, for $n \geq M$ (where M is the same as above) we have

$$\begin{aligned} \left| \int_a^b f - \int_a^b f_n \right| &= \left| \int_a^b \big(f(x) - f_n(x)\big)\, dx \right| \\ &\leq \frac{\epsilon}{2(b-a)}(b-a) = \frac{\epsilon}{2} < \epsilon. \end{aligned}$$

Therefore $\{\int_a^b f_n\}$ converges to $\int_a^b f$. ☐

Corollary: Suppose that $f_n\colon [a,b] \to \mathbb{C}$ are Riemann integrable and suppose that

$$\sum_{n=1}^{\infty} f_n(x)$$

converges uniformly. Then the series is Riemann integrable on $[a,b]$ and

$$\int_a^b \sum_{n=1}^{\infty} f_n(x)\, dx = \sum_{n=1}^{\infty} \int_a^b f_n(x)\, dx$$

Example: Let us show how to integrate a Fourier series. used for convenience.

$$\int_0^x \sum_{n=1}^{\infty} \frac{\cos(nt)}{n^2}\, dt = \sum_{n=1}^{\infty} \int_0^x \frac{\cos(nt)}{n^2}\, dt = \sum_{n=1}^{\infty} \frac{\sin(nx)}{n^3}$$

The swapping of integral and sum is possible because of uniform convergence, which we have proved before using the M test.

Note that we can swap integrals and limits under far less stringent hypotheses, but for that we would need a stronger integral than Riemann integral. E.g. the Lebesgue integral.

Differentiation

Example: Let $f_n(x) = \frac{1}{1+nx^2}$. We know this converges pointwise to a function that is zero except at the origin. Now note that

$$f_n'(x) = \frac{-2nx}{(1+nx^2)^2}$$

No matter what x is $\lim_{n\to\infty} f_n'(x) = 0$. So the derivatives converge pointwise to 0 (not uniformly on any closed interval containing 0), while the limit of f_n is not even differentiable at 0 (not even continuous).

Example: Let $f_n(x) = \frac{\sin(n^2 x)}{n}$, then $f_n \to 0$ uniformly on $\mathbb{R}$ (easy). However

$$f_n'(x) = n\cos(n^2 x)$$

And for example at $x = 0$, this doesn't even converge. So we need something stronger than uniform convergence of f_n.

For complex-valued functions, again, the derivative of $f\colon [a,b] \to \mathbb{C}$ where $f(x) = u(x) + iv(x)$ is

$$f'(x) = u'(x) + iv'(x).$$

Theorem 7.17: Suppose $f_n\colon [a,b] \to \mathbb{C}$ are differentiable on $[a,b]$. Suppose that there is some point $x_0 \in [a,b]$ such that $\{f_n(x_0)\}$ converges, and such that $\{f_n'\}$ converge uniformly. Then there is a differentiable function f such that $\{f_n\}$ converges uniformly to f, and furthermore

$$f'(x) = \lim_{n\to\infty} f_n'(x) \qquad \text{for all } x \in [a,b].$$

Proof. Let $\epsilon > 0$ be given. We have that $\{f_n(x_0)\}$ is Cauchy so there is an N such that for $n, m \geq N$ we have

$$|f_n(x_0) - f_m(x_0)| < \epsilon/2$$

We also that that $\{f_n'\}$ converge uniformly and hence is uniformly Cauchy, so also assume that for all x we have

$$|f_n'(x) - f_m'(x)| < \frac{\epsilon}{2(b-a)}$$

Now we have the mean value theorem (for complex-valued function just apply the theorem on both real and imaginary parts separately) for the differentiable function $f_n - f_m$, so for x_0, x in $[a, b]$ there is a t between x_0 and x such that

$$(f_n(x_0) - f_m(x_0)) - (f_n(x) - f_m(x)) = (f_n'(t) - f_m'(t))(x_0 - x)$$

and so

$$|(f_n(x_0) - f_m(x_0)) - (f_n(x) - f_m(x))| = |f_n'(t) - f_m'(t)|\ |x_0 - x| < \frac{\epsilon}{2(b-a)}\ |x_0 - x| \leq \frac{\epsilon}{2}$$

and so by inverse triangle inequality

$$|f_n(x) - f_m(x)| \leq |(f_n(x_0) - f_m(x_0)) - (f_n(x) - f_m(x))| + |f_n(x_0) - f_m(x_0)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

This was true for all $x \in [a, b]$ so $\{f_n\}$ is uniformly Cauchy and so the sequence converges uniformly to some $f\colon [a, b] \to \mathbb{C}$.

So what we are still missing is that f is differentiable and that its derivative is the right thing.

Call $g\colon [a, b] \to \mathbb{C}$ the limit of $\{f_n'\}$. Fix $x \in [a, b]$ and take the functions

$$\varphi_n(t) = \begin{cases} \frac{f_n(x)-f_n(t)}{x-t} & \text{if } t \neq x, \\ f_n'(x) & \text{if } t = x, \end{cases}$$

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

The functions $\varphi_n(t)$ are continuous (in particular continuous at x). Given an $\epsilon > 0$ we get that there is some N such that for $m, n \geq N$ we have that $|f_n'(y) - f_m'(y)| < \epsilon$ for all $y \in [a, b]$. Then by the same logic as above using the mean value theorem we get for all t

$$|(f_n(x) - f_n(t)) - (f_m(x) - f_m(t))| = |(f_n(x) - f_m(x)) - (f_n(t) - f_m(t))| < \epsilon\,|x - t|$$

Or in other words

$$|\varphi_n(t) - \varphi_m(t)| < \epsilon \qquad \text{for all } t.$$

So $\{\varphi_n\}$ converges uniformly. The limit is therefore a continuous function. It is easy to see that for a fixed $t \neq x$, $\lim \varphi_n(t) = \varphi(t)$ and since g is the limit of $\{f_n'\}$ also $\lim \varphi_n(x) = \varphi(x)$. Therefore, φ is the limit and is therefore continuous. This means that f is differentiable at x and that $f'(x) = g(x)$. As x was arbitrary we are done. ☐

Note that continuity of f_n' would make the proof far easier and could be done by fundamental theorem of calculus, and the passing of limit under the integral. The trick above is that we have no such assumption.

