104: Introduction to Analysis

Nir Elber

Fall 2021

Contents

1 Pinning Down the Reals **3**
1.1 August 25 . . . 3
1.2 August 30 . . . 4
1.3 September 1 . . . 7
1.4 September 8 . . . 11
1.5 September 13 . . . 15

2 Convergence **19**
2.1 September 15 . . . 19
2.2 September 20 . . . 22
2.3 September 22 . . . 26
2.4 September 27 . . . 32
2.5 September 29 . . . 32
2.6 October 4 . . . 38
2.7 October 6 . . . 44

3 Continuity **49**
3.1 October 13 . . . 49
3.2 October 18 . . . 51
3.3 October 20 . . . 56
3.4 October 25 . . . 62
3.5 October 27 . . . 67
3.6 November 1 . . . 73
3.7 November 3 . . . 78

4 Differentiation **83**
4.1 November 10 . . . 83
4.2 November 15 . . . 86
4.3 November 17 . . . 86

5 Integration **92**
5.1 November 22 . . . 92
5.2 November 29 . . . 96
5.3 December 1 . . . 100

Theme 1: Pinning Down the Reals

1.1 August 25

Let's go ahead and begin.

1.1.1 Logistics

Email is `asharma18@berkeley.edu`. The course website is `ocf.berkeley.edu/ãsharma/Math104`. Namely, we are not using bCourses. Office hours are (tentatively) 3:15PM–4:45PM in 833 Evans and Saturday 2PM–4PM online. Check the website for the Zoom link.

We're using *Elementary Analysis* by Ross, and we will follow the book pretty closely. Most of the things we do will be relatively known, but we will be adding rigor as we go through. For example, the homework is from the book and largely in order.

Do attempt to turn in the homework nontrivially early due to technical issues. Turning in will be done completely online. Please put different sections of the homework to different pages. This will help if a section needs to be moved to a later homework. Also, please take the homework seriously; it will gauge if you are keeping up with the homework. Similarly, working together is somewhat discouraged if you are not actually learning.

Rigor will be important in this class. For example, if asked to find the derivative of $f(x) = x^2$ using the limit definition of the derivative, then one should know the limit definition. These sorts of things are important in this class.

Example 1.1. We can compute

$$\int_1^2 \frac{1}{x}\,dx = \log 2 - \log 1 = \log 2.$$

However, we cannot compute $\int_{-1}^2 1/x\,dx$ because it has problems at 0.

The point of this example is that $\int_{-1}^2 1/x\,dx$ seems like it should be $\log 2$, but this is very false because $\int 1/x\,dx$ is an entire class of functions. Definitions are important here.

1.1.2 Foreshadowing

We will be doing induction proofs in this class. Here's a different kind of thing.

Theorem 1.2 (Extreme value theorem)**.** A continuous function $f : I \to \mathbb{R}$ defined on a closed and bounded interval I achieves an absolute maximum and minimum.

Explaining why this is true requires some thought. We can show that the image of f is bounded because the domain is compact, so if it has no maximum, we could create an infinitely ascending sequence with no bound in the image of f.

This turns out to be quite subtle. Upgrading from closed and bounded intervals is nontrivial because finding the condition "compact" is quite difficult to do. For example, we cannot set $D = \mathbb{Q} \cap [0, 1]$ because of something like $f(x) = |x - \sqrt{2}|$, which achieves no absolute minimum on D.

As another kind of example, we can create a function which is nowhere continuous, which is somewhat nonintuitive given most people's graphical understanding of functions. For example,

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

The main problem here is that

$$\lim_{x \to x_0} f(x)$$

does not exist here (this requires a δ-ε proof, sadly), so it cannot be continuous by the definition of continuous.

Can we create a function which is continuous everywhere but somewhere not differentiable? Here $f(x) = |x|$. This construction can be generalized to any finite set of points S we want to not be differentiable by

$$\sum_{a \in S} |x - a|.$$

Question 1.3. How bad can we make our set of not differentiable points? Can we make our bad points $\mathbb{Q}$? How about all of $\mathbb{R}$?

Here's another question. Consider the sequence $3, 3.1, 3.14, 3.141, \ldots,$ which consists of the truncated Does this sequence converge? Well, we need a good definition of convergence. Cauchy sequences, for example, will do the trick here for $\mathbb{R}$, but this does not converge in $\mathbb{Q}$. Does $\mathbb{R}$ have this same problem? More precisely, this is the question we are asking.

Question 1.4. Suppose a sequence in $\mathbb{R}$ "converges." Must it converge to a real number?

Using fancy words, we are asking if $\mathbb{R}$ is metrically complete. We could ask the same question for $\mathbb{Z}$.

One issue here is that we don't have a good definition of a limit point without a number we actually converge to. For example, perhaps we want convergence to means something like "the terms get closer and closer together," which is connected to Cauchy sequences.

1.2 August 30

We are mostly preluding for the time being, but let's go ahead and jump in with it.

1.2.1 Prelude: Motivating Rigor

Also recall from last time that we can have a sequence of rational numbers which approaches something rational. However, we seem more dubious to having a sequence of real numbers approach something which is not a real number. Part of the problem with putting this doubt to rest is that we don't have a good definition of "convergence" yet.

As a related idea, suppose we have a sequence $\{a_k\}_{k \in \mathbb{N}}$ of only positive real numbers, then the sum

$$\sum_{k=1}^{\infty} a_k$$

is either convergent or diverges to $+\infty$. However, if we permit some of these to be negative, then the sum might just not be well-defined. Viewing the infinite series as a sequence of partial sums, we see that we are hoping that a sequence of real numbers (the partial sums) to only converge to real numbers when they converge at all. So if we care about infinite series convergence types, then we must be careful here.

So we hope quite strongly that convergent real sequences only converge to real numbers. This is tied to *completeness*, which we will do later.

Recall the following function.

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

This function is not continuous at any single point, though it might not be completely apparent because limit laws do not help us here. Namely, all of our limits are bad. For example, let's try to think about

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

Does this even make sense? How would we go about evalauting this? Our Calculus I and II tricks do not really apply here. To make our lives more concrete, let's imagine something like

$$a_n := \begin{cases} 1 & n \text{ is even}, \\ 0 & n \text{ is odd}. \end{cases}$$

Now how would we evaluate $\lim_{n\to\infty} a_n$? Does this even make sense?

The correct way to approach this is via a δ-ε proof, which is a bit hard. As an example, how do we prove something like

$$\lim_{x\to 2} x^2 = 4?$$

Well, we could "just look at it" so that it's not very interesting. But this rigor is reasonable because of garbage like $\lim_{n\to\infty} f(x)$. We need comfort with δ-ε proofs because we need to be able to use them when we don't have other options.

There are some other questions we might be interested in.

- Can you make a functoin continuous at only one point?
- Can you make a function continuous at only the rational numbers?
- Can you make a function continuous at onle the irrational numbers?

Note that we really need δ-ε for these proofs because what other tools do we have against the pathologies we need to create?

As another example, recall the Intermediate value theorem.

Theorem 1.5 (Intermediate value)**.** Fix $f : [a,b] \to \mathbb{R}$ a continuous function. Then, for each y between $f(a)$ and $f(b)$, there exists an x such that $f(x) = y$.

Now, can we prove this? Intuitively, this is clear, but that is not a proof. Similarly, how could we prove the various convergence tests? For example, how could we prove the ratio test? the root test? What would happen if the limits don't exist?

The real point of this class is to add rigor to our various intuitions. There are things that we know, such as the fact that

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

diverges, which are somewhat difficult to prove. But if we wanted to do this rigorously, it is not clear how we could talk about the partial sums of this series. Or, we know that

$$\sum_{k=0}^{\infty} \frac{1}{2k+1} - \frac{1}{2k+2}$$

will converge by (say) the alternating series test, but what about

$$\sum_{k=0}^{\infty} \frac{1}{2k+1} + \frac{1}{2k+2} - \frac{1}{2k+3} - \frac{1}{2k+4}?$$

This requires some trickery; the trickery here turns out to be that we shold group the terms as given and then test convergence.

Here is another series.

$$f(x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots .$$

This series converges at $x = 1$. However, when we take its derivative, we get

$$f'(x) = 1 - x + x^2 - x^3 + \cdots ,$$

which diverges at $x = 1$. Things are stranger at the ends of its interval of convergence, which is roughly because we cannot exchange an infinite sum with an integral whenever we want: the original series is merely conditionally convergent at $x = 1$.

Have yet another question. If we have two everywhere differentiable functions $f, g : \mathbb{R} \to \mathbb{R}$ have identically equal derivateives, must they differ by a constant? Well, sure: $\delta := (f - g)$ has derivative which vanishes everywhere, which implies that the function is constant by the Mean value theorem. We could probably do this by hand with limits if we tried hard enough, but so it goes.

Continuing with our theme, note

$$\int_0^1 x^2\,dx = \frac{1}{3} \qquad \text{and} \qquad \int_0^1 \frac{1}{x}\,dx = \text{diverges}.$$

However, the integral of the product is $\frac{1}{2}$ and notably well-defined. Or even worse, what is the integral of

$$\int_0^1 1_{\mathbb{Q}}(x)\,dx?$$

I think this will turn out to not be well-defined under Riemann integration (though Lebesgue would like to know your location). In general, it is not clear how we should integrate poorly-behaved functions: the Fundamental theorem of calculus only tells us what to for continuous functions. In this class, we would like to build a general criterion for integrable. For example, can we have an integrable function which is nowhere continuous?

And here is a last example. Imagine the plane $\mathbb{R}^2$ equipped with its usual (Euclidean) distance. Pick up two points (x_1, y_1) and (x_2, y_2) so that the distance between them is

$$d((x_1, y_1), (x_2, y_2)) := \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

However, why can't we define distance more pathologically, by

$$d_{\text{taxicab}}((x_1, y_1), (x_2, y_2)) : +|x_2 - x_1| + |y_2 - y_1|?$$

This distance function seems fairly nice. But how nice is it? Or what about the lonely, distance, defined by

$$d_{\text{discrete}}((x_1, y_1), (x_2, y_2)) := \begin{cases} 1 & (x_1, y_1) \neq (x_2, y_2), \\ 0 & (x_1, y_1) = (x_2, y_2). \end{cases}$$

There are lots of these things that we could play around with. Now, how do these things change convergence? We would like convergence to mean that "the distance between is going to 0," but in the discrete metric, it doesn't like it's converging. Sadness ensues. In fact, the only convergent sequences in the lonely metric are eventually constant ones.

1.2.2 Bring Natural

Let's get started with the book now. Our story begins with the natural numbers $\mathbb{N}$, which consist of the positive integers in this class.

One property of $\mathbb{N}$ we care about is that $n \in \mathbb{N}$ implies $n + 1 \in \mathbb{N}$. This appears trivial but is in fact quite fundamental. For example, there is certainly no largest number. Similarly, this gives us rise to induction: if $S \subseteq \mathbb{N}$ with $1 \in S$ and $n \in S \implies n + 1 \in S$, then we are able to conclude that $S = \mathbb{N}$. Somehow, we've gotten all natural numbers! More precisely, we have the following.

Axiom 1.6. Suppose that $P(n)$ is a proposition in $n \in \mathbb{N}$. If $P(1)$ is true, and $P(k)$ implies $P(k+1)$ for any $k \in \mathbb{N}$, then $P(n)$ is true for all $n \in \mathbb{N}$.

Let's do an example of induction.

Proposition 1.7 (Ross 1.7)**.** We prove that $7^n - 6n - 1$ is divisible by 36 for all $n \in \mathbb{N}$.

Proof. We proceed by induction. Our base case is $n = 1$, for which the statement reads $7^1 - 6 \cdot 1 - 1 = 7 - 6 - 1 = 0 = 0 \cdot 36$, so we are safe.

Now suppose that $36 \mid 7^k - 6k - 1$ so that we want to prove $36 \mid 7^{k+1} - 6(k+1) - 1$. The trick is to write

$$\begin{aligned} 7^{k+1} - 6(k+1) - 1 &= 7 \cdot 7^k - 6k - 7, \\ &= 7 \cdot \left(7^k - 6k - 1\right) + 7 \cdot 6k + 7 - 6k - 7 \\ &= 7 \cdot \left(7^k - 6k - 1\right) + 36k, \end{aligned}$$

which is now divisible by 36 by the inductive hypothesis. ■

Here is another example.

Proposition 1.8. We prove that $4^n > 3^n + 2^n$ for each $n \geq 2$.

Proof. Our base case is $n = 2$, for which this reads $16 > 9 + 4 = 13$, so we are safe. Now suppose that the statement is true for n, and we want to show it is true for $n+1$. Well, note that

$$4^{n+1} = 4 \cdot 4^n > 4\left(3^n + 2^n\right) = 4 \cdot 3^n + 4 \cdot 2^n.$$

Now the right-hand side is termwise greater than $3^{n+1} + 2^{n+1}$, so we finish the inductive step. ■

In contrast, how could we go about showing the following?

Proposition 1.9. It is true that $4^x > 3^x + 2^x$ for real numbers $x \geq 2$.

A direct induction won't work here, though an inductive spirit might: for example, if we show that it is true in $[2, 3]$ to start, then we could do an $x \mapsto x + 1$ move to get all real numbers. The more direct way to do this is to say that 4^x is increasing faster than $2^x + 3^x$, though this is somewhat difficult to rigorize.

1.3 September 1

I went to office hours today; it was pretty fun. I exhibited a continuous, surjective function $[0, 1) \to (0, 1)$ and felt smart.

1.3.1 Philosophy about Induction

Let's study the sequence

$$S_n = \sum_{k=1}^{n} \sin k.$$

We would like to have closed form for this series, but it is not at all obvious how to obtain one. And, for example, induction doesn't really help us find such a formula: induction only helps us verify truth, not discover it. This is important to keep in mind.