Example: Let us see what this means for Fourier series

$$\sum_{n=-\infty}^{\infty} a_n e^{inx}$$

and suppose that for some C and some $\alpha > 2$ that for all nonzero $n \in \mathbb{Z}$ we have

$$|a_n| \leq \frac{C}{|n|^\alpha}.$$

Not only does the series converge, but it is also continuously differentiable, and the derivative is obtained by differentiation termwise.

The trick to this is that 1) At $x = 0$ we converge (duh!). 2) If we differentiate the partial sums the resulting partial sums converge absolutely as we are looking at sums such as

$$\sum_{n=m}^{M} i n a_n e^{inx}$$

where we are letting m go to $-\infty$ and M go to ∞. So for the coefficients we have

$$|ina_n| \leq \frac{C}{|n|^{\alpha-1}}$$

and the differentiated series will converge uniformly by the M-test.

Example: Let us construct a continuous nowhere differentiable function. (Theorem 7.18)

Define

$$\varphi(x) = |x| \qquad \text{for } x \in [-1,1]$$

we can extend the definition to all of $\mathbb{R}$ by making φ 2-periodic ($\varphi(x) = \varphi(x+2)$). $\varphi\colon \mathbb{R} \to \mathbb{R}$ is continuous as $|\varphi(x) - \varphi(y)| \leq |x - y|$ (not hard to prove).

As $\sum \left(\frac{3}{4}\right)^n$ converges and $|\varphi(x)| \leq 1$ for all x, we have by the M-test that

$$f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \varphi(4^n x)$$

converges uniformly and hence is continuous.

Fix x and define

$$\delta_m = \pm\frac{1}{2}4^{-m}$$

where the sign is chosen in such a way so that there is no integer between $4^m x$ and $4^m(x + \delta_m)$, which can be done since $4^m |\delta_m| = \frac{1}{2}$.

If $n > m$, then as $4^n \delta_m$ is an even integer. Then as φ is 2-periodic we get that

$$\gamma_n = \frac{\varphi\big(4^n(x+\delta_m)\big) - \varphi(4^n x)}{\delta_m} = 0$$

Furthermore, because there is no integer between $4^m x \pm 1/2$ and $4^m x$ we have that $|\varphi(4^m x \pm 1/2) - \varphi(4^m x)| = |4^m x \pm 1/2) - 4^m x| = 1/2$. Therefore

$$|\gamma_m| = \left|\frac{\varphi(4^m x \pm 1/2) - \varphi(4^m x)}{\pm(1/2)4^{-m}}\right| = 4^m.$$

Similarly, if $n < m$ we, since $|\varphi(s) - \varphi(t)| \leq |s - t|$

$$|\gamma_n| = \left|\frac{\varphi\big(4^n x \pm (1/2)4^{n-m}\big) - \varphi(4^n x)}{\pm(1/2)4^{-m}}\right| \leq \left|\frac{\pm(1/2)4^{n-m}}{\pm(1/2)4^{-m}}\right| = 4^n$$

And so

$$\begin{aligned}
\left|\frac{f(x+\delta_m) - f(x)}{\delta_m}\right| &= \left|\sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \frac{\varphi\big(4^n(x+\delta_m)\big) - \varphi(4^n x)}{\delta_m}\right| = \left|\sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \gamma_n\right| \\
&= \left|\sum_{n=0}^{m}\left(\frac{3}{4}\right)^n \gamma_n\right| \\
&\geq \left|\left(\frac{3}{4}\right)^m \gamma_m\right| - \left|\sum_{n=0}^{m-1}\left(\frac{3}{4}\right)^n \gamma_n\right| \\
&\geq 3^m - \sum_{n=0}^{m-1} 3^n = 3^m - \frac{3^m - 1}{3 - 1} = \frac{3^m + 1}{2}
\end{aligned}$$

It is obvious that $\delta_m \to 0$ as $m \to \infty$, but $\frac{3^m+1}{2}$ goes to infinity. Hence f cannot be differentiable at x.

We will see later that such an f is a uniform limit of not just differentiable functions but actually it is the uniform limit of polynomials.

Equicontinuity

We would like an analogue of Bolzano-Weierstrass, that is every bounded sequence of functions has a convergent subsequence. Matters will not be as simple of course even for continuous functions. Recall from last semester that not every bounded set in the metric space $C(X, \mathbb{R})$ was compact, so a bounded sequence in the metric space $C(X, \mathbb{R})$ may not have a convergent subsequence.

Definition: Let $f_n \colon X \to \mathbb{C}$ be a sequence. We say that $\{f_n\}$ is *pointwise bounded* if for every $x \in X$, there is an $M_x \in \mathbb{R}$ such that

$$|f_n(x)| \leq M_x \qquad \text{for all } n \in \mathbb{N}$$

We say that $\{f_n\}$ is *uniformly bounded* if there is an $M \in \mathbb{R}$ such that

$$|f_n(x)| \leq M \qquad \text{for all } n \in \mathbb{N} \text{ and all } x \in X$$

Note that uniform boundedness is the same as boundedness in the metric space $C(X, \mathbb{C})$ (assuming that X is compact)

We don't have the machinery yet to easily exhibit a sequence of continuous functions on $[0, 1]$ that is uniformly bounded but contains no subsequence converging even pointwise. Let us just state that $f_n(x) = \sin(2\pi nx)$ is one such sequence.

We also of course have that $f_n(x) = x^n$ is a sequence that is uniformly bounded, but contains no sequence that converges uniformly (it does converge pointwise though).

When the domain is countable, matters are easier. We will also use the following theorem as a lemma later. It will use a very common and useful diagonal argument.

Theorem 7.23: Let X be countable and $f_n \colon X \to \mathbb{C}$ give a pointwise bounded sequence of functions, then $\{f_n\}$ has a subsequence that converges pointwise.

Proof. Let $\{x_k\}$ be an enumeration of the elements of X. The sequence $\{f_n(x_1)\}_{n=1}^{\infty}$ is bounded and hence we have a subsequence which we will denote by $f_{1,k}$ such that $\{f_{1,k}(x_1)\}$ converges. Next $\{f_{1,k}(x_2)\}$ has a subsequence that converges and we denote that subsequence by $\{f_{2,k}(x_2)\}$. In general we will have a sequence $\{f_{m,k}\}_k$ that makes $\{f_{m,k}(x_j)\}_k$ converge for all $j \leq m$ and we will let $\{f_{m+1,k}\}_k$ be the subsequence of $\{f_{m,k}\}_k$ such that $\{f_{m+1,k}(x_{m+1})\}_k$ converges (and hence it converges for all x_j for $j = 1, \ldots, m+1$) and we can rinse and repeat.

Finally pick the sequence $\{f_{k,k}\}$. This is a subsequence of the original sequence $\{f_n\}$ of course. Also for any m, except for the first m terms $\{f_k, k\}$ is a subsequence of $\{f_{m,k}\}_k$ and hence for any m the sequence $\{f_{k,k}(x_m)\}_k$ converges. □

For larger than countable sets, we need the functions of the sequence to be related. We look at continuous functions, and the concept we need is equicontinuity.

Definition: a family $\mathcal{F}$ of functions $f \colon X \to \mathbb{C}$ is said to be equicontinuous on X if for every $\epsilon > 0$, there is a $\delta > 0$ such that if $x, y \in X$ with $d(x, y) < \delta$ we have

$$|f(x) - f(y)| < \epsilon \qquad \text{for all } f \in \mathcal{F}.$$

One obvious fact is that if $\mathcal{F}$ is equicontinuous, then every element of f is uniformly continuous. Also obvious is that any finite set of uniformly continuous functions is an equicontinuous family. Of course the interesting case is when $\mathcal{F}$ is an infinite family such as a sequence of functions. Equicontinuity of a sequence is closely related to uniform convergence.

First a proposition, that we have proved for compact intervals before.

Proposition: If X is compact and $f \colon X \to \mathbb{C}$ is continuous, then f is uniformly continuous.

Proof. Let $\epsilon > 0$ be given. If we take all $z \in \mathbb{C}$, the open sets of the form $f^{-1}\big(B(z, \epsilon/2)\big)$ cover X. By the Lebesgue covering lemma as X is compact, there exists a $\delta > 0$ such that for any $x \in X$, $B(x, \delta)$ lies in some member of the cover, in other words there is some $z \in \mathbb{C}$ such that

$$B(x, \delta) \subset f^{-1}\big(B(z, \epsilon/2)\big)$$

Now since $f(x) \in B(z, \epsilon/2)$, then by triangle inequality we have $B(z, \epsilon/2) \subset B\big(f(x), \epsilon\big)$ and so

$$B(x, \delta) \subset f^{-1}\big(B(f(x), \epsilon)\big)$$

But that is precisely what it means for f to be uniformly continuous. $\square$

Theorem 7.24: If X is compact, $f_n \in C(X, \mathbb{C})$, and $\{f_n\}$ converges uniformly, then $\{f_n\}$ is equicontinuous.

Proof. Let $\epsilon > 0$ be given. As f_n converge uniformly, there is an integer N such that for all $n \geq N$ we have

$$|f_n(x) - f_N(x)| < \epsilon/3 \qquad \text{for all } x \in X.$$

Now $\{f_1, f_2, \ldots, f_N\}$ is a finite set of uniformly continuous functions and so as we mentioned above an equicontinuous family. Hence there is a $\delta > 0$ such that

$$|f_j(x) - f_j(y)| < \epsilon/3 < \epsilon$$

whenever $d(x, y) < \delta$ and $1 \leq j \leq N$.

Now take $n > N$. Then for $d(x, y) < \delta$ we have

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.$$

$\square$

Proposition: A compact metric space X contains a countable dense subset.

Proof. For each $n \in \mathbb{N}$ we have that there are finitely many balls of radius $1/n$ that cover X (as X is compact). That is, for every n, there exists a finite set of points $x_{n,1}, x_{n,2}, \ldots, x_{n,k_n}$ such that

$$X = \bigcup_{j=1}^{k_n} B(x_{n,j}, 1/n)$$

So consider the set S of all the points $x_{n,j}$. As S is a countable union of finite sets and therefore countable. For every $x \in X$ and every $\epsilon > 0$, there exists an n such that $1/n < \epsilon$ and an $x_{n,j} \in S$ such that

$$x \in B(x_{n,j}, 1/n) \subset B(x_{n,j}, \epsilon).$$

Hence $x \in \overline{S}$, so $\overline{S} = X$ and S is dense. $\square$

We can now prove the very useful Arzelà-Ascoli theorem about existence of convergent subsequences.

Theorem 7.25: (Arzelà-Ascoli) Let X be compact, $f_n \in C(X, \mathbb{C})$, and let $\{f_n\}$ be pointwise bounded and equicontinuous. Then $\{f_n\}$ is uniformly bounded and $\{f_n\}$ contains a uniformly convergent subsequence.

Basically, an equicontinuous sequence in the metric space $C(X, \mathbb{C})$ that is pointwise bounded is bounded (in $C(X, \mathbb{C})$) and furthermore contains a convergent subsequence in $C(X, \mathbb{C})$.

Proof. Let us first show that the sequence is uniformly bounded.

By equicontinuity we have that there is a $\delta > 0$ such that for all $x \in X$

$$B(x,\delta) \subset f_n^{-1}\big(B(f_n(x),1)\big)$$

Now X is compact, so there exists $x_1, x_2, \ldots, x_k$ such that

$$X = \bigcup_{j=1}^{k} B(x_j,\delta)$$

As $\{f_n\}$ is pointwise bounded there exist $M_1, \ldots, M_k$ such that for $j = 1, \ldots, k$ we have

$$|f_n(x_j)| \leq M_j$$

for all n. Let $M = 1 + \max\{M_1, \ldots, M_k\}$. Now given any $x \in X$, there is a j such that $x \in B(x_j,\delta)$. Therefore, for all n we have $x \in f_n^{-1}(B(f_n(x_j),1))$ or in other words

$$|f_n(x) - f_n(x_j)| < 1$$

By reverse triangle inequality,

$$|f_n(x)| < 1 + |f_n(x_j)| \leq 1 + M_j \leq M$$

And as x was arbitrary, $\{f_n\}$ is uniformly bounded.