1.3.2 More Classes of Numbers

From $\mathbb{N}$, the next class we care about is $\mathbb{Z} := \{a-b : a, b \in \mathbb{N}\}$. Then we have $\mathbb{Q} := \{p/q : p, q \in \mathbb{Q}\}$. Formally, we define

$$\mathbb{Q} := \frac{\{(p,q) \in \mathbb{Z}^2 : q \neq 0\}}{\{(p_1, q_1) \sim (p_2, q_2) : \exists c, d : cp_1 = cp_2 \text{ and } dq_1 = dq_2\}}.$$

Alternatively, a number is rational if and only if its decimal expansion is eventually periodic.

Then there are numbers which are not rational: for example, $\sqrt{2} \notin \mathbb{Q}$. We would like to prove this; let's start with a smaller question.

Proposition 1.10. There does not exist an integer n for which $n^2 = 50$.

Proof. We can bound $7^2 < 50 < 8^2$, so $7 < \sqrt{50} < 8$. ■

Why can't we try the same thing for $\mathbb{Q}$? Well, there's just too many (infinitely many) rational numbers between any given two rational numbers. This is not a finite computation here: $\mathbb{Z}$ is nice because it is discrete.

Anyways, let's prove $\sqrt{2} \notin \mathbb{Q}$.

Proposition 1.11. We show $\sqrt{2} \notin \mathbb{Q}$.

Proof. Suppose for the sake of contradiction $\sqrt{2} \in \mathbb{Q}$ so that $\sqrt{2} = p/q$ such that $p, q \in \mathbb{N}$ and $\gcd(p,q) = 1$ (the fraction is reduced). Technically, it suffices to assume that at least one of m, n is odd. But now $q\sqrt{2} = p$ implies that

$$p^2 = 2q^2.$$

Now, p^2 is even, so p is even, so $p = 2p_0$. But then we can rearrange to

$$2p_0^2 = q^2,$$

so q^2 is even, so q is even. However, this violates our assumption that p and q were both even, which is a contradiction. ■

Remark 1.12. The above more or less convinces us that $\mathbb{R}$ is a useful thing to look at: it has numbers like $\sqrt{2}$ which we want but are not immediately accessible via $\mathbb{Q}$.

Similar logic could show that $\sqrt{3}$ or $\sqrt[3]{2}$ are irrational.

After $\mathbb{Q}$, the next class that we care about is the set of algebraic numbers.

Algebraic

Definition 1.13 (Algebraic). We say that a real number α is *algebraic* if and only if there exists a polynomial $p(x) \in \mathbb{Z}[x]$ such that $p(\alpha) = 0$.

Example 1.14. We see that all rational numbers are algebraic: for $\alpha := a/b \in \mathbb{Q}$, we see that α is a root of $p(x) := bx - a \in \mathbb{Z}[x]$.

Example 1.15. We see that there are algebraic numbers which are not rational: for $\alpha := \sqrt{2} \notin \mathbb{Q}$, we see that α is a root of $p(x) := x^2 - 2 \in \mathbb{Z}[x]$.

So we see that $\mathbb{Q} \subsetneq \mathbb{A}$. In fact, most of our friends are algebraic, such as $\sqrt[3]{6}$ or $\sqrt{\sqrt[3]{5} - \sqrt[4]{2}}$ similar.

1.3.3 Rational Root Theorem

The super-powered version of Proposition 1.11 is the Rational root theorem, which turns sieving through $\mathbb{Q}$ into a finite computation.

Theorem 1.16 (Rational root)**.** Fix a polynomial

$$f(x) = \sum_{k=0}^{n} c_k x^k \in \mathbb{Z}[x]$$

with $c_n \neq 0$ and $c_0 \neq 0$. Now suppose that $q = a/b$ is a root of $f(x)$, where $a, b \in \mathbb{Z}$ and $\gcd(a, b) = 1$. Then we claim $a \mid c_n$ and $b \mid c_0$.

Proof. The main point is to plug in a/b into $f(x)$. We see that

$$0 = \sum_{k=0}^{n} c_k \left(\frac{a}{b}\right)^k,$$

so

$$0 = \sum_{k=0}^{n} c_k a^k b^{n-k}.$$

Now, the idea is to isolate

$$c_n a^n = b \cdot \sum_{k=0}^{n-1} -c_k a^k b^{n-k-1},$$

so it follows $b \mid c_n a^n$. However, $\gcd(a, b) = 1$ now forces $\boxed{b \mid c_n}$.

For the other divisibility, we similarly isolate

$$c_0 b^n = a \cdot \sum_{k=1}^{n} -c_k a^{k-1} b^{n-k},$$

so $a \mid c_0 b^n$. But again, $\gcd(a, b) = 1$ forces $\boxed{a \mid c_0}$, finishing. ■

Corollary 1.17. Suppose $\alpha \in \mathbb{Q}$ is the root of a monic polynomial $f(x) \in \mathbb{Z}[x]$. Then $\alpha \in \mathbb{Z}$.

Proof. Writing $\alpha = a/b$ as a reduced fraction, we see that we must have $b \mid \pm 1$ by Theorem 1.16, so $b = \pm 1$. Thus, $\alpha = \pm a \in \mathbb{Z}$, finishing. ■

This is surprisingly powerful. In particular, the Rational root theorem gives us a way to determine if an algebraic integer is rational by turning it into a finite computation: we only have to check the rational numbers with a bounded numerator or denominator. Here are some examples.

Example 1.18. We show that $\sqrt[3]{6} \notin \mathbb{Q}$. Well, it's a root of the monic polynomial

$$x^3 - 6,$$

so $\sqrt[3]{6} \in \mathbb{Q}$ implies $\sqrt[3]{6} \notin \mathbb{Z}$, which is false because $2 < \sqrt[3]{6} < 3$.

Corollary 1.19. Suppose that $n \in \mathbb{Z}$ has $\sqrt{n} \notin \mathbb{Z}$. Then $\sqrt{n} \notin \mathbb{Q}$.

Proof. We show the contrapositive. Indeed, $\sqrt{n} \in \mathbb{Q}$ would imply $\sqrt{n} \in \mathbb{Z}$ by Corollary 1.17: $\sqrt{n}$ is a root of the monic polynomial $x^2 - n$. ∎

Example 1.20. We show that $\alpha := \sqrt{2+\sqrt{2}} \notin \mathbb{Q}$. Note that $\alpha^2 = 2 + \sqrt{2}$, so $\left(\alpha^2 - 2\right)^2 = 2$. Thus, α is a root of the (monic) polynomial

$$f(x) = \left(x^2 - 2\right)^2 - 2 = x^4 - 4x^2 + 2.$$

By Corollary 1.17, we see that any rational root is in $\{\pm 1, \pm 2\}$, but we can check by hand that none of them work.

Example 1.21. We note that $\alpha := \sqrt{4+2\sqrt{3}} - \sqrt{3}$ looks very irrational, but it's actually 1. Note that $\left(1+\sqrt{3}\right)^2 = 4 + 2\sqrt{3}$, so $\alpha = (1+\sqrt{3}) - \sqrt{3} = 1$.

If we wanted to, we could write out a polynomial for the sake of completeness. Note $\alpha + \sqrt{3} = \sqrt{4+2\sqrt{3}}$, so

$$\alpha^2 + 2\alpha\sqrt{3} + 3 = 4 + 2\sqrt{3}.$$

Then $2\sqrt{3}(\alpha - 1) = 1 - \alpha^2$, and we can square both sides to get the polynomial $12(\alpha-1)^2 - \left(\alpha - 1\right)^2 = 0$.

Note that even our attempt to find a polynomial, which comes from trying to show that $\sqrt{4+2\sqrt{3}} - \sqrt{3}$ is irrational, we can look at the polynomial we created, named

$$12(x-1)^2 - \left(x-1\right)^2,$$

has a double root of $x = 1$, so this guides us that maybe the number we were looking at what 1.

1.3.4 Transcendental Talk

People spent a while thinking that $\mathbb{Q}$ has everything we could ever want. Now we are brought to place where $\mathbb{A}$ seems to have everything we could ever want. However, there are real numbers which are not algebraic. The numbers π and e are the typical examples, but how about $2^{\sqrt{2}}$? This is known as the Gelfond–Schneider constant, and it turns out to be not algebraic as well.

The simplest example to prove is

$$L := \sum_{k=1}^{\infty} \frac{1}{10^{n!}}.$$

(This is a "Liouville number.") The problem here is that L has really great rational approximations, even better than any algebraic number could hope for. This sum even provides us a sequence (via the partial sums) that are each algebraic (in fact, rational) that converges to a transcendental number.

So it looks like we are repeatedly able to have sequences which seem to converge (in some decimal expansion sense) but are not converging to a number in our set: it happened for $\mathbb{Q}$, and now it happened for $\mathbb{A}$. So now we have the real numbers to fix this problem.

Real numbers

Definition 1.22 (Real numbers)**.** We define the real numbers as any decimal expansion.

The point here is that we could define $\mathbb{Q}$ as eventually periodic decimals, and then algebraic numbers also have good decimal expansions (or at least rational approximations), but we hope that adding in all real numbers—all decimal expansions—we note that all of our sequences which converge (in some decimal expansion sense) will actually converge to a real number. This is more or less what it means to converge (in some decimal expansion sense).

Example 1.23. There are a lot of real numbers. For example, $0.1234567891011121314151617\ldots$ is not a rational number: its decimal expansion can never repeat. (If it repeated, then the form of our natural numbers would be too restricted.) However, this number is not very useful.

Example 1.24. Another real number we don't care about is

$$\sum_{k=1}^{n} \frac{1}{10^{k^2}}$$

is also irrational because its decimal expansion is again not eventually periodic: for any period length, we can always find a string of constant zeroes of length twice the period, so the period must only consist of zeroes, but this number has 1s as far down in the decimal expansion as we please.

Warning 1.25. We make the first homework due on September 12th.

1.4 September 8

1.4.1 Constructing $\mathbb{R}$

Recall the sequence of rationals

$$1, 1.4, 1.141, \ldots \to \sqrt{2}.$$

The point here is that we have a sequence of rational numbers which converge but not to a real number. Similarly, we can see that

$$2^1, 2^{1.4}, 2^{141}, \ldots \to 2^{\sqrt{2}}$$

produces a sequence of algebraic integers which converge to a non-algebraic number. (We are taking on faith that $2^{\sqrt{2}}$ is not algebraic; this is hard to show.) This is sad.

We would like to stop our sequences from producing new numbers. It turns out that this is a reasonable definition of $\mathbb{R}$.

Reals, I

Definition 1.26 (Reals, I)**.** We can define $\mathbb{R}$ as the set of all numbers which have a rational sequence converging to them.

However, this is not rigorous, and it's not even obvious if $\mathbb{R}$ is closed under taking more convergent sequences.

Let's begin to add rigor. We start with a field.

Field

Definition 1.27 (Field)**.** A field a set F together with two operations $+$ and $\cdot$ which can add, subtract, multiply, and divide in a way that makes sense.

We won't be rigorous about this definition either because this is not an algebra class. It's in the book if you want it. However, the structure we do care about is that $\mathbb{R}$ has an order.

Ordered field

Definition 1.28 (Ordered field)**.** An ordered field is a field F together with a total ordering $\leq$ which behaves nicely with the field operations.

Again, this formal definition is in the book if you want it.

Warning 1.29. There is danger in this definition. For example, we could try to order subsets of (say) $\mathbb{R}$ by containment, but then not all subsets are comparable (e.g., $\{1\}$ and $\{2\}$). Or we could try to order subsets by cardinality, but then we lose trichotomy (e.g., $\{1\}$ and $\{2\}$).

Remark 1.30. We could view all of this adding structure as "throwing out" fake real numbers. For example, there are lots of fields which aren't $\mathbb{R}$, so if we want to define $\mathbb{R}$, we need more structure. For example, adding in the structure of being an ordered field prevents us from confusing $\mathbb{C}$ (which isn't ordered) with $\mathbb{R}$.

Remark 1.31. The homework will contain "axiomatic torture." For example, we will have to prove $0 < 1$ from the ordered field axioms. On one hand, this is painful because it looks obvious; on the other hand, this is necessary because we are trying to be rigorous and hence have to be careful.

1.4.2 Absolute Value

Recall the following definition.

Absolute value

Definition 1.32 (Absolute value)**.** We have that, for $x \in \mathbb{R}$,

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

We would like to know things about the absolute value, such as $|x| \geq 0$ for all $x \in \mathbb{R}$. To have tools, we must prove them; here is one such tool we need.

Proposition 1.33. For $a, b \in \mathbb{R}$, we have $|a| + |b| \geq |a + b|$.

Proof. We see that $a \leq |a|$ and $b \leq |b|$ always, so it follows $a + b \leq |a| + |b|$. We also know that $-a \leq |a|$ and $-b \leq |b|$, so it again follows $-(a+b) \leq |a|+|b|$. However, $|a+b| \in \{\pm(a+b)\}$, so we get $|a+b| \leq |a|+|b|$. ∎

Corollary 1.34 (Triangle inequality)**.** Given $a, b, c \in \mathbb{R}$, we have that

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

Proof. Set $a = x - y$ and $b = y - z$ so that the given inequality is equivalent to

$$|a + b| \stackrel{?}{\leq} |a| + |b|.$$

This is exactly Proposition 1.33. ∎

Remark 1.35. The title "triangle inequality" makes this sound geometric, and indeed it is geometric and will apply in larger contexts.

Let's do an exercise, for fun.

Exercise 1.36 (Ross 3.5)**.** The following are true. Fix $a, b \in \mathbb{R}$.

(a) $|b| \le a$ if and only if $-a \le b \le a$.

(b) $||a| - |b|| \le |a - b|$.

Proof. We take these one at a time.

(a) In one direction, we first assume $|b| \le a$. We use the same trick as in Proposition 1.33. We see that $b \le |b|$ and $-b \le |b|$ both, so it follows

$$b \le |b| \le a \qquad \text{and} \qquad -b \le |b| \le a.$$

The first inequality gives $b \le a$, and the second inequality gives $b \ge -a$, from which the result follows.

In the other direction, we first assume $-a \le b \le a$. Then it follows $b \le a$ and $-a \le b$. This second inequality yields $-b \le a$, so it follows $|b| \le a$ because $|b| \in \{\pm b\}$.