Next, pick a countable dense set S. By Theorem 7.23, we find a subsequence $\{f_{n_j}\}$ that converges pointwise on S. Write $g_j = f_{n_j}$ for simplicity. Note that $\{g_n\}$ is equicontinuous.

Let $\epsilon > 0$ be given, then pick $\delta > 0$ such that for all $x \in X$

$$B(x,\delta) \subset g_n^{-1}\big(B(g_n(x), \epsilon/3)\big)$$

By density of S, every $x \in X$ is in some $B(y,\delta)$ for some $y \in S$, and by compactness of X, there is a finite subset $\{x_1, \ldots, x_k\}$ of S such that

$$X = \bigcup_{j=1}^{k} B(x_j,\delta).$$

Now as there are finitely many points and we know that $\{g_n\}$ converges pointwise on S, there exists a single N such that for all $n, m \geq N$ we have for all $j = 1, \ldots, k$

$$|g_n(x_j) - g_m(x_j)| < \epsilon/3.$$

Let $x \in X$ be arbitrary. There is some j such that $x \in B(x_j,\delta)$ and so we have for all $i \in \mathbb{N}$

$$|g_i(x) - g_i(x_j)| < \epsilon/3$$

and so $n, m \geq N$ that

$$|g_n(x) - g_m(x)| \leq |g_n(x) - g_n(x_j)| + |g_n(x_j) - g_m(x_j)| + |g_m(x_j) - g_m(x)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.$$

□

Corollary: Let X be a compact metric space. Let $S \subset C(X,\mathbb{C})$ be a closed, bounded and equicontinuous set. Then S is compact.

The theorem says that S is sequentially compact and we know that means compact in a metric space.

Last semester we have seen that in $C([0,1],\mathbb{C})$ (well actually we had real valued functions but the idea is exactly the same), the closed unit ball $C(0,1)$ in $C([0,1],\mathbb{C})$ was not compact. Hence it cannot be an equicontinuous set.

Corollary: Suppose that $\{f_n\}$ is a sequence of differentiable functions on $[a,b]$, $\{f_n'\}$ is uniformly bounded, and there is an $x_0 \in [a,b]$ such that $\{f_n(x_0)\}$ is bounded. Then there exists a uniformly convergent subsequence $\{f_{n_j}\}$.

Proof. The trick is to use the mean value theorem. If M is the uniform bound on $\{f_n'\}$, then we have by the mean value theorem

$$|f_n(x) - f_n(y)| \leq M\,|x-y|$$

So all the f_n's are Lipschitz with the same constant and hence equicontinuous.

Now suppose that $|f_n(x_0)| \leq M_0$ for all n. By inverse triangle inequality we have for all x

$$|f_n(x)| \leq |f_n(x_0)| + |f_n(x) - f_n(x_0)| \leq M_0 + M\,|x-x_0| \leq M_0 + M(b-a)$$

so $\{f_n\}$ is uniformly bounded.

We can now apply Arzelà-Ascoli to find the subsequence. □

A consequence of the above corollary and the Fundamental Theorem of Calculus is that given some fixed $g \in C([0,1],\mathbb{C})$, the set of functions

$$\{F \in C([0,1],\mathbb{C}) : F(x) = \int_0^x g(t)f(t)\,dt,\ f \in C([0,1],\mathbb{C}),\ \|f\|_u \leq 1\}$$

has compact closure (*relatively compact*). That is, the *operator* $T\colon C([0,1],\mathbb{C}) \to C([0,1],\mathbb{C})$ given by

$$T\big(f\big)(x) = F(x) = \int_0^x g(t)f(t)\,dt$$

takes the unit ball centered at 0 in $C([0,1],\mathbb{C})$ into a relatively compact set. We often say that this means that the operator is compact, and such operators are very important (and very useful).

Stone-Weierstrass

Perhaps surprisingly, even a very badly behaving continuous function is really just a uniform limit of polynomials. We cannot really get any "nicer" as a function than a polynomial.

Theorem 7.26 (Weierstrass approximation theorem) If $f\colon [a,b] \to \mathbb{C}$ is continuous, then there exists a sequence $\{p_n\}$ of polynomials converging to f uniformly on $[a,b]$. Furthermore, if f is real-valued, we can find real-valued p_n.

Proof. For $x \in [0,1]$ define

$$g(x) = f\big((b-a)x+a\big) - f(a) - x\big(f(b)-f(a)\big).$$

If we can prove the theorem for g and find the $\{p_n\}$ for g, we can prove it for f since we simply composed with an invertible affine function and added an affine function to f, so we can easily reverse the process and apply that to our p_n, to obtain polynomials approximating f.

So g is defined on $[0,1]$ and $g(0) = g(1) = 0$. We can now assume that g is defined on the whole real line for simplicity by defining $g(x) = 0$ if $x < 0$ or $x > 1$.

Let

$$c_n = \left(\int_{-1}^1 (1-x^2)^n\,dx\right)^{-1}$$

Then define

$$q_n(x) = c_n(1-x^2)^n$$

so that $\int_{-1}^1 q_n(x)\,dx = 1$.

We note that q_n are peaks around 0 (ignoring what happens outside of $[-1,1]$) that get narrower and narrower. You should plot a few of these to see what is happening. A classic approximation idea is to do a *convolution* integral with peaks like this. That is we will write for $x \in [0,1]$,

$$p_n(x) = \int_0^1 g(t)q_n(t-x)\,dt \quad \left(= \int_{-\infty}^{\infty} g(t)q_n(t-x)\,dt\right).$$

Because q_n is a polynomial we can write

$$q_n(t-x) = a_0(t) + a_1(t)x + \cdots + a_{2n}(t)x^{2n}$$

for some functions $a_k(t)$. Since the integral is in terms of t, the xs just pop out by linearity of the integral, and so we obtain that p_n is a polynomial in x. Finally if $g(t)$ is real-valued then the functions $g(t)a_j(t)$ are real valued and hence p_n has real coefficients (so the "furthermore" part of the theorem will hold).

As q_n is a peak, the integral only sees the values of g that are very close to x and it does a sort of average of them. When the peak gets narrower, we do this average closer to x and hence we expect to get close to the value of $g(x)$. Really we are approximating a δ-function (if you have heard this concept). We could really do this with any polynomial that looks like a little peak near zero. This just happens to be the simplest one. We only need this behavior on $[-1,1]$ as the convolution sees nothing further than this as g is zero outside $[0,1]$.

We still need to prove that this works. First we need some handle on the size of c_n. Note that for $x \in [0,1]$

$$(1-x^2)^n \geq (1-nx^2).$$

This fact can be proved with a tiny bit of calculus as the two expressions are equal at $x = 0$ and $\frac{d}{dx}\left((1-x^2)^n - (1-nx^2)\right) = -2nx(1-x^2)^{n-1} + 2nx$, which is nonnegative for $x \in [0,1]$. Furthermore $1 - nx^2 \geq 0$ for $x \in [0, 1/\sqrt{n}]$.

$$\begin{aligned} c_n^{-1} &= \int_{-1}^{1} (1-x^2)^n \, dx \\ &= 2\int_0^1 (1-x^2)^n \, dx \\ &\geq 2\int_0^{1/\sqrt{n}} (1-x^2)^n \, dx \\ &\geq 2\int_0^{1/\sqrt{n}} (1-nx^2) \, dx \\ &= \frac{4}{3\sqrt{n}} > \frac{1}{\sqrt{n}}. \end{aligned}$$

so $c_n < \sqrt{n}$.

Let's see how small g is if we ignore some small bit around the origin, which is where the peak is. Given any $\delta > 0$, $\delta < 1$, we have for $\delta \leq |x| \leq 1$ that

$$q_n(x) \leq \sqrt{n}(1-\delta^2)^n.$$

Then it is easy to see (e.g. by the ratio test) that $\sqrt{n}(1-\delta^2)^n$ goes to 0 as n goes to infinity. For $x \in [0,1]$ we note

$$p_n(x) = \int_0^1 g(t)q_n(t-x)\,dt = \int_{-x}^{1-x} g(t+x)q_n(t)\,dt = \int_{-1}^{1} g(t+x)q_n(t)\,dt$$

the second equality follows by assuming that g is zero outside of $[0,1]$.

Let $\epsilon > 0$ be given. As $[0,1]$ is compact, we have that g is uniformly continuous. Pick $\delta > 0$ such that if $|x-y| < \delta$ then

$$|g(x) - g(y)| < \frac{\epsilon}{2}.$$

This actually works for all $x, y \in \mathbb{R}$ because g is just zero outside of $[0,1]$

Let M be such that $|g(x)| \leq M$ for all x. Furthermore pick N such that for all $n \geq N$ we have

$$4M\sqrt{n}(1-\delta^2)^n < \frac{\epsilon}{2}.$$

Note that $\int_{-1}^{1} q_n(t)\,dt = 1$ and $q_n(x) \geq 0$ on $[-1,1]$ so

$$\begin{aligned}
|p_n(x) - g(x)| &= \left| \int_{-1}^{1} g(t+x) q_n(t)\,dt - g(x) \int_{-1}^{1} q_n(t)\,dt \right| \\
&= \left| \int_{-1}^{1} \big(g(t+x) - g(x)\big) q_n(t)\,dt \right| \\
&\leq \int_{-1}^{1} |g(t+x) - g(x)|\, q_n(t)\,dt \\
&= \int_{-1}^{-\delta} |g(t+x) - g(x)|\, q_n(t)\,dt + \int_{-\delta}^{\delta} |g(t+x) - g(x)|\, q_n(t)\,dt + \int_{\delta}^{1} |g(t+x) - g(x)|\, q_n(t)\,dt \\
&\leq 2M \int_{-1}^{-\delta} q_n(t)\,dt + \frac{\epsilon}{2} \int_{-\delta}^{\delta} q_n(t)\,dt + 2M \int_{\delta}^{1} q_n(t)\,dt \\
&\leq 2M\sqrt{n}(1-\delta^2)^n(1-\delta) + \frac{\epsilon}{2} + 2M\sqrt{n}(1-\delta^2)^n(1-\delta) \\
&< 4M\sqrt{n}(1-\delta^2)^n + \frac{\epsilon}{2} < \epsilon.
\end{aligned}$$

□

Think about the consequences of the theorem. If you have any property that gets preserved under uniform convergence and it is true for polynomials, then it must be true for all continuous functions.

Let us note an immediate application of the Weierstrass theorem. We have already seen that countable dense subsets can be very useful.

Corollary: The metric space $C([a,b],\mathbb{C})$ contains a countable dense subset.

Proof. Without loss of generality suppose that we are dealing with $C([a,b],\mathbb{R})$ (why can we?). The real polynomials are dense in $C([a,b],\mathbb{R})$. If we can show that any real polynomial can be approximated by polynomials with rational coefficients, we are done. This is because there are only countably many rational numbers and so there are only countably many polynomials with rational coefficients (a countable union of countable sets is still countable).

Further without loss of generality, suppose that $[a,b] = [0,1]$. Let

$$p(x) = \sum_{k=0}^{n} a_k x^k$$

be a polynomial of degree n where $a_k \in \mathbb{R}$. Given $\epsilon > 0$, pick $b_k \in \mathbb{Q}$ such that $|a_k - b_k| < \frac{\epsilon}{n+1}$. Then if we let

$$q(x) = \sum_{k=0}^{n} b_k x^k$$

we have

$$|p(x) - q(x)| = \left| \sum_{k=0}^{n} (a_k - b_k) x^k \right| \leq \sum_{k=0}^{n} |a_k - b_k|\, x^k \leq \sum_{k=0}^{n} |a_k - b_k| < \sum_{k=0}^{n} \frac{\epsilon}{n+1} = \epsilon$$

□

Funky remark: While we will not prove this, we note that the above corollary implies that $C([a,b],\mathbb{C})$ has the same cardinality as the real numbers, which may be a bit surprising. The set of all functions $[a,b] \to \mathbb{C}$ has cardinality that is strictly greater than the cardinality of $\mathbb{R}$, it has the cardinality of the power set of $\mathbb{R}$. So the set of continuous functions is a very tiny subset of the set of all functions.