(b) Note that $|a| \le |a - b| + |b|$ by Proposition 1.33 and $|b| \le |b - a| + |a|$ for the same reason. It follows $\pm(|a| - |b|) \le |a - b|$, which gives the result. ■

1.4.3 Talking Bounds

Let's rigorize Definition 1.26. The problem with the rationals and the algebraic numbers is that they have lots of "gaps." Well, should $\mathbb{R}$ have these gaps? We hope not. Proving this requires some care, and it will be critical to defining $\mathbb{R}$.

Our story begins by revising the idea of maximum.

Maximum, minimum

Definition 1.37 (Maximum, minimum)**.** Fix $A \subseteq \mathbb{R}$ a subset.

- We define $\max A$ to be an element of A such that $a \in A$ implies $a \le \max A$.
- We define $\min A$ to be an element of A such that $a \in A$ implies $a \ge \min A$.

We can show that every finite set has a well-defined maximum and minimum, say by induction. However, not all sets have a maximum and/or minimum: for example, $A := \{x \in \mathbb{R} : 0 < x < 1\}$ "should" have 0 and 1 as its minimum and maximum respectively, but they are not actually elements of A. To make our descriptions easier, we take the following definitions.

Interval notation

Definition 1.38 (Interval notation)**.** Given $a, b \in \mathbb{R}$, we define

$$\begin{aligned}(a, b) &:= \{x \in \mathbb{R} : a < x < b\},\\ [a, b) &:= \{x \in \mathbb{R} : a \le x < b\},\\ (a, b] &:= \{x \in \mathbb{R} : a < x \le b\},\\ [a, b] &:= \{x \in \mathbb{R} : a \le x \le b\}.\end{aligned}$$

There is something different here between $\mathbb{Q}$ and $\mathbb{R}$. In $\mathbb{Q}$, the set $[0, \sqrt{2})$ has no maximum, and we cannot add elements of $\mathbb{Q}$ to fix its predicament. However, in $\mathbb{R}$, even though $[0, \sqrt{2})$ still has no maximum, we see that we could add just $\sqrt{2}$ to fix it.

What's going on with $\mathbb{R}^+ := \{x \in \mathbb{R} : x > 0\}$. It has no maximum, but it does have a minimum. The difference here is boundedness.

Boundedness

Definition 1.39 (Boundedness)**.** Fix a nonempty subset $A \subseteq \mathbb{R}$.

- We say that $c \in \mathbb{R}$ is an *upper bound* of A if and only if $a \in A$ implies $a \leq c$. We say that A is *upper-bounded*.
- We say that $c \in \mathbb{R}$ is a *lower bound* of A if and only if $a \in A$ implies $c \leq a$. We say that A is *lower-bounded*.

We say that A is *bounded* if it has both a lower bound and an upper bound.

There are lots of bounded sets which don't have maximum or minimum, but all the bounded sets we can think of feel like they should.

1.4.4 The Completeness Axiom

Let's think about what "feel like they should" really means. Imagine A is bounded above, so that we want it to have a maximum or something like it. However, there are lots of upper bounds (if c is an upper bound, then $c+1$ works), and it's not clear what is the "best" upper bound in the same way that a maximum is clearly the "best" upper bound. For now, fix

$$B := \{x \in \mathbb{R} : x \text{ is an upper bound for } A\}.$$

How should we choose the best upper bound from B? It doesn't have a good maximum, but it does look like it has a minimum!

Example 1.40. For $(0,1)$ or even $[0,1]$, its upper bounds are $[1,\infty)$, which has 1 as its minimum.

Warning 1.41. This is a very special property of $\mathbb{R}$! All we've said is that $\mathbb{R}$ is a real number, but (say) $\mathbb{Q}$ is another ordered field, and in $\mathbb{Q}$, the set $(0,\sqrt{2})$ has $(\sqrt{2},\infty)$ as its upper bounds, which has no minimum!

This discussion motivates the following definition.

Supremum, Infimum

Definition 1.42 (Supremum, Infimum)**.** Fix a nonempty subset $A \subseteq \mathbb{R}$.

- Let U be the set of upper bounds of A. If U has a minimum $\sup A$, then we say that $\sup A$ is the *supremum* of A.
- Let L be the set of lower bounds of A. If L has a maximum $\inf A$, then we say that $\inf A$ is the *infimum* of A.

Example 1.43. If a set has a maximum, then it has a supremum (which is the maximum).

We suspect that all sets bounded above will have a supremum (and all sets bounded below have an infimum). We literally don't have the tools to prove this (our only tool is ordered field), so we make it an axiom.

Axiom 1.44 (Completeness)**.** Any nonempty subset of $\mathbb{R}$ which is bounded above has a supremum.

As an exercise, we can show that any set bounded below has an infimum. It turns out that Axiom 1.44 is exactly what we need to pin down our definition of $\mathbb{R}$. For example, we have the following.

Corollary 1.45. An infinite series with positive terms which does not diverge to ∞ must converge.

Proof. For completeness, let $\{a_k\}_{k=1}^{\infty}$ be our sequence of positive terms. The trick is to look at the set of partial sums

$$S := \left\{\sum_{k=1}^{n} a_k : n \in \mathbb{N}\right\}.$$

This set S has a supremum, which is what the series converges to. We cannot make this terribly rigorous yet because we don't have a good definition of "converge." ■

(After this he started talking philosophy about the Intermediate value theorem, which we're not going to get to for quite some time. We would like to prove it, but we don't have a workable definition of continuous yet.)

1.5 September 13

Alright, take two using VS Code. Hopefully this goes better.

1.5.1 The Archimedean Property

Recall the completeness axiom.

Axiom 1.46 (Completeness)**.** Every set of real numbers bounded above as a supremum.

Recall that this is also equivalent to every set bounded below having an infimum.

Let's relate this to the following.

Proposition 1.47 (Density of $\mathbb{Q}$)**.** Between any two distinct real numbers, there is a rational number.

We would like to know why this is true. This turns out to be related to the following.

Proposition 1.48 (Archimedean)**.** For any a, b positive real numbers, there exists a positive integer n such that $na > b$.

In other words, the multiples of a "just keep going." Let's prove this.

Proof of Proposition 1.48. Suppose for the sake of contradiction that this is false. Then the set

$$S = \{na : n \in \mathbb{N}\}$$

has b as an upper bound: $na \le b$ for each $n \in \mathbb{N}$. It follows that S has a supremum, which we name M. In particular, $M - a$ is not an upper bound for S, so there exists some $n_0 \in \mathbb{N}$ such that $n_0 a > M - a$. But then

$$M < (n_0 + 1)a \in S,$$

which contradicts the fact that M was expected to be an upper bound. ■

Warning 1.49. Note that the Archimedean property is not true in all metric spaces. For example, this fails in $\mathbb{Q}_p$.

Now, let's use the Archimedean property to show Proposition 1.47.

Proof of Proposition 1.47. This requires some care. The idea is to "tile" the real numbers by some rational number less than $b - a > 0$. Indeed, $b - a > 0$ implies that there exists some $n \in \mathbb{N}$ such that $n(b - a) > 1$. (Yes, $1 > 0$.)

Thus, $nb > 1 + na$. We would like to place an integer between na and nb. Rigorizing this requires some care, so we will not do so here, but picking up $k \in \mathbb{Z}$ between na and nb, we see

$$an < k < bn,$$

So $a < \frac{k}{n} < b$. ■

In fact, density of $\mathbb{Q}$ gives us the following.

Proposition 1.50. Any real number is the limit of some sequence of rational numbers.

Proof. Fix r our real number. Then, given any positive integer n, we define q_n as a real number between $r - \frac{1}{n}$ and $r + \frac{1}{n}$. Then we see that

$$|r - q_n| < \frac{1}{n}$$

for each $n \in \mathbb{N}$. So this distance goes to 0, yielding convergence. ■

This turns out to be quite useful because it gives us some small handle on real numbers: at the very least they are all the limit of some sequence of rational numbers.

Remark 1.51. Our metric here matters. If we use the lonely metric, where

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

Here, the only convergent sequences are ones which are eventually constant, so $\mathbb{Q}$ is not dense in $\mathbb{R}$.

Anyways, let's do some example problems.

Proposition 1.52 (Ross 4.5)**.** Fix S a nonempty subset of $\mathbb{R}$ with $\sup S \in S$. Then $\sup S = \max S$.

Proof. Note $s \in S$ implies that $s \leq \sup S$ because $\sup S$ is an upper bound for S. However, $\sup S \in S$ implies that $\sup S = \max S$ as well because maximum is unique. ■

Proposition 1.53 (Ross 4.11)**.** Fix $a, b \in \mathbb{R}$ with $a < b$. Prove that there are infinitely many rational numbers in (a, b).

Proof. Suppose that there are only finitely many rational numbers in (a, b). Surely there is at least one, named q, and surely there are at least two because there is a rational number in (a, q).

Now, because there are only finitely many rational numbers in (a, b), this set of rational numbers has a minimum, named q_0. But then we know there is a rational number in (a, q_0), which violates the minimality of q_0. ■

Remark 1.54. We can remove the contradiction by actually exhibiting the sequence of rational numbers with $q_0 \in (a, b)$ and then recursively defining $q_{k+1} \in (a, q_k)$.

Proposition 1.55 (Ross 4.15)**.** Suppose that $a \leq b + 1/n$ for each $n \in \mathbb{N}$. Then $a \leq b$.

Proof. We show the contrapositive: suppose $a > b$, and we show that there exists $n \in \mathbb{N}$ such that $a \leq b + 1/n$. It follows $a - b > 0$ so that there exists $n \in \mathbb{N}$ with $n(a - b) > 1$, which implies $a > b + \frac{1}{n}$ for some $n \in \mathbb{N}$. ■

1.5.2 Talking $+\infty$ and $-\infty$

It's going to be convenient to be able to talk about $+\infty$ and $-\infty$ in this class, mostly for the sake of intervals and bounding.

Intervals with infinities

Definition 1.56 (Intervals with infinities)**.** We define the interval $(a, +\infty) := \{x \in \mathbb{R} : x > a\}$ and the other intervals with $+\infty$ and $-\infty$ similarly.

Note that there are dangers here: we cannot really do arithmetic with $\pm\infty$. Sometimes we can (e.g., $2 \cdot +\infty = +\infty$), but sometimes we cannot; for example, what is $0 \cdot \infty$?

Warning 1.57. Do not write $[5, \infty]$. You cannot have a closed interval of real numbers actually include ∞.

Anyways, what we are getting out of our ∞ is full completeness.

Supremum and infimum, II

Definition 1.58 (Supremum and infimum, II)**.** Fix S a nonempty set of real numbers. If S is bounded above, we use the definition of $\sup S$ from earlier. Otherwise, we define $\sup S := +\infty$.

Similarly, if S is bounded below, we use the definition of $\inf S$ from earlier. Otherwise, we define $\inf S := -\infty$.

It follows that every set has a supremum and an infimum.

Something else that $\pm\infty$ does is that it helps us disambiguate what "limit doesn't exists" means. For example, being told that a function $f(x)$ has

$$\lim_{x \to 0} f(x) \notin \mathbb{R}$$

could mean all sorts of things: it could be $\infty, -\infty$, too oscillatory, etc. Being given this information is just one way that we can track of this information.

Anyways, let's do some examples.

Proposition 1.59. Given nonempty sets $A, B \subseteq \mathbb{R}$, we have that $\sup(A + B) = \sup A + \sup B$, where $A + B := \{a + b : a \in A \text{ and } b \in B\}$.

Proof. We have to do casework on the supremums being finite or infinite.

- In one case, suppose that at least one of A or B has infinite supremum. Without loss of generality, $\sup A = +\infty$, which is equivalent to A not being bounded above. We claim that $\sup(A + B) = +\infty$ as well.

  Well, for any real number $r \in \mathbb{R}$, and fixing some $b \in B$, there exists $a \in A$ such that $a > r - b$ because A is not bounded above. But then $a + b > r$, so $a + b \in A + B$ exceeds any finite bound $r \in \mathbb{R}$.

- Otherwise, we may fix $\alpha = \sup S$ and $\beta = \sup S$ real numbers. Note that $\alpha + \beta$ is an upper bound for $A + B$ because, for any $a \in A$ and $b \in B$, we have $a \leq \alpha$ and $b \leq \beta$ so that

  $$a + b \leq \alpha + \beta.$$

  It follows that $\sup A + \sup B \geq \sup(A + B)$ because $\sup(A + B)$ is the least upper bound.

  In the other direction, the trick is to show that $\sup(A + B) - \sup A \geq \sup B$, for which it suffices to show that $\sup(A + B) - \sup A$ is an upper bound for B. Well, for any $b \in B$ and $a \in A$, we see that

  $$a + b \leq \sup(A + B)$$

  definitionally. It follows that $\sup(A+B)-b$ is an upper bound for A always, so $\sup A \leq \sup(A+B)-b$, so $b \leq \sup(A + B) - \sup A$. So indeed, $\sup(A + B) - \sup A$ is an upper bound for B, finishing. ■

Proposition 1.60. Fix $S \subseteq \mathbb{R}$ a nonempty subset. Then $\inf S \leq \sup S$.

Proof. Fix some $s \in S$. Then we have $\inf S \leq s$ always (even in case of infinities) as well as $s \leq \sup S$ (even in case of infinities), so $\inf S \leq \sup S$ by transitivity (even in case of infinities). We won't rigorize this, but it would essentially have to be casework. ■

1.5.3 Philosophy

The final section of Chapter 1 in Ross is §6. Roughly speaking, it has to do with the construction of $\mathbb{R}$ from Dedekind cuts. Essentially what this does is prove the Completeness axiom for a particular set so that we can be sure that the real numbers exist, but this confidence will not be relevant to our story. So next class we are talking about sequences and convergence.

Theme 2: Convergence

2.1 September 15

2.1.1 Review

We quickly review a proof of the following fact.

Lemma 2.1. Fix A and B nonempty sets bounded above. Then $\sup(A+B) = \sup A + \sup B$.