Next thing we do is that we will want to abstract away what is not really necessary and prove a general version of this theorem. We have shown that the polynomials are dense in the space of continuous functions on a compact interval. So what kind of families of functions are also dense? Furthermore, what if we let the domain be an arbitrary metric space, then we no longer have polynomials.

The resulting theorem is the Stone-Weierstrass theorem. We will need a special case of the Weierstrass theorem though.

Corollary: Let $[-a, a]$ be an interval. Then there is a sequence of real polynomials $\{p_n\}$ that converges uniformly to $|x|$ on $[-a, a]$ and such that $p_n(0) = 0$ for all n.

Proof. As $f(x) = |x|$ is continuous and real-valued on $[-a, a]$ we definitely have some real polynomials $\tilde{p}_n$ that converge to f. Let

$$p_n(x) = \tilde{p}_n(x) - \tilde{p}_n(0)$$

Obviously $p_n(0) = 0$.

We know $\lim \tilde{p}_n(0) = 0$. Given $\epsilon > 0$, let N be such that for $n \geq N$ we have $|\tilde{p}_n(0)| < \epsilon/2$ and also that $\big|\tilde{p}_n(x) - |x|\,\big| < \epsilon/2$. Now

$$\big|p_n(x) - |x|\big| = \big|\tilde{p}_n(x) - \tilde{p}_n(0) - |x|\big| \leq \big|\tilde{p}_n(x) - |x|\big| + |\tilde{p}_n(0)| < \epsilon/2 + \epsilon/2 = \epsilon.$$

$\square$

Following the proof of the theorem, we see that we can always make the polynomials from the Weierstrass theorem have a fixed value at one point, so it works not just for $|x|$.

Definition: A family $\mathcal{A}$ of complex-valued functions $f \colon X \to \mathbb{C}$ is said to be an *algebra* (sometimes complex algebra or algebra over $\mathbb{C}$) if for all $f, g \in \mathcal{A}$ and $c \in \mathbb{C}$ we have

(i) $f + g \in \mathcal{A}$,
(ii) $fg \in \mathcal{A}$, and
(iii) $cg \in \mathcal{A}$.

If we talk of an algebra of real-valued functions (sometimes real algebra or algebra over $\mathbb{R}$), then of course we only need the above to hold for $c \in \mathbb{R}$.

We say that $\mathcal{A}$ is *uniformly closed* if the limit of every uniformly convergent sequence in $\mathcal{A}$ is also in $\mathcal{A}$. Similarly, a set $\mathcal{B}$ of all uniform limits of uniformly convergent sequences in $\mathcal{A}$ is said to be the *uniform closure* of $\mathcal{A}$.

If X is a compact metric space and $\mathcal{A}$ is a subset of the metric space $C(X, \mathbb{C})$, then the uniform closure of $\mathcal{A}$ is just the metric space closure $\overline{\mathcal{A}}$ in $C(X, \mathbb{C})$, of course.

For example, the set $\mathcal{P}$ of all polynomials is an algebra in $C(X, \mathbb{C})$, and we have shown that its uniform closure is all of $C(X, \mathbb{C})$.

Theorem 7.29: Let $\mathcal{A}$ be an algebra of bounded functions on a set X, and let $\mathcal{B}$ be its uniform closure. Then $\mathcal{B}$ is a uniformly closed algebra.

Proof. Let $f, g \in \mathcal{B}$ and $c \in \mathbb{C}$. Then there are uniformly convergent sequences of functions $\{f_n\}$, $\{g_n\}$ in $\mathcal{A}$ such that f is the uniform limit of $\{f_n\}$ and g is the uniform limit of $\{g_n\}$.

What we want is to show that $f_n + g_n$ converges uniformly to $f + g$, $f_n g_n$ converges uniformly to fg, and cf_n converges uniformly to cf. As f and g are bounded, then we have seen that $\{f_n\}$ and $\{g_n\}$ are also uniformly bounded. Therefore there is a single number M such that for all $x \in X$ we have

$$|f_n(x)| \leq M, \qquad |g_n(x)| \leq M, \qquad |f(x)| \leq M, \qquad |g(x)| \leq M.$$

The following estimates are enough to show uniform convergence.

$$\big|\big(f_n(x) + g_n(x)\big) - \big(f(x) + g(x)\big)\big| \leq |f_n(x) - f(x)| + |g_n(x) - g(x)|$$

$$\begin{aligned} |f_n(x)g_n(x) - f(x)g(x)| &\leq |f_n(x)g_n(x) - f_n(x)g(x)| + |f_n(x)g(x) - f(x)g(x)| \\ &\leq M\,|g_n(x) - g(x)| + M\,|f_n(x) - f(x)| \end{aligned}$$

$$|cf_n(x) - cf(x)| = |c|\;|f_n(x) - f(x)|$$

Hence $f+g \in \mathcal{B}$, $fg \in \mathcal{B}$ and $cf \in \mathcal{B}$.

Next we want to show that $\mathcal{B}$ is uniformly closed. Suppose that $\{f_n\}$ is a sequence in $\mathcal{B}$ converging uniformly to some function $f\colon X \to \mathbb{C}$. As $\mathcal{B}$ is the closure of $\mathcal{A}$ we find a $g_n \in \mathcal{A}$ for every n such that for all $x \in X$ we have

$$|f_n(x) - g_n(x)| < {}^1\!/_n$$

So given $\epsilon > 0$ find N such that for all $n \geq N$ we have $|f_n(x) - f(x)| < {}^\epsilon\!/_2$ and also such that ${}^1\!/_n < {}^\epsilon\!/_2$. Then

$$|g_n(x) - f(x)| \leq |g_n(x) - f_n(x)| + |f_n(x) - f(x)| < {}^1\!/_n + {}^\epsilon\!/_2 < \epsilon$$

Or if we had shown that the set of bounded functions on X is a metric space, then the last assertion would follow directly as in a metric space closure of a closed set is the set itself. (Rudin does this). $\square$

Now let us distill the right properties of polynomials that were sufficient for an approximation theorem.

Definition: Let $\mathcal{A}$ be a family of complex-valued functions defined on a set X.