Proof. We show that $\sup(A+B) \leq \sup A + \sup B$ and $\sup(A+B) \geq \sup A + \sup B$.

- To show that $\sup(A+B) \leq \sup A + \sup B$, we note that $a \in A$ and $b \in B$ implies that $a + b \leq \sup A + \sup B$, so $\sup A + \sup B$ is an upper bound for $A + B$.
- Now we show that $\sup(A+B) \leq \sup A + \sup B$. Well, fix $y := \sup(A+B)$ so that $y \geq a + b$ for each $a \in A$ and $b \in B$ and
$$y - a \geq b.$$
It follows $y - a$ is an upper bound for B, so $y - a \geq \sup B$, and so $a \leq y - \sup B$, so $y - \sup B$ is an upper bound for A, and $\sup A \leq y - \sup B$, finishing. ■

This is essentially the proof I wrote down last class; he had given a different one.

2.1.2 Limits, Informally

To establish we need rigor, consider the following question.

Question 2.2. Are limits unique?

Certainly they should be, but there is something to proven here, and we don't quite know how to prove this without good definitions.

But before we jump there, let's get comfortable with some of our ideas of limits.

Example 2.3. Fix $a_n := \sin n/n$. We show $a_n \to 0$ as $n \to \infty$.

Indeed, an idea is to use the Squeeze theorem. Namely, we can bound
$$-\frac{1}{n} \leq \frac{\sin n}{n} \leq \frac{1}{n},$$
so sending $n \to \infty$ takes $\sin n/n \to 0$ because $\pm 1/n$ goes to 0.

Example 2.4. Fix $a_n := \frac{2^{n+1}+5}{2^n-7}$. Then $a_n \to 2$ as $n \to \infty$.

Indeed, the trick is to write
$$a_n := \frac{2^{n+1}+5}{2^n-7} \cdot \frac{2^{-n}}{2^{-n}} = \frac{2+5\cdot 2^{-n}}{1-7\cdot 2^{-n}}.$$
Then we see the numerator goes to 2 and the denominator goes to 1, so we see that $a_n \to 2/1 = 2$.

However, not everything is obvious. Consider the sequence

$$a_n := \begin{cases} \frac{n}{n+1} & n \text{ is odd,} \\ 1 & n \text{ is even.} \end{cases}$$

An idea here is that, if the total sequence converges, then any subsequence must converge to the same value. So we can see that the odd case goes to 1, and the even case is 1 as well, so we would like to say this is 1, but this is not a proof. Similarly, surely

$$b_n := \begin{cases} \frac{1}{n} & n \text{ is odd,} \\ 1 & n \text{ is odd,} \end{cases}$$

does not have a limit, but we need a definition.

2.1.3 Limits, Formally

So we move towards a definition.

Limits

Definition 2.5 (Limits). Fix $\{x_k\}_{k=1}^{\infty}$ a sequence of $\mathbb{R}$. Then we say that

$$\lim_{n\to\infty} x_n = \lim x_n = L \in \mathbb{R}$$

if and only if, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $n > N$ implies $|x_n - L| < \varepsilon$.

Roughly speaking, we are saying that, for any error bound ε (such as, say $\varepsilon = 10^{-10!}$), there is a bound N in the sequence such that all points after N are ε-close to L.

Let's use this to kill one of our examples.

Exercise 2.6. Fix

$$a_n := \begin{cases} \frac{n}{n+1} & n \text{ is odd,} \\ 1 & n \text{ is even.} \end{cases}$$

Then $a_n \to 1$ as $n \to \infty$.

Proof. We claim that the limit is $L := 1$. Well, fix any ε so that we want N such that

$$|a_n - 1| < \varepsilon$$

for $n > N$. Well, we can compute that

$$|a_n - 1| \le \left|\frac{n}{n+1} - 1\right| = \frac{1}{n+1} < \frac{1}{n}.$$

So we want $\frac{1}{n} < \varepsilon$, for which $N > 1/\varepsilon$ suffices. We can find an integer N by the Archimedean property. ■

Remark 2.7. To prove that a limit exists, we have to guess its value and then prove it. Namely, it is hopeless if you guessed the wrong number. Conversely, to prove a limit doesn't exist, we have to show that no L can be a limit.

And so let's kill our other example.

Exercise 2.8. Fix

$$b_n := \begin{cases} \frac{1}{n} & n \text{ is odd,} \\ 1 & n \text{ is odd,} \end{cases}$$

Then b_n has no limit.

Proof. Suppose for the sake of contradiction that the limit is $L \neq 1$. Then we find an $\varepsilon := \frac{|L-1|}{2}$ such that and N has $n := \max\{2, 2N\}$ such that

$$|b_n - L| = |1 - L| > \varepsilon.$$

Tracking our quantifiers through, we see that we have shown that the limit of b_n is not 1.

It remains to show that the limit is not 1. Well, fixing $\varepsilon = 1/2$, we see that any N has $n = \max\{3, 2N+1\}$ such that

$$|b_n - 1| = \left|\frac{1}{n} - 1\right| = \frac{n-1}{n} > \frac{1}{2},$$

where the last inequality holds by cross-multiplying. ■

Remark 2.9. Intuitively, we can begin to see why $1_{\mathbb{Q}}$ is nowhere continuous: no matter how close a neighborhood we look at a point, there will be rationals and irrationals around.

And with our actual definition, we can answer Question 2.2.

Proposition 2.10. The limit of a sequence a_n as $n \to \infty$ is unique.

Proof. Suppose that L_1 and L_2 are both limits of a_n as $n \to \infty$, and suppose $L_1 \neq L_2$ for the sake of contradiction. Then set $\varepsilon := \frac{|L_1 - L_2|}{3} > 0$ so that there exists N_1 and N_2 such that

$$n > N_1 \implies |a_n - L_1| < \varepsilon \qquad \text{and} \qquad n > N_2 \implies |a_n - L_2| < \varepsilon.$$

But taking $n > \max\{N_1, N_2\}$, we see that $|L_1 - L_2| < |L_1 - a_n| + |a_n - L_2| < 2\varepsilon < |L_1 - L_2|$ is our contradiction. In other words, a_n has to be in $(L_1 - \varepsilon, L_1 + \varepsilon) \cap (L_2 - \varepsilon, L_2 + \varepsilon)$, which is empty. ■

Let's do some more examples, for fun.

Exercise 2.11 (Ross 7.5(a))**.** We compute

$$\lim_{n\to\infty} \sqrt{n^2+1} - n = 0.$$

Proof. The trick is to write

$$\lim_{n\to\infty} \left(\sqrt{n^2+1} - n\right) \cdot \frac{\sqrt{n^2+1}+n}{\sqrt{n^2+1}+n} = \lim_{n\to\infty} \frac{1}{\sqrt{n^2+1}+n}.$$

Now the top is 1 and the bottom goes to ∞, so the limit vanishes. ■

Exercise 2.12. We prove that

$$\lim_{n\to\infty} \frac{2n-1}{3n+2} = \frac{2}{3}.$$

Proof. We work backwards. The trick is to first study our difference

$$\left|\frac{2n-1}{3n+2} - \frac{2}{3}\right| = \left|\frac{3(2n-1) - 2(3n+2)}{3(3n+2)}\right| = \frac{7}{3(3n+2)}.$$

We see that we can bound this by $\frac{1}{n}$ because $\frac{7}{3(3n+2)} < \frac{1}{n}$ is equivalent to $7n < 9n+6$. So choosing $\boxed{N := 1/\varepsilon}$ so that $n > N$ implies

$$\left|\frac{2n-1}{3n+2} - \frac{2}{3}\right| < \frac{1}{n} < \frac{1}{N} < \varepsilon,$$

finishing. ■

Remark 2.13 (Philosophy)**.** In some deep sense, the reason that we need the precise definition of the limit is that things might get more complicated than what, say, limit laws can control. For example, limits in multiple variables are not really susceptible to limit laws when continuous.

Exercise 2.14 (Ross 8.3)**.** Fix s_n a sequence of nonnegative numbers such that $a_n \to 0$ as $n \to \infty$. Then $\sqrt{a_n} \to 0$ as $n \to \infty$.

Proof. Well, fix some $\varepsilon > 0$ so that we want to find N such that $n > N$ implies that $|\sqrt{s_n}| < \varepsilon$, which is equivalent to $s_n < \varepsilon^2$ because s_n is nonnegative.[1] But $s_n \to 0$, there exists N such that $|s_n| < \varepsilon^2$ for $n > N$, finishing. ■

Remark 2.15. Our N given ε is in some sense "less effective" than in previous examples because our information about s_n is less effective.

Exercise 2.16. We show that
$$\lim_{n\to\infty} \cos\left(\frac{n\pi}{3}\right) \notin \mathbb{R}.$$

Proof. The issue is that our sequence is periodic but nonconstant. Well, the image of the sequence lives in $\{\pm 1/2, \pm 1\}$, so we fix our bad $\varepsilon = 1/5$ smaller than the smallest pairwise distance.

Now suppose for the sake of contradiction we have L is our limit. Then there is N such that $n > N$ implies $\left|\cos\left(\frac{n\pi}{3}\right) - L\right| < \varepsilon$. But then all terms past N live in $(L-\varepsilon, L+\varepsilon)$, which cannot contain all of the $\{\pm 1/2, \pm 1\}$. Explicitly,
$$2 = |1 - -1| \le |1 - L| + |-1 - L| = \left|\cos\left(\frac{6N\pi}{3}\right) - L\right| + \left|\cos\left(\frac{(6N+3)\pi}{3}\right) - L\right| < 2\varepsilon < 2,$$
which is a contradiction. ■

Remark 2.17. At a high level, what is happening is that the distance between consecutive terms is not going to 0, and this violating convergence. This idea will return.

2.2 September 20

I skipped a talk on the Riemann hypothesis to come to this class. This had better be good.

2.2.1 Limit Laws

So we can prove nice things with our rigorous definition of a limit, and by "nice" I mean "obvious." Let's have an example.

Proposition 2.18. A convergent sequence in $\mathbb{R}$ is always bounded.

This is the kind of thing that looks obvious, but proving it formally is annoying

[1] This equivalence is nontrivial but annoying, so we ignore it.

Proof. Fix our sequence $\{a_n\}_{n\in\mathbb{N}}$ of real numbers which converges to L. Now, for any (!) $\varepsilon > 0$, we have an N such that $n > N$ implies

$$|a_n - L| < \varepsilon.$$

It follows that $L - \varepsilon < a_N < L + \varepsilon$ by considering cases. Then consider

$$M := \max\{a_1, \ldots, a_N, L + \varepsilon\}.$$

Then we claim each $n \in \mathbb{N}$ has $a_n < M$. Indeed, if $n \leq N$, then $n \leq M$ because of the maximum. Otherwise $n > N$ so that $a_n < L + \varepsilon$. ■

Let's jump into a limit law.

Proposition 2.19. Suppose that the real sequences $\{s_n\}_{n\in\mathbb{N}}$ and $\{t_n\}_{n\in\mathbb{N}}$ converge to the real numbers s and t respectively. Then

$$\lim_{n\to\infty} s_n t_n = st.$$

Proof. This requires a careful application of the triangle inequality. We show that $|s_n t_n - st| \to 0$. Then we write

$$|s_n t_n - st| = |s_n t_n - s_n t + s_n t - st| \leq |s_n t_n - s_n t| + |s_n t - st|$$

by the triangle inequality. Then we can bound

$$|s_n t_n - st| = |s_n| \cdot |t_n - t| + |t| \cdot |s_n - s|.$$

Now, taking everything as $n \to \infty$, we have $|s_n| \cdot |t_n - t|$ has absolute value bounded by $||s| + 1| \cdot |t_n - t| \to 0$ for sufficiently large n. Similarly, $|t| \cdot |s_n - s|$ will also go to 0.

The end of the proof requires some care, so we will provide a few more details for the first case: we will show that $|s_n| \cdot |t_n - t| \to 0$ more rigorously. Fix any $\varepsilon > 0$. Then there is some N_1 for which $n > N_1$ implies

$$|t_n - t| < \frac{\varepsilon}{|s| + 1},$$

and there is some N_2 for which $n > N_2$ implies $|s_n| < |s| + 1$. Then $N := \max\{N_1, N_2\}$ has $n > N$ implies

$$|s_n| \cdot |t_n - t| < (|s| + 1) \cdot \frac{\varepsilon}{|s| + 1} = \varepsilon,$$

which is what we wanted. ■

Remark 2.20. This is essentially the same trick that proves the product rule for derivatives: we write

$$\begin{aligned}
(fg)'(x) &= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x+h)}{h} + \lim_{h\to 0} \frac{f(x)g(x+h) - f(x)g(x)}{h} \\
&= f'(x)g(x) + f(x)g'(x).
\end{aligned}$$

Let's do an exercise.

Exercise 2.21. We show that

$$\lim_{n\to\infty} n^{1/n} = 1.$$

In calculus, we might want to take the derivative of the numerator and denominator, but we won't do that here.

Proof. We consider the sequence $s_n := n^{1/n} - 1$; we show that $s_n \to 0$. Note that $n \geq 1$ implies $n^{1/n} \geq 1$ implies $s_n \geq 0$.

Now, the trick is to write, for $n > 1$, that

$$n = (1+s_n)^n = 1 + \binom{n}{1} s_n + \binom{n}{2} s_n^2 + \cdots \geq \frac{n(n-1)}{2} s_n^2.$$

It follows that $s_n^2 \leq \frac{2}{n-1}$. However, $\sqrt{\frac{2}{n-1}}$ goes to 0 with some effort[2], so s_n must also go to 0, finishing. ■

2.2.2 Back to $\pm\infty$

What should it mean for a sequence $\{s_n\}_{n\in\mathbb{N}}$ to have

$$\lim_{n\to\infty} s_n = +\infty?$$

Note that this definition requires some care. For example, the sequence

$$s_n := \begin{cases} 0 & n \text{ is even,} \\ n & n \text{ is odd,} \end{cases}$$

is mostly increasing and is not bounded above, but it does not go to ∞. The rigorous definition is as follows.

$s_n \to \infty$

Definition 2.22 ($s_n \to \infty$)**.** Fix $\{s_n\}_{n\in\mathbb{N}}$ a sequence of real numbers. Then $\lim s_n = \infty$ if and only if, for every M, there exists an N such that $n > N$ implies

$$s_n > M.$$

Intuitively, s_n must keep climbing and eventually never go back. Note this does not mean monotonic.

Remark 2.23. The above definition uses an M and not a ε because M is not intended to be a small number; M is supposed to be big. We could use $1/\varepsilon$ if we wanted, but this is unnecessarily awkward.

Warning 2.24. To prove

$$\lim_{n\to\infty} s_n = +\infty,$$

do not fix $\varepsilon > 0$ and try to find N such that $n > N$ implies $|s_n - \infty| < \varepsilon$. The issue here is that $|s_n - \infty|$ does not

Similarly, we have the following, almost dual definition.

$s_n \to -\infty$

Definition 2.25 ($s_n \to -\infty$)**.** Fix $\{s_n\}_{n\in\mathbb{N}}$ a sequence of real numbers. Then $\lim s_n = \infty$ if and only if, for every M, there exists an N such that $n > N$ implies

$$s_n < M.$$

In this definition, M is intended to be a very negative number.

Now that we've added $\pm\infty$ to our vocabulary, we would like to codify this.

Has a limit

Definition 2.26 (Has a limit)**.** Given a sequence $\{s_n\}_{n\in\mathbb{N}}$, we say that s_n *has a limit* if and only if

$$\lim_{n\to\infty} s_n \in \mathbb{R} \cup \{\pm\infty\}.$$

[2] We will not do this here, but we just have to show that $\sqrt{\frac{2}{n-1}}$ gets arbitrarily small.

Note that having a limit is not the same as converging; converging excludes $\pm\infty$. Having a limit is more about excluding sequences which oscillate too much. The language is careful here: converging is synonymous with the limit existing, which is distinct from the sequence having a limit.

Anyways, let's do some exercises.

Proposition 2.27. Fix $\{s_n\}_{n\in\mathbb{N}}$ a sequence of positive real numbers. Then

$$\lim_{n\to\infty} s_n = \infty \iff \lim_{n\to\infty} \frac{1}{s_n} = 0.$$

Proof. Again, this statement is too general for us to be able to do anything other than the formal definition of a limit. We show the directions one at a time.

- Fix s_n with $s_n \to \infty$, and fix any $\varepsilon > 0$. Then $1/\varepsilon > 0$, so there exists N such that $n > N$ implies

  $$s_n > 1/\varepsilon > 0.$$

  From this it follows $0 < 1/s_n < \varepsilon$, so $|1/s_n| < \varepsilon$. So our N witnesses that $1/s_n \to 0$.

- In the other direction, fix s_n with $1/s_n \to 0$, and fix any M. If $M \le 0$, then there is some N such that $1/s_n < 1$ so that $s_n > 1 > M$.

  Otherwise, $M > 0$, so we can take $\varepsilon = 1/M > 0$ and find N so that $n > N$ implies

  $$1/s_n < \varepsilon.$$

  It follows $s_n > 1/\varepsilon = M$, so our N witnesses $s_n \to \infty$. ■

Exercise 2.28. Fix $t_1 = 1$ and define $\{t_n\}_{n\in\mathbb{N}}$ recursively by

$$t_{n+1} = \frac{t_n^2 + 2}{2t_n}$$

for $n \ge 1$. Then, given $\lim t_n$ is a real number, we can compute the limit is $\sqrt{2}$.

Proof. The point is to take $n \to \infty$ on both sides of our recursion.

$$t_{n+1} = \frac{t_n^2 + 2}{2t_n}. \tag{$*$}$$

Now, we can show that

$$\lim_{n\to\infty} t_{n+1} = \lim_{n\to\infty} t_n =: t$$

by shifting over by one (we won't be rigorous here). We would like to use limit laws to finish, but the denominator of $(*)$ has t_n in it, which we need to verify does not vanish. Well, we claim

$$t_n \ge 1$$

by induction. Surely this is true for $n = 1$; then by induction, we note

$$\frac{t_n^2 + 2}{2t_n} = \frac{t_n}{2} + \frac{1}{t_n}.$$

If $t_n \ge 2$, then the first term is at least 1; if $2 \ge t_n \ge 1$, then the sum of the two terms is at least $\frac{1}{2} + \frac{1}{2} = 1$.

Now we can use limit laws to note that $(*)$ implies

$$t = \lim_{n\to\infty} t_{n+1} = \frac{\lim\limits_{n\to\infty} t_n^2 + 2}{\lim\limits_{n\to\infty} 2t_n} = \frac{t^2+2}{2t}$$

after using limit laws. Solving, we see that $2t^2 = t^2 + 2$, so $t^2 = 2$, so $t \in \{\pm\sqrt{2}\}$. Then we see $t \geq 1$ implies $\boxed{t = \sqrt{2}}$ is forced. ■

Remark 2.29. It is a very strong assumption that the limit exists, which is what lets the above argument function. For example, if we change $t_1 = 1$ to $t_1 = -1$, then the limit "should" be $-\sqrt{2}$ (all terms are negative), but it's not obvious that the limit should converge anymore.

Exercise 2.30 (Ross 9.11(a))**.** Fix $\{s_n\}_{n\in\mathbb{N}}$ and $\{t_n\}_{n\in\mathbb{N}}$ sequences such that $s_n \to \infty$ and $\inf\{t_n\}_{n\in\mathbb{N}} > -\infty$. Then

$$\lim_{n\to\infty} (s_n + t_n) = \infty.$$

Proof. Fix any lower bound $m \in \mathbb{R}$ for $\{t_n\}_{n\in\mathbb{N}}$. Now, fix any $M > 0$ so that we want to find N such that $n > N$ implies

$$s_n + t_n > M.$$

Well, we have that $t_n \geq m$ so that the above condition is equivalent to

$$s_n > M - m.$$

To finish, $s_n \to \infty$ promises us some N such that $n > N$ implies $s_n > M - m$, which is exactly what we wanted. ■

Exercise 2.31 (Ross 9.15)**.** Fix $a \in \mathbb{R}$ a positive real number. Then

$$\lim_{n\to\infty} \frac{a^n}{n!} = 0.$$

Proof. For brevity, fix $d_n := a^n/n!$. Consider $N \in \mathbb{N}$ such that $N > 2a$; the main idea is that after N, d_n is being (more than) halved each time and therefore geometrically decreasing. So let's start bounding at d_N: we see

$$d_{N+k} = d_N \prod_{\ell=1}^{k} \frac{a}{N+\ell} \leq d_N \prod_{\ell=1}^{k} \frac{a}{N} \leq d_N \prod_{\ell=1}^{k} \frac{1}{2} = d_N \left(\frac{1}{2}\right)^k.$$

But now, as $k \to \infty$, we have $(1/2)^k \to 0$, so $d_N(1/2)^k \to 0$ as well.[3] It follows that $d_n \to 0$ by squeezing. ■

Remark 2.32. One of the motivations for taking $N > 2a$ is that, after this point, the sequence is definitely (monotonically!) decreasing, and being bounded below by 0, we can be fairly confident that the limit exists. This idea will come up again.

2.3 September 22

Here we go again.

[3] It is surprisingly technical to show that $(1/2)^k \to 0$, but it can be done.

2.3.1 Clarifications

We quickly return to some terminology. For a sequence $\{a_k\}_{k\in\mathbb{N}}$ of real numbers, we say that the sequence "has a limit" if

$$\lim_{n\to\infty} a_n \in \mathbb{R} \cup \{\pm\infty\},$$

and $\{a_k\}_{k\in\mathbb{N}}$ "converges" if

$$\lim_{n\to\infty} a_n \in \mathbb{R}.$$

The ambiguous terminology is what it means for a limit to "exist." Does "exist" force $\mathbb{R}$ or $\mathbb{R}\cup\{\pm\infty\}$?

Warning 2.33. For this course, we will say that the limit of a sequence $\{a_k\}_{k\in\mathbb{N}}$ *exists* if and only if

$$\lim_{n\to\infty} a_n \in \mathbb{R} \cup \{\pm\infty\}.$$

So this is another thing we have to keep track of. The main distinction here matters most on the homework; on exams, just ask the professor what is meant.

2.3.2 Monotonic Sequences

The idea here is that sequences whose terms "get closer" as the sequence goes on should mean that the sequence should converge in $\mathbb{R}$. However, this is a bit annoying to prove because actually finding the limit is a bit annoying, and we still need to rigorize "get closer."

To rigorize "get closer," we begin by talking about monotonic sequences.

Flavors of monotonic

Definition 2.34 (Flavors of monotonic). Fix $\{a_k\}_{k\in\mathbb{N}}$ a sequence of real numbers.

- We say $\{a_k\}_{k\in\mathbb{N}}$ is *increasing* if $a_k \le a_{k+1}$ for each k.
- We say $\{a_k\}_{k\in\mathbb{N}}$ is *strictly increasing* if $a_k < a_{k+1}$ for each k.
- We say $\{a_k\}_{k\in\mathbb{N}}$ is *decreasing* if $a_k \ge a_{k+1}$ for each k.
- We say $\{a_k\}_{k\in\mathbb{N}}$ is *strictly decreasing* if $a_k > a_{k+1}$ for each k.

If any of the above are satisfied, we say that $\{a_k\}_{k\in\mathbb{N}}$ is *monotonic*.

Here is an important result.

Theorem 2.35. Any bounded, monotonic sequence in $\mathbb{R}$ converges.

Proof. Let's say $\{a_k\}_{k\in\mathbb{N}}$ is bounded and increasing (without loss of generality). Note the sequence is upper-bounded and hence has a supremum, which we name a. We show that

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

Well, take any $\varepsilon > 0$. Then $(a-\varepsilon, a]$ must contain some a_N, for otherwise $a_k \le a-\varepsilon < a$ for each a_k, so $a-\varepsilon$ would be a lesser upper bound than a.

Now, for each $n > N$, we have

$$a - \varepsilon < a_N < a_n \le a,$$

so it follows that $|a_n - a| < \varepsilon$, completing the proof. ■

So the above result is nice, but it's pretty restricted in scope: most sequences are not going to be monotonic. Of course, the above logic still works if our sequence is "eventually" monotonic, but still most sequences are not monotonic.

However, if we expand our view, we can ask if every sequence has some monotonic subsequence, which would mean that every sequence has some notion of getting close to somewhere. Sure, it's possible that different subsequences converge to different places, which is annoying, but it's still information. At a high level, our goal is to describe how sequences approach numbers, and keeping track of multiple limits is one way this happens.

Remark 2.36. Theorem 2.35 fails in $\mathbb{Q}$ because we don't have completeness: $3, 3.1, 3.14, 3.141$ is a bounded sequence of rational numbers which does not converge to a rational number.

Exercise 2.37. Consider the sequence $\{s_k\}_{k\in\mathbb{N}}$ defined recursively by $s_1 = 1$ and

$$s_{n+1} = \sqrt{2+s_n}$$

for $n \in \mathbb{N}$. Then we can prove $\lim s_n = 2$.

Proof. The proof is in two steps.

1. We show that the limit exists, also in two steps: we show that $\{s_k\}_{k\in\mathbb{N}}$ is bounded and monotonic.
   (a) We claim that $1 \le s_n \le 2$ for each n, by induction. Of course this is true for $n = 1$. Then, if $1 \le s_n \le 2$, it follows
   $$3 \le 2 + s_n \le 4,$$
   so
   $$1 \le \sqrt{3} \le s_{n+1} \le 2,$$
   which finishes.
   (b) We show that $\{s_k\}_{k\in\mathbb{N}}$ is increasing. The trick is to show that
   $$\sqrt{2+x} \overset{?}{\ge} x$$
   for any $x \in [1, 2]$. Well, this is equivalent to $x + 2 \ge x^2$, which is equivalent to $(x + 1)(x - 2) \le 0$, which is true on $x \in [-1, 2]$ because $x + 1 \ge 0$ and $x + 2 \le 0$ here.
   It follows that $s_{n+1} = \sqrt{2+s_n} \ge s_n$ from the bounding above.
2. Now let the limit be L because we know it exists. Then taking both sides
   $$s_{n+1} = \sqrt{2+s_n}$$
   as $n \to \infty$, we find $L = \sqrt{2+L}$. Thus, $0 = L^2 - (2 - L) = (L + 1)(L - 2)$, so $L \in \{-1, 2\}$. But the s_k are all larger than 1, so $L \ge 1$, so we have $\boxed{L = 2}$. ■

Remark 2.38. This is quite remarkable! Actually finding this limit by some explicit formula would be quite annoying to do, and indeed, we did not have to.

Remark 2.39. Theorem 2.35 kind of shows that all decimal expansions are real numbers. If we accept that rational numbers are real, then the fact that all bounded, monotonic sequences converge, then

$$3, 3.1, 3.14, 3.141, \ldots$$

must converge to some real number, no matter the decimals. (In particular, this sequences is increasing.)

Here is the other side of Theorem 2.35.

Proposition 2.40. Any unbounded, increasing sequence must go to ∞.

Proof. We show the contrapositive; take $\{a_k\}_{k\in\mathbb{N}}$ which does not go to ∞. Then there exists some M such that there is no N such that $n > N$ implies

$$a_n > M.$$

In particular, pushing our quantifier exchange through, for all N, there is some $n > N$ such that $a_n \le M$. However, $n > N$ gives $a_n \le a_N$, so for all N, we conclude that $a_N \le M$. (Here we used that $\{a_k\}_{k\in\mathbb{N}}$ is increasing!)

Thus, $\{a_k\}_{k\in\mathbb{N}}$ is upper-bounded by M. We can take a_1 as a lower bound to finish. ■

So it follows after some more work that any monotonic sequence "has a limit."