We say $\mathcal{A}$ *separates points* if for every $x, y \in X$, with $x \neq y$ there is a function $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.

We say $\mathcal{A}$ *vanishes at no point* if for every $x \in X$ there is an $f \in \mathcal{A}$ such that $f(x) \neq 0$.

Examples: The set $\mathcal{P}$ of polynomials separates points and vanishes at no point on $\mathbb{R}$. That is, $1 \in \mathcal{P}$ so it vanishes at no point. And for $x, y \in \mathbb{R}$, $x \neq y$, just take $f(t) = (t - x)$, then $f(x) = 0$ and $f(y) = y - x \neq 0$.

The set of functions of the form

$$f(t) = C + \sum_{n=1}^{k} \cos(nt)$$

does not separate points if we let the domain be any interval of the form $[-a, a]$. That's because $f(-x) = f(x)$ for all x.

The set of polynomials with no constant term vanishes at the origin.

Theorem 7.31: Suppose that $\mathcal{A}$ is an algebra of functions on X, that separates points and vanishes at no point. Suppose x, y are distinct points of X and $c, d \in \mathbb{C}$. Then there is an $f \in \mathcal{A}$ such that

$$f(x) = c, \qquad f(y) = d.$$

If $\mathcal{A}$ is a real algebra, the theorem holds as well when c, d are real.

Proof. There must exist an $g, h, k \in \mathcal{A}$ such that

$$g(x) \neq g(y), \quad h(x) \neq 0, \quad k(y) \neq 0.$$

write

$$f = c\frac{\big(g - g(y)\big)h}{\big(g(x) - g(y)\big)h(x)} + d\frac{\big(g - g(x)\big)k}{\big(g(y) - g(x)\big)k(y)} = c\frac{gh - g(y)h}{g(x)h(x) - g(y)h(x)} + d\frac{gk - g(x)k}{g(y)k(y) - g(x)k(y)}$$

Do note that we are not dividing by zero (clear from the first formula). Also from the first formula we see that $f(x) = c$ and $f(y) = d$. By the second formula we see that $f \in \mathcal{A}$ (as $\mathcal{A}$ is an algebra). $\square$

Theorem 7.32: (Stone-Weierstrass, real version) Let X be a compact metric space and $\mathcal{A}$ an algebra of real-valued continuous functions on X, such that $\mathcal{A}$ separates points and vanishes at no point. Then the uniform closure of $\mathcal{A}$ is all of $C(X, \mathbb{R})$.

The proof is divided into several claims. Write $\mathcal{B}$ for the uniform closure of $\mathcal{A}$. We wish to show that $\mathcal{B} = C(X, \mathbb{R})$.

Claim 1: If $f \in \mathcal{B}$ then $|f| \in \mathcal{B}$.

Proof. f is bounded (continuous on a compact set) for example $|f(x)| \leq M$ for all $x \in X$.

Let $\epsilon > 0$ be given. By the corollary to Weierstrass theorem there exists a real polynomial $c_1 y + c_2 y^2 + \cdots + c_n y^n$ (vanishing at $y = 0$) such that

$$\left| |y| - \sum_{j=1}^{N} c_j y^j \right| < \epsilon$$

for all $y \in [-M, M]$. Because $\mathcal{A}$ is an algebra (note there is no constant term in the polynomial) we have that

$$g = \sum_{j=1}^{N} c_j f^j \in \mathcal{A}.$$

As $|f(x)| \leq M$ we have that

$$\big| |f(x)| - g(x) \big| = \left| |f(x)| - \sum_{j=1}^{N} c_j (f(x))^j \right| < \epsilon.$$

So $|f|$ is in the closure of $\mathcal{B}$, which is closed, so $|f| \in \mathcal{B}$. □

Claim 2: If $f \in \mathcal{B}$ and $g \in \mathcal{B}$ then $\max(f, g) \in \mathcal{B}$ and $\min(f, g) \in \mathcal{B}$, where

$$\big(\max(f,g)\big)(x) = \max\{f(x), g(x)\} \qquad \big(\min(f,g)\big)(x) = \min\{f(x), g(x)\}$$

Proof. Write:

$$\max(f, g) = \frac{f+g}{2} + \frac{|f-g|}{2},$$

and

$$\min(f, g) = \frac{f+g}{2} - \frac{|f-g|}{2}.$$

As $\mathcal{B}$ is an algebra we are done. □

The claim is of course true for the minimum or maximum of any finite collection of functions as well.

Claim 3: Given $f \in C(X, \mathbb{R})$, $x \in X$ and $\epsilon > 0$ there exists a $g_x \in \mathcal{B}$ with $g_x(x) = f(x)$ and

$$g_x(t) > f(t) - \epsilon \qquad \text{for all } t \in X.$$

Proof. Let $x \in X$ be fixed. By Theorem 7.31, for every $y \in X$ we find an $h_y \in \mathcal{A}$ such that

$$h_y(x) = f(x), \qquad h_y(y) = f(y)$$

As h_y and f are continuous, the set

$$J_y = \{t \in X : h_y(t) > f(t) - \epsilon\}$$

is open (it is the inverse image of an open set by a continuous function). Furthermore $y \in J_y$. So the sets J_y cover X.

Now X is compact so there exist finitely many points $y_1, \ldots, y_n$ such that

$$X = \bigcup_{j=1}^{n} J_{y_j}.$$

Let

$$g_x = \max(h_{y_1}, h_{y_2}, \ldots, h_{y_n})$$

By Claim 2, g_x is in $\mathcal{B}$. It is easy to see that

$$g_x(t) > f(t) - \epsilon$$

for all $t \in X$, since for any t there was at least one h_{y_j} for which this was true.

Furthermore $h_y(x) = f(x)$ for all $y \in X$, so $g_x(x) = f(x)$. □

Claim 4: If $f \in C(X, \mathbb{R})$ and $\epsilon > 0$ is given then there exists an $h \in \mathcal{B}$ such that

$$|f(x) - h(x)| < \epsilon$$

Proof. For any x find function g_x as in Claim 3.