2.3.3 Introducing limsup and liminf

We hope to use monotone sequences to gain some control on all sequences. Here is one idea: fix $\{s_k\}_{k\in\mathbb{N}}$ a sequence of real numbers and define

$$t_N := \sup\{s_n : n > N\}.$$

Note that it's possible for t_N to be ∞, and in fact, if one of them is ∞, then all of them are ∞.[4]

So for now, we fix $\{s_k\}_{k\in\mathbb{N}}$ to be bounded above so that each $\{s_n : n > N\}$ is also bounded above, so each t_N is a real number. Now we consider

$$\lim_{N\to\infty} t_N.$$

This certainly "exists" because $\{t_N\}_{N\in\mathbb{N}}$ is decreasing and hence monotonic: the fact that $N_1 \le N_2$ implies $\{s_n : n > N_2\} \subseteq \{s_n : n > N_1\}$ implies $t_{N_1} \le t_{N_2}$ because t_{N_1} will upper-bound $\{s_n : n > N_2\}$.

This gives us the following definition.

$\limsup$ and $\liminf$

Definition 2.41 ($\limsup$ and $\liminf$)**.** Fix $\{s_k\}_{k\in\mathbb{N}}$ a sequence of real numbers. Then we define

$$\limsup_{n\to\infty} s_k := \lim_{N\to\infty} \sup\{s_n : n > N\}.$$

Similarly,

$$\liminf_{n\to\infty} s_k := \lim_{N\to\infty} \inf\{s_n : n > N\}.$$

Warning 2.42. If $\{s_k\}_{k\in\mathbb{N}}$ is not bounded above, then $\limsup$ will be $+\infty$. If $\{s_k\}_{k\in\mathbb{N}}$ is something like $s_k = -k$, then

Regardless of the above warning, we do know that $\limsup$ and $\liminf$ will always exist from the preceding discussion.

Let's do an example.

Proposition 2.43. Suppose that $\{s_k\}_{k\in\mathbb{N}}$ is a sequence with $s_k \to s$ as $k \to \infty$. Then we claim that

$$\limsup s_n = \liminf s_n = s.$$

Proof. We show that $\limsup s_n = s$, and the $\liminf$ argument is similar. We also won't talk about $s \in \{\pm\infty\}$: $s = +\infty$ just means that s_n is not bounded above, and $s = -\infty$ can be done with some effort as well.

So let's say $s \in \mathbb{R}$. Fix any $\varepsilon > 0$. Then there is some N for which $n > N$ implies $s - \varepsilon < s_n < s + \varepsilon$, which implies that

$$\sup\{s_n : n > N\} \in (s - \varepsilon, s + \varepsilon).$$

This shows that $\lim_{N\to\infty} \sup\{s_n : n > N\} = s$. ■

[4] Roughly speaking, this is because the "finite truncation" that t_N does will not mess with the long-term behavior of $\{s_k\}_{k\in\mathbb{N}}$.

In fact, the converse is also true.

Proposition 2.44. Suppose that $\{s_k\}_{k\in\mathbb{N}}$ is a sequence with

$$\limsup_{n\to\infty} s_n = \liminf_{n\to\infty} s_n = s.$$

Then $s_k \to s$ as well.

Proof. We outline. Again, we ignore $\pm\infty$, though they are hard. The idea is that, for any $\varepsilon > 0$, there is some N_1 and N_2 such that $N > \max\{N_1, N_2\}$ implies

$$s - \varepsilon < \inf\{s_n : n > N\} \leq \sup\{s_n : n > N\} < s + \varepsilon.$$

Then any $n > N$ implies $s - \varepsilon < s_n < s + \varepsilon$ as well, finishing. ■

In general, the best we can say is that

$$\inf\{s_n : n > N\} \leq \sup\{s_n : n > N\},$$

so it follows that $\limsup s_n \geq \liminf s_n$.

2.3.4 Cauchy Sequences

We are beginning to realize our goal of tracking how sequences converge. For example, if we are told that $\{s_k\}_{k\in\mathbb{N}}$ has $\limsup s_k = 1$ and $\liminf s_k = 0$. Then, roughly speaking, $\{s_k\}_{k\in\mathbb{N}}$ should "oscillate" in some sense between 0 and 1, even in the long term.

So, for example, $\{s_k\}_{k\in\mathbb{N}}$ has infinitely many terms at least $3/4$ and at most $1/4$, so there are infinitely many pairs with distance at most $1/2$.

With this in mind, the next big idea here is that of a Cauchy sequence.

Cauchy sequence

Definition 2.45 (Cauchy sequence). Fix $\{s_k\}_{k\in\mathbb{N}}$ a sequence of real numbers. We say that $\{s_k\}_{k\in\mathbb{N}}$ is a *Cauchy sequence* if and only if, for each $\varepsilon > 0$, there exists some N such that $n, m > N$ implies

$$|s_n - s_m| < \varepsilon.$$

The advantage here is that Cauchy sequences let us focus locally on the the terms only without having to "guess" a limit as we've been doing. However, this is somewhat worse because we now have two indices n and m to keep track of.

We note the following.

Proposition 2.46. Any converging sequence is a Cauchy sequence.

Proof. Suppose that $\{s_k\}_{k\in\mathbb{N}}$ converges to some $s \in \mathbb{R}$. Then, for any ε, there exists some N such that $n > N$ implies

$$|s_n - s| < \varepsilon/2.$$

This factor of $1/2$ is a trick that will help us momentarily. Indeed, for $n, m > N$, it follows

$$|s_n - s_m| \leq |s_n - s| + |s - s_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

which proves $\{s_k\}_{k\in\mathbb{N}}$ is converging. ■

Remark 2.47. The above proof does not use the completeness axiom, so this is also true in, say, $\mathbb{Q}$.

What about the converse? Indeed, it is true.

Proposition 2.48. Any Cauchy sequence converges to some real number.

Proof. We outline; fix $\{s_k\}_{k\in\mathbb{N}}$ a Cauchy sequence. The trick is to show that

$$\limsup_{n\to\infty} s_n = \liminf_{n\to\infty} s_n,$$

which roughly holds because, after long enough, the $\sup\{s_n : n > N\}$ and $\inf\{s_n : n > N\}$ need to get closer together because of the Cauchy condition. ■

Remark 2.49. The above proof does use completeness, and it is not true in $\mathbb{Q}$, using the typical examples. It is an interesting question which kinds of spaces have that all Cauchy sequences converge.

Let's do an exercise.

Exercise 2.50 (Ross 10.7)**.** Fix $S \subseteq \mathbb{R}$ a bounded, nonempty subset such that $\sup S \notin S$. Then there is a sequence $\{s_k\}_{k\in\mathbb{N}} \subseteq S$ which converges to $\sup S$.

Proof. Note that, for each $k \in \mathbb{N}$, we can fix $\varepsilon = 1/k$ and find some $s_k \in S$ such that $\sup S - 1/k < s_k$ because $\sup S - 1/k$ cannot be an upper bound. Then we can show

$$\lim_{k\to\infty} s_k = \sup S$$

in the usual manner. The point is that

$$\sup S - \frac{1}{k} < s_k < \sup S + \frac{1}{k},$$

so we can take $N = 1/\varepsilon$ in our formal proof. We won't write this out fully. ■

Remark 2.51. The above proof is not very interesting, but it's leading towards the following question: if $\{s_k\}_{k\in\mathbb{N}}$, can we find a subsequence converging to $\limsup s_k$?

Exercise 2.52 (Ross 10.11)**.** Define the sequence $\{t_k\}_{k\in\mathbb{N}}$ by $t_1 := 1$ and

$$t_{n+1} = \left(1 - \frac{1}{4n^2}\right) t_n$$

for $n \in \mathbb{N}$.

Proof. We can show inductively that this sequence is positive, and it is decreasing because $1 - \frac{1}{4n^2} < 1$. ■

It's an interesting question to evaluate this limit. I think it turns out to be $2/\pi$ due to some kind of Weierstrass factorization argument.

2.4 September 27

I didn't take notes today because I went to Professor Beneish's talk on arithmetic statistics. It was probably the most fun I've had all semester, so no, I do not have regrets. Anyways, I am told we covered §11 and §12 during lecture.

2.5 September 29

Here we go. We're talking about §13 today. It's optional but important.

2.5.1 Metrics

One of the main idea is to expand the theory we've built over $\mathbb{R}$ to work more generally. What is something nice that $\mathbb{R}$ has? Well, $\mathbb{R}$ has a good notion of "distance."

Metric space

Definition 2.53 (Metric space)**.** Given a set X, we say that a function $d : X^2 \to \mathbb{R}_{\geq 0}$ is a *metric* if and only if it satisfies the following conditions; fix any $x, y, z \in X$.

- Distance-zero: $d(x, x) = 0$, and $d(x, y) > 0$ if $x \neq y$. In other words, $d(x, y) = 0$ if and only if $x = y$.
- Symmetry: $d(x, y) = d(y, x)$.
- Triangle inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

In this case, we call X a *metric space.*

Warning 2.54. To be a metric, we must satisfy all of the above conditions. They are annoying, but they are necessary.

A while ago, we defined other distance functions; they were metrics. Recall the following examples.

Example 2.55. On $\mathbb{R}^2$, the function

$$d_{\text{Euclid}}\big((x_1, y_1), (x_2, y_2)\big) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2^2)}$$

is the usual, Euclidean metric.

Example 2.56. On $\mathbb{R}^2$, the function

$$d_{\text{taxi}}\big((x_1, y_1), (x_2, y_2)\big) = |x_1 - x_2| + |y_1 - y_2|$$

is called the taxicab metric. Physically speaking, this is the distance we have to go if we can only walk along "streets" parallel the axes.

Example 2.57. On $\mathbb{R}^2$, the function

$$d_{\text{tele}}\big((x_1, y_1), (x_2, y_2)\big) = \begin{cases} 1 & (x_1, y_1) \neq (x_2, y_2), \\ 0 & (x_1, y_1) = (x_2, y_2). \end{cases}$$

is called the tele-metric. Physically speaking, this is the distance we have to go in the internet: it's all a click away.

We want these metrics to be nice with each other. For example, if a sequence of points converges to the origin using d_{Euclid}, then does it converge to the origin in d_{taxi}? What about d_{tele}?

Anyways, let's do a more nontrivial example.

Exercise 2.58 (Ross 13.3)**.** Let B be the set of bounded sequences in $\mathbb{R}$. Then, given two sequences $\{x_k\}_{k\in\mathbb{N}}, \{y_k\}_{k\in\mathbb{N}} \in B$, we define

$$d(x, y) = \sup_k |x_k - y_k|.$$

Then d is a metric on B.

Proof. We can check the conditions one at a time.

- Distance-zero: the only way to make the supremum of the differences zero is to make everything zero, so all elements are equal.
- Symmetry: the absolute value respects negation.
- Triangle inequality: this is checked by force. ■

2.5.2 Convergence Ideas

To answer these questions, we need a good notion of convergence. Recall the definition in $\mathbb{R}$.

Convergence in $\mathbb{R}$

Definition 2.59 (Convergence in $\mathbb{R}$)**.** A sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ *converges* to $x \in \mathbb{R}$ if and only if, for each $\varepsilon > 0$, there exists N such that $n > N$ implies

$$|x - x_n| < \varepsilon.$$

To generalize this to more general metric spaces, we note that $|x - x_n|$ is really just the distance between x and x_n. Here is the general notion for metric spaces.

Convergence in metric spaces

Definition 2.60 (Convergence in metric spaces)**.** Fix X a metric space with metric d. Then a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ *converges* to $x \in X$ if and only if, for each $\varepsilon > 0$, there exists N such that $n > N$ implies

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

Note this is essentially the same definition as in $\mathbb{R}$.

How about Cauchy sequences? Here was our definition in $\mathbb{R}$.

Cauchy in $\mathbb{R}$

Definition 2.61 (Cauchy in $\mathbb{R}$)**.** A sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ is *Cauchy* if and only if, for each $\varepsilon > 0$, there exists N such that $m, n > N$ implies

$$|x_m - x_n| < \varepsilon.$$

Again, to generalize, we swap out distance in $\mathbb{R}$ with general distance.

Cauchy in metricspaces

Definition 2.62 (Cauchy in metricspaces)**.** A sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ is *Cauchy* if and only if, for each $\varepsilon > 0$, there exists N such that $m, n > N$ implies

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

2.5.3 Completeness

Completeness of $\mathbb{R}$ was another part of our story here. In some sense, this came down to all Cauchy sequences converged, which are notions we have defined. So we have the following.

Complete

Definition 2.63 (Complete)**.** A metric space X is *complete* if every Cauchy sequence converged.

Note that all convergent sequences are Cauchy[5], so the reverse direction is the kicker.

Example 2.64. We showed that $\mathbb{R}$ was complete, which roughly came from the Completeness axiom.

Non-Example 2.65. We know that $\mathbb{Q}$ is not complete: take the sequence $\{\lfloor \pi n\rfloor / n\}_{n\in\mathbb{N}}$.

Remark 2.66. We can define $\mathbb{R}$ as equivalence classes of Cauchy sequences in $\mathbb{Q}$, if we wanted.

What about $\mathbb{R}^2$ (with the Euclidean metric)? is it complete? In particular, does every Cauchy sequence converge?

Proposition 2.67. Fix $D \in \mathbb{N}$. Then $\mathbb{R}^D$ is a complete metric space.

Proof. Suppose that $\{x_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence so that we need to show it converges. For concreteness, given $x_k \in \mathbb{R}^D$, we let $\pi_\ell x_k$ be the ℓth coordinate.

We claim that $\{\pi_\ell x_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence. Indeed, for any $\varepsilon > 0$, we know there is some N so that $n, m > N$ implies

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

But

$$|\pi_\ell x_n - \pi_\ell x_m| \le \sqrt{\sum_{k=1}^{D} (\pi_k x_m - \pi_k x_n)^2} = d(x_n, x_m) < \varepsilon,$$

so indeed, $|\pi_\ell x_n - \pi_\ell x_m| < \varepsilon$, making $\{\pi_\ell x_k\}_{k\in\mathbb{N}}$ a Cauchy sequence.

Now, because each coordinate projects into a Cauchy sequence, each coordinate will converge eventually because $\mathbb{R}$ is complete, so we combine these converging sequences into

$$y := (y_\ell)_{\ell=1}^{D},$$

where $\pi_\ell x_n \to y_\ell$ as $n \to \infty$. We claim that $x_n \to y$ as $n \to \infty$. Now, for any $\varepsilon > 0$, we know that we can find an N_ℓ for each coordinate such that $n > N_\ell$ implies

$$|\pi_\ell x_n - y_\ell| < \frac{\varepsilon}{\sqrt{D}}.$$

Then, for $n > \max_\ell\{N_\ell\}$, we have

$$d(x_n, y) = \sqrt{\sum_{k=1}^{D} (\pi_\ell x_\ell - y_\ell)^2} < \sqrt{\sum_{k=1}^{D} \left(\frac{\varepsilon}{\sqrt{D}}\right)^2} = \sqrt{D \cdot \frac{\varepsilon^2}{D}} = \varepsilon,$$

verifying that $x_n \to y$ as $n \to \infty$. ■

We can also define bounded; we just give the definition.

[5] If $a_n \to a$, then for any $\varepsilon > 0$, find N for which $n > N$ implies $|a_n - a| < \varepsilon$. Then $n, m > N$ implies $|a_n - a_m| < \varepsilon$.

Bounded

Definition 2.68 (Bounded). A subset $S \subseteq X$ is "bounded" if and only if there exists $x_0 \in X$ and $r \in R$ such that $x \in S$ implies $d(x, x_0) < r$. In other words, we can put S in a box.

One important thing we did in $\mathbb{R}$ is that every bounded sequence has a convergent subsequence. Is this true in $\mathbb{R}^n$? Sure: we can find a convergent subsequence for the first coordinate, then find a convergent subsequence of that for the second coordinate, and so on. We'll leave this area with a question.

Question 2.69. Is this true for general complete, metric spaces?

The answer turns out to be no; see here. In short, we can use the tele-metric; here the sequence

$$1, 2, 3, \ldots$$

is bounded because all elements of $\mathbb{R}$ are a distance of 1 away from (say) 0. However, this sequence does not converge because it is not Cauchy: for any N, we can find unequal $n, m > N$ so that $d(n, m) = 1$.

2.5.4 Open Sets

We want to generalize "open" and "closed" intervals in $\mathbb{R}$, which connects much deeper inside topology. Starting easy, the open interval $(-1, 1)$ seems like it should generalize to the unit circle minus the edge in $\mathbb{R}^2$. Explicitly, we want

$$\left\{x \in \mathbb{R}^2 : d(x, 0) < 1\right\}.$$

So now we're using distances instead of order again. So we have the following definition.

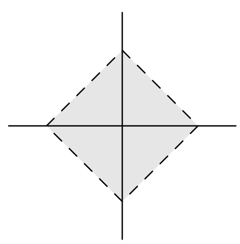
Open sphere

Definition 2.70 (Open sphere). Fix X a metric space, and take $x \in X$ and $r > 0$. Then we define the *open sphere* centered at x of radius r as

$$s_r(x) = \{y \in X : d(x, y) < r\}.$$

We are using "sphere" here because we might want to work in very funny metric spaces.

These can be funny. For example, here is a "sphere" in the taxicab metric of $\mathbb{R}^2$.



And in the tele-metric, the spheres are points or the full space.

Anyways, we now move towards defining and "open" set.

Interior

Definition 2.71 (Interior). Fix X a metric space. Given a subset $S \subseteq X$, we say that $x \in S$ is in the *interior* of S if and only if there exists $r > 0$ such that $s_r(x) \subseteq S$. The set of interior points is notated S°.

And here we are.

Open

Definition 2.72 (Open). Fix X a metric space. Then U is an open set if and only if U is its interior.

Example 2.73. The open interval $(0, 1)$ is open: given any $x \in (0, 1)$, we can set $r := \frac{1}{2}\min\{x, 1 - x\}$ so that $s_x(r) \subseteq (0, 1)$. So each element of $(0, 1)$ is in its interior.

Non-Example 2.74. The interval $[0,1)$ is not open: there is no r such that $s_r(0) \subseteq [0,1)$.

Let's actually prove something.

Proposition 2.75. Fix X a metric space. Then X and $\varnothing$ are both open.

Proof. We do these one at a time.

- For any $x \in X$, we note that
$$s_1(x) = \{y \in X : d(x,y) < 1\} \subseteq X$$
by definition. So all points of X are in the interior of X.

- $\varnothing$ is open because all elements of $\varnothing$ are in the interior (and bananas). ■

We also have the following, which we don't do in class (as they are homework).

Proposition 2.76. An arbitrary union of open sets is open. A finite intersection of open sets is open.

Proof. This is on the homework. We do remark that intersections are finite because, in some sense, we want the "smallest" interior sphere over all of our open sets, but we can only take minimums in finite ses. For example,
$$\bigcap_{k=1}^{\infty}(-1/k, 1/k) = \{0\}$$
is not open. ■

2.5.5 Closed

Now we can define "closed."

Closed **Definition 2.77** (Closed). Fix a metric space X. Then $V \subseteq X$ is *closed* if and only if $X \setminus V$ is open.

Example 2.78. The interval $[0,1]$ is closed because its complement is $(-\infty,0)\cup(1,\infty)$, which is the union of two open intervals and hence is open.

Non-Example 2.79. The interval $[0,1)$ is not open, from earlier. Also, its complement is $(-\infty,0)\cup[1,\infty)$ is also not open because of 1. So $[0,1)$ is neither open nor closed.

While we're here, we note that we can turn around Proposition 2.76.

Proposition 2.80. Any finite union of closed sets is closed. Any arbitrary intersection of closed sets is closed.

Proof. Take the complement of the statements in Proposition 2.76. ■

In light of the above proposition, we have the following.

Closure

Definition 2.81 (Closure)**.** Fix X a metric space and S a subset. Then we define the *closure* of S to be

$$\overline{S} := \bigcap_{S \subseteq V} V,$$

where the intersection is over closed sets V containing S.

Because this is an arbitrary intersection of closed sets, $\overline{S}$ is closed, and we note that $\overline{S}$ will contain any closed set around S, so $\overline{S}$ is in some sense the "smallest" closed set around S.

We have the following result.

Proposition 2.82. A subset $S \subseteq X$ of a metric space X is closed if and only if $\overline{S} = S$.

Proof. We leave this as an exercise. ∎

What happens when there's some discrepancy between the two?

Boundary

Definition 2.83 (Boundary)**.** Fix X a metric space and S a subset. Then we define the *boundary points* of S to be

$$\partial S := \overline{S} \setminus S^\circ.$$

Example 2.84. If $S = (0, 1)$ is closed, $\partial S = \{0, 1\}$.

Why do we care? We have the following.

Proposition 2.85. Fix X a metric space and S a subset. Then any $x \in \overline{S}$ has a sequence of points in S converging to x.

Proof. For any ε, we claim that find $y \in S$ such that $d(x, y) < \varepsilon$. Indeed, we show that every $y \in S$ has $d(x, y) \geq \varepsilon$, then $x \notin \overline{S}$. Consider the ball

$$s_\varepsilon(x) = \{x' \in X : d(x, x') < \varepsilon\}.$$

By hypothesis, $S \cap s_\varepsilon(x) = \varnothing$, so $X \setminus s_\varepsilon(x)$ is a closed set containing S. So

$$\overline{S} \subseteq X \setminus s_\varepsilon(x),$$

which does not contain x. So indeed, $x \notin \overline{S}$.

So for each $n \in \mathbb{N}$, we can find $x_n \in S$ such that

$$d(x, x_n) < 1/n.$$

Then we can see that the sequence $\{x_n\}_{n\in\mathbb{N}}$ converges to x. Indeed, for any $\varepsilon > 0$, we can set $N := 1/\varepsilon$ so that $n > N$ implies

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

which is what we wanted. ∎

2.5.6 Compactness

We start with the following assertion.

Proposition 2.86. Fix X a complete metric space. Fix $V_1 \supseteq V_2 \supseteq \cdots$ a descending sequence of nonempty, bounded, closed sets. Then

$$V := \bigcap_{k=1}^{\infty} V_k$$

is also nonempty, bounded, and closed.

Proof. Closed is by arbitrary intersection. Bounded is by taking a bound of V_1 to bound V.

The meat here is showing that V is nonempty. Well, find some $x_k \in V_k$ for each k. Then $\{x_k\}_{k\in\mathbb{N}}$ lives in a complete metric space and hence has a convergent subsequence, which by abuse of notation we call $\{y_k\}_{k\in\mathbb{N}}$. Say it converges to y.

We claim $y \in V$. The point is that, after long enough, $\{y_k\}_{k\in\mathbb{N}}$ lives completely in any fixed $V_\bullet$, so $y \in V_\bullet$ for any fixed $V_\bullet$. Thus, $y \in V$, finishing. ■

We now define compact and will return to the above shorty.

Compact

Definition 2.87 (Compact)**.** Fix X a metric space and $S \subseteq X$. Then S is *compact* if and only if, for every open cover $\mathcal{U}$ on top of S, there exists a finite subcover $\mathcal{U}_0$ which still covers S.

Non-Example 2.88. We have that $\mathbb{R}$ is not compact. For example, the open cover

$$\{(-n, n) : n \in \mathbb{N}\}$$

has no finite subcover: any finite number will miss sufficiently large real numbers.

Non-Example 2.89. The interval $(0, 1)$ is not compact. For example,

$$\{(1/n, 1) : n \in \mathbb{N}\}$$

has no finite subcover: any finite number will miss sufficiently small real numbers.

Example 2.90. Ihe interval $[0, 1]$ is compact. This is deep, and we will not write out the proof here.

Showing that something is compact is quite difficult: we need to deal with all open covers at once, which is hard to handle. Regardless, there is the following theorem.

Theorem 2.91 (Heine–Borel)**.** In $\mathbb{R}^n$, a subset is compact if and only if it is closed and bounded.

This is amazing! Compactness was very hard to handle, but we have good feelings for what closed and bounded should mean.

2.6 October 4

It's october: that unapparent summer air in early fall.

2.6.1 Infinite Series

Aside from our story into metric spaces, we have been talking about sequences: what can we say about limits, what can we say about going to infinity, etc.

From talking about sequences, we are now going to talk about infinite series. Per usual, we need to start by defining what we mean.

Infinite series

Definition 2.92 (Infinite series)**.** Fix $\{a_k\}_{k\in\mathbb{N}}$ a sequence of real numbers. Define the sequence of partial sums $\{s_n\}_{n\in\mathbb{N}}$ by

$$s_n := \sum_{k=1}^{n} a_k.$$

Then we write the *infinite series*

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} s_n$$

if the s_k have a limit. Note we are allowing the series to equal $\pm\infty$.

Warning 2.93. If the partial sums have no limit, then the series simply does not have a meaning. For example,

$$\sum_{k=1}^{\infty}(-1)^k$$

is a collection of meaningless symbols.

Defined like this, we see that talking about infinite series is quite similar to talking about sequences. The challenges are often the same, but our tools for simple sequences tend to be more robust.

At a high level, the issue here is that the partial sums are more or less defined recursively by

$$s_{k+1} = s_k + a_{k+1},$$

and recursive formulae are difficult to work with. For example, recall how much trouble we had showing that the sequence $\{r_k\}_{k\in\mathbb{N}}$ defined by

$$r_1 = 1 \qquad \text{and} \qquad r_{k+1} = \sqrt{2+r_k}.$$

We essentially had to appeal the fact that monotone, bounded sequences converge, which is much harder than it seems.

Example 2.94. Consider the series

$$\sum_{k=1}^{\infty}\frac{1}{k}.$$

The individual terms $\{1/k\}_{k\in\mathbb{N}}$ are somewhat simple, but trying to find any kind of explicit formula is difficult.

Remark 2.95 (Nir)**.** We can estimate the harmonic series pretty well by

$$\sum_{k=1}^{N}\frac{1}{k} = 0.577 + \log N,$$

which is something but not everything.

2.6.2 Motivation

Let's start with the ratio test.

Proposition 2.96 (Ratio test, I)**.** Fix $\{a_k\}_{k\in\mathbb{N}}$ a sequence of real numbers. Further suppose that we can find

$$L := \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|.$$

Then we have the following.

- If $L<1$, then $\sum a_n$ converges.
- If $L>1$, then $\sum a_n$ diverges.

Note that there is a really big hypothesis in this statement: we need the limit for L to exist, which is potentially problematic.

Example 2.97. For example,

$$a_n = \begin{cases} 2^{-n} & n \text{ is odd}, \\ 3^{-n} & n \text{ is odd}, \end{cases}$$

certainly has $\sum a_n$ convergent (it's bounded above by 2^{-n}), but the consecutive terms are quite bouncy.

To motivate our discussion, let's start with the following example: we define $\{a_n\}_{n\in\mathbb{N}}$ by

$$a_n := \begin{cases} 1/n & n \text{ is odd}, \\ -2^n & n \text{ is even}, \end{cases}$$

which will still cause $\sum a_n$ to diverge because it's mostly $(-2)^n$s, which are huge and negative. Here, both the ratio test and the root test have limits which do not exist, but this very clearly diverges.

We would like to buf our ratio and root tests to accommodate for this. The trick is to not look at $\lim$ but $\limsup$ and $\liminf$ so that we don't have non-existence problems. We will hit here later but not quite now; stay tuned.

2.6.3 Some Tests

Let's start with some basic tests.