Let

$$V_x = \{t \in X : g_x(t) < f(t) + \epsilon\}.$$

The sets V_x are open as g_x and f are continuous. Furthermore as $g_x(x) = f(x)$, $x \in V_x$, and so the sets V_x cover X. Thus there are finitely many points $x_1, \ldots, x_n$ such that

$$X = \bigcup_{j=1}^{n} V_{x_j}.$$

Now let

$$h = \min(g_{x_1}, \ldots, g_{x_n}).$$

we see that $h \in \mathcal{B}$ by Claim 2. Similarly as before (same argument as in Step 3) we have that for all $t \in X$

$$h(t) < f(t) + \epsilon.$$

Since all the g_x satisfy $g_x(t) > f(t) - \epsilon$ for all $t \in X$, so does h. Therefore, for all t

$$-\epsilon < h(t) - f(t) < \epsilon,$$

which is the desired conclusion. □

The conclusion follows from Claim 4. The claim states that an arbitrary continuous function is in the closure of $\mathcal{B}$ which itself is closed. Claim 4 is the conclusion of the theorem. So the theorem is proved.

Example: The functions of the form

$$f(t) = \sum_{j=1}^{n} c_j e^{jt},$$

for $c_j \in \mathbb{R}$, are dense in $C([a, b], \mathbb{R})$. We need to note that such functions are a real algebra, which follows from $e^{jt} e^{kt} = e^{(j+k)t}$. They separate points as e^t is one-to-one, and $e^t > 0$ for all t so the algebra does not vanish at any point (we will prove all these properties of exponential in the next chapter).

In general if we have a family of functions that separates points and does not vanish at any point, we can let these function *generate* an algebra by considering all the linear combinations of arbitrary multiples of such functions. That is we basically consider all real polynomials of such functions (where the polynomials have no constant term). For example above, the algebra is generated by e^t, we simply consider polynomials in e^t.

Warning! You could similarly show that the set of all functions of the form

$$\frac{a_0}{2} + \sum_{n=1}^{N} a_n \cos(nt)$$

is an algebra (you would have to use some trig identities to show this). When considered on for example $[0, \pi]$, the algebra separates points and vanishes nowhere so Stone-Weierstrass applies. You *do not* want to conclude from this that every continuous function on $[0, \pi]$ has a uniformly convergent Fourier cosine series. That is not true. In fact there exist continuous functions whose Fourier series does not converge even pointwise. We do have formulas for the coefficients, but those coefficients will not give us a convergent series. The trick is that the sequence of functions in the algebra that we obtain by Stone-Weierstrass is not necessarily the sequence of partial sums of a series.

Same warning applies to polynomials as well. Just because a function is a uniform limit of polynomials doesn't mean that it is a uniform limit of a power series (which we will see in the next chapter).

If we wish to have Stone-Weierstrass to hold for complex algebras, we must make an extra assumption.

Definition: An algebra $\mathcal{A}$ is self-adjoint, if for all $f \in \mathcal{A}$, the function $\bar{f}$ defined by $\bar{f}(x) = \overline{f(x)}$ is in $\mathcal{A}$, where by the bar we mean the complex conjugate.

Theorem 7.32: (Stone-Weierstrass, complex version) Let X be a compact metric space and $\mathcal{A}$ an algebra of complex-valued continuous functions on X, such that $\mathcal{A}$ separates points, vanishes at no point, and is self-adjoint. Then the uniform closure of $\mathcal{A}$ is all of $C(X, \mathbb{C})$.

Proof. Suppose that $\mathcal{A}_{\mathbb{R}} \subset \mathcal{A}$ is the set of the real-valued elements of $\mathcal{A}$. If $f = u + iv$ where u and v are real-valued, then we note that

$$u = \frac{f + \bar{f}}{2}, \qquad v = \frac{f - \bar{f}}{2i}.$$

So $u, v \in \mathcal{A}$ as $\mathcal{A}$ is a self-adjoint algebra, and since they are real-valued we get that $u, v \in \mathcal{A}_{\mathbb{R}}$.

If $x \neq y$, then find an $f \in \mathcal{A}$ such that $f(x) \neq f(y)$. If $f = u + iv$, then it is obvious that either $u(x) \neq u(y)$ or $v(x) \neq v(y)$. So $\mathcal{A}_{\mathbb{R}}$ separates points.

Similarly, for any x find $f \in \mathcal{A}$ such that $f(x) \neq 0$. If $f = u + iv$, then either $u(x) \neq 0$ or $v(x) \neq 0$. So $\mathcal{A}_{\mathbb{R}}$ vanishes at no point.

Obviously $\mathcal{A}_{\mathbb{R}}$ is a real algebra, and satisfies the hypotheses of the real Stone-Weierstrass theorem. So given any $f = u + iv \in C(X, \mathbb{C})$, we can find $g, h \in \mathcal{A}_{\mathbb{R}}$ such that $|u(t) - g(t)| < \epsilon/2$ and $|v(t) - h(t)| < \epsilon/2$ and so

$$\bigl|f(t) - \bigl(g(t) + ih(t)\bigr)\bigr| = \bigl|u(t) + iv(t) - \bigl(g(t) + ih(t)\bigr)\bigr| \leq |u(t) - g(t)| + |v(t) - h(t)| < \epsilon/2 + \epsilon/2 = \epsilon.$$

So $\mathcal{A}$ is dense in $C(X, \mathbb{C})$. $\square$

An example of why we need the self-adjoint will be in the homework.

Here is an interesting application. We'll do it for the complex version, but the same thing can be done with the real version. It turns out when working with functions of two variables, one would really like to work with functions of the form $f(x)g(y)$ rather than $F(x, y)$. We can't quite do that, but we do have the following.

Application: Any continuous function $F \colon [0,1] \times [0,1] \to \mathbb{C}$ can be approximated uniformly by functions of the form

$$\sum_{j=1}^{n} f_j(x) g_j(y)$$

where $f_j \colon [0,1] \to \mathbb{C}$ and $g_j \colon [0,1] \to \mathbb{C}$ are continuous.

Proof. It is not hard to see that the functions of the above form are a complex algebra. It is equally easy to show that they vanish nowhere, separate points, and the algebra is self-adjoint. As $[0,1] \times [0,1]$ is compact we can apply Stone-Weierstrass to obtain the result. $\square$