Proposition 2.98 (Geometric series)**.** Fix $a,r\in\mathbb{R}$. Then

$$\sum_{k=0}^{\infty} ar^k = \begin{cases} \frac{a}{1-r} & |r|<1, \\ \text{diverges} & |r|\ge 1. \end{cases}$$

Proof. Omitted. The idea is that, for any given N, we have

$$\sum_{k=0}^{N} ar^k = \frac{a\left(1-r^{N+1}\right)}{1-r},$$

from which we can take the explicit formula as $N\to\infty$. ■

Proposition 2.99 (Cauchy criterion)**.** Fix $\{a_n\}_{n\in\mathbb{N}}$. Then the series $\sum a_n$ converges if and only if, for each $\varepsilon > 0$, there exists N such that $n, m > N$ implies

$$|s_n - s_m| < \varepsilon,$$

where $\{s_n\}_{n\in\mathbb{N}}$ is the sequence of

Proof. This is equivalent to asserting that the partial sums form a Cauchy sequence, so they converge because $\mathbb{R}$ is complete. ■

Proposition 2.100 (Divergence test)**.** Fix $\{a_n\}_{n\in\mathbb{N}}$. If $\lim a_n$ is nonzero, then $\sum a_n$ diverges.

Proof. This follows from the Cauchy criterion. Suppose $\sum a_n$ converges to s. Then, for any $\varepsilon > 0$, there exists N such that $n, m > N$ implies

$$|s_n - s_m| < \varepsilon/2.$$

But then ■

This lets us jump into the comparison test.

Proposition 2.101 (Comparison test)**.** Fix $\{a_n\}_{n\in\mathbb{N}}$ a sequence of nonnegative terms. Then suppose that

$$|b_k| \le a_k$$

for each k, eventually.

(a) If $\sum a_n$ converges, then $\sum b_n$ converges.

(b) If $\sum a_n$ diverges to ∞, then $\sum b_n$ diverges to ∞.

Proof. Here we go.

(a) We use the Cauchy criterion. Fix any $\varepsilon > 0$. Then there exists N such that $n > m > N$ implies

$$\left|\sum_{k=m}^{n} a_k\right| = \sum_{k=m}^{n} a_k < \varepsilon.$$

Then

$$\left|\sum_{k=m}^{n} b_k\right| \le \sum_{k=m}^{n} |b_k| \le \sum_{k=m}^{n} a_k < \varepsilon$$

by the triangle inequality, so $\sum b_k$ satisfies the Cauchy criterion as well, so $\sum b_k$ converges.

(b) Omitted. ■

Example 2.102. We know that

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

converges. It follows that

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2}$$

also converges by the Comparison test, though we have no idea what the actual sum is.

Warning 2.103. The above statements are about "eventually" because, in series, we only care about the long-term behavior. Note that this also means we are by design not caring very much about the actual value of a series.

This lets us talk about the following.

Absolute convergence

Definition 2.104 (Absolute convergence)**.** We say that a series $\sum a_n$ is *absolutely convergent* if and only if $\sum |a_n|$ is also convergent.

We can justify the word "convergent" in the above definition because of the following.

Proposition 2.105. Suppose $\sum |a_n|$ is absolutely convergent, then $\sum a_n$ also converges.

Proof. Apply the Comparison test with a_n and $|a_n|$ so that we need to check $|a_n| \leq |a_n|$, which is true. ■

2.6.4 Ratio and Root Tests

And now we move into the harder tests.

Proposition 2.106. Fix $\{a_n\}_{n\in\mathbb{N}}$ a sequence of real numbers, and define $s_n := |a_{n+1}/a_n|$. We have the following.

- If $\limsup s_n < 1$, then the series is absolutely converging.
- If $\liminf s_n > 1$, then the series is diverging.

Proof. Omitted. The idea is to compare with a geometric series. ■

This covers the case where the limit does not exist: no matter what, we will have values to check because $\limsup$ and $\liminf$ always exist. However, it is still quite possible that

$$\liminf s_n < 1 < \limsup s_n,$$

in which case we still get no information.

Non-Example 2.107. Consider

$$a_n := \begin{cases} 1/n & n \text{ is odd,} \\ -2^n & n \text{ is odd.} \end{cases}$$

Then $\liminf$ of the ratios is 0 and $\limsup$ of the ratios is ∞, so the Ratio test does not help us here.

Regardless, it's nice to have the stronger conditions so that we can be sure the problem is with the Ratio test itself and not our application of it.

And here is the buffed version of the Root test.

Proposition 2.108. Fix $\{a_n\}_{n\in\mathbb{N}}$ a sequence of real numbers, and define

$$\alpha := \limsup_{n\to\infty} |a_n|^{1/n}.$$

Then we have the following.

(a) If $\alpha > 1$, then the series $\sum a_n$ diverges.

(b) If $\alpha < 1$, then the series $\sum a_n$ is absolutely converging.

Proof. Again omitted. The idea is still to compare with a geometric series. ■

This is quite nice because there are no gaps as seen between $\liminf$ and $\limsup$ of the Ratio test.

Example 2.109. Consider

$$a_n := \begin{cases} 1/n & n \text{ is odd}, \\ -2^n & n \text{ is odd}. \end{cases}$$

Then $\limsup |a_n|^{1/n} = 2$, so $\sum a_n$ diverges. Finally we have a proof.

2.6.5 Some Examples

Let's do some examples.

Exercise 2.110 (Ross 14.7)**.** Fix $\{a_n\}_{n\in\mathbb{N}}$ is a sequence of nonnegative terms with $\sum a_n$ converging. The, taking $p > 1$, the series $\sum a_n^p$ also converges.

Proof. Use the comparison test. For N large enough, we can say that $n > N$ implies $|a_n| < 1$. Then, after this point,

$$|a_k^p| < |a_k|^p < |a_k|,$$

so $\sum a_n^p$ is eventually smaller than $\sum a_n$. So we are done by the Comparison test. ■

The above exercise is a good example of why we want the Comparison test to have the "eventually." Namely, we had to wait until a point where the terms are less than 1. For example, it is quite possible for $\sum a_n^p$ to have a larger sum:

$$\sum_{k=0}^{\infty} \frac{1}{2^k} = 2 \qquad \text{but} \qquad \sum_{k=0}^{\infty} \left(\frac{1}{2^k}\right)^2 = \frac{13}{3}.$$

Exercise 2.111 (Ross 14.9)**.** Fix $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ are equal aside from finitely many terms. Then $\sum a_n$ converges if and only if $\sum b_n$ converges.

Proof. Use the Cauchy criterion. (Note that we don't use the Comparison test because these need not be nonnegative.) Namely, if $n > N$ has $a_n = b_n$, then $\sum a_n$ satisfies the Cauchy criterion after N if and only if $\sum b_n$ satisfies the Cauchy criterion after N. ■

2.6.6 Integral Test

We have the following.

Proposition 2.112. Fix $\{a_n\}_{n\in\mathbb{N}}$ a decreasing sequence of terms such that $f(n) = a_n$ is a decreasing function. Then $\sum a_n$ has the same convergence/divergence as

$$\int_1^\infty f(x)\,dx.$$

Proof. The book has a nice geometric proof. The main idea is that, because $f(x)$ is decreasing, we have

$$\int_n^{n+1} f(x)\,dx \le f(n) \le \int_{n-1}^{n} f(x)\,dx$$

for $n \in \mathbb{N}$. Summing over $n \in \mathbb{N}$, we find that

$$\int_1^\infty f(x)\,dx \le \sum_{n=1}^\infty f(n) \le f(1) + \int_2^\infty f(x)\,dx$$

which gives the result after some pushing. ■

The prototypical example is the harmonic series.

Example 2.113. The series $\sum 1/n$ diverges because

$$\int_1^\infty \frac{1}{x}\,dx = \log x\Big|_1^\infty = \infty.$$

Example 2.114. The series $\sum 1/n^{1.01}$ converges because

$$\int_1^\infty \frac{1}{x^{1.01}}\,dx = \frac{x^{-0.01}}{-0.01}\Big|_1^\infty = 100 < \infty$$

even though the series is very close to the harmonic series. In fact, the value of the integral is a good first-order approximation of the series.

2.6.7 Logistics

The final exam has the following properties.

- It will be in-class, for 90 minutes.
- It will be six questions. Questions will have one or two parts.
- It will cover the section that we have had homework over by the time the exam occurs. Namely, we will cover Chapters 1 and 2.
- It will be completely closed-book.
- Be prepared.

2.7 October 6

We're finishing up §15 and §16 today.

2.7.1 Integral Test

We've probably seen the integral test in other classes. Let's start with the following example.

Proposition 2.115 (p-series test)**.** The series

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

if and only if $p > 1$.

Proof. This is done by the integral test. Essentially,

$$\int_2^\infty \frac{1}{x^p}\,dx$$

converges if and only if $p > 1$. ■

Remark 2.116. This result is quite important because they provide nice series to compare to. For example, we have some amount of control over any polynomials now.

2.7.2 Alternating Series Test

We start by defining our sequences.

Alternating

Definition 2.117 (Alternating)**.** A sequence $\{a_k\}_{k\in\mathbb{N}}$ is *alternating* if and only if a_{k+1} has a different sign from a_k for each $k \in \mathbb{N}$.

Warning 2.118. Merely having a mix of positive and negative terms is not enough to be alternating. It must alternate at every integer.

And here is our test.

Proposition 2.119 (Alternating series test)**.** Suppose $\{a_k\}_{k\in\mathbb{N}}$ is a sequence of nonnegative terms such that

- $a_k \to 0$ as $k \to \infty$, and
- $\{a_k\}_{k\in\mathbb{N}}$ is decreasing.

Then

$$\sum_{k=1}^{\infty}(-1)^{k-1}a_k$$

converges.

Note that this is almost as strong as we could want: surely $a_k \to 0$ as $k \to \infty$ is necessary for convergence, and all extra we have asked for is decreasing. In exchange for this strength, we are

Proof. We use the Cauchy criterion. Fix any $\varepsilon > 0$. Now, there is some N for which $k > N$ implies $a_k < \varepsilon$ because $a_k \to 0$. Then we claim that $n > m > N$ implies

$$\left|\sum_{k=m+1}^{n}(-1)^{k-1}a_k\right| < \varepsilon.$$

The point is that an induction can show

$$\sum_{k=m+2}^{n}(-1)^{k-m-2}a_k > 0$$

for any n, so

$$\left|\sum_{k=m+1}^{n}(-1)^{k-1}a_k\right| \le |a_n| < \varepsilon,$$

which is what we needed. ■

Remark 2.120. The Cauchy criterion is nice here because it provides a somewhat general machinery for how we could prove similar statements. For example, we could imagine using the same machine to show

$$\sum_{k=1}^{\infty} \sin\left(\frac{2\pi k}{10}\right) a_k$$

converges when $\{a_k\}_{k\in\mathbb{N}}$ is a nonnegative, decreasing sequence of real numbers which goes to 0.

As an example, note that the alternating harmonic series

$$\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}$$

converges by the Alternating series test. However, it is possible to rearrange the terms around to make this diverge to $+\infty$ (or anything in $\mathbb{R}\cup\{\pm\infty\}$ for that matter). For example, write positive terms $\frac{1}{2k+1}$ until we are above 1, then write down $-1/2$. Then continue writing down positive terms until we are above 2, and write down $-1/3$. Then continue in this matter.

The problem here is that the alternating harmonic series is conditionally convergent.

Conditionally convergent

Definition 2.121 (Conditionally convergent)**.** Fix $\{a_k\}_{k\in\mathbb{N}}$ a sequence of real numbers. If $\sum a_n$ converges but $\sum |a_n|$ does not converge, then $\sum a_n$ is called *conditionally convergent*.

It is a fact that rearranging the terms in a conditionally convergent series can be rearranged to make the sum whatever we want. In contrast, absolutely convergent series do not have this problem: any rearrangement is safe.

Anyways, let's do an exercise.

Exercise 2.122. The series

$$\sum_{k=2}^{\infty} \frac{1}{k(\ln k)^p}$$

will converge if and only if $p > 1$.

Proof. We use the integral test because it can cover most p at once. Comparison test might seem viable, but because we need to cover the entire real spectrum of p, we have to be careful. Set $f(x) = \frac{1}{x(\ln x)^p}$ which is decreasing and hence safe.

We can compute, for $p \neq 1$,

$$\int_2^{\infty} \frac{dx}{x(\ln x)^p} = \int_{\ln 2}^{\infty} \frac{du}{u^p} = \frac{u^{-p+1}}{-p+1}\bigg|_{\ln 2}^{\infty}.$$

If $p > 1$, then the exponent is negative, so the series converges. If $p < 1$, then the exponent is positive, so the series diverges. And for $p = 1$, we have

$$\int_{\ln 2}^{\infty} \frac{du}{u} = \ln u\bigg|_2^{\infty} = +\infty,$$

so we diverge again. ■

Exercise 2.123. Suppose $\{a_k\}_{k\in\mathbb{N}}$ is a decreasing sequence with $\sum a_k$ convergent. Then $\lim na_n = 0$.

Proof. Note that some care is required: doing this by contradiction is difficult because the $\lim na_n \neq 0$ is not very helpful.

We have $a_k \geq 0$ for each k because otherwise $\lim a_k$ will be less than 0 or nonexistent, breaking the divergence test. Now, by the Cauchy criterion, for each $\varepsilon > 0$, we can find N such that $n > m > N$ implies

$$\sum_{k=m+1}^{n} a_k < \varepsilon.$$

The point here is that all of these terms are at least a_n because the $a_\bullet$ decrease, so $(n-m)a_n < \varepsilon$. For example, we may take $n = 2m$ so that $ma_{2m} < \varepsilon$, which means $(2m)a_{2m} < \varepsilon$. So this shows that $(2m)a_{2m} \to 0$, which is good enough because the $a_\bullet$ are decreasing. ■

2.7.3 Talking Reals

Let's talk about §16; it's nice but fairly irrelevant. Namely, it is perhaps not worth studying closely for the purposes of the class.

We are building the real numbers. Let's recall some ways to define the rational numbers.

Definition 2.124. A rational number is the ratio of two integers.

Definition 2.125. A rational number is an eventually repeating decimal expansion.

Then we can define our real numbers as decimal expansions, using the idea of extending the second definition. However, this is somewhat subtle: how do we show that this expansion is unique? Well, the answer is that this is false:

$$0.0\overline{9} = 0.10.$$

To rigorize this, we note that

$$0.d_1d_2\ldots = \sum_{k=1}^{\infty} \frac{d_k}{10^k}.$$

Then if two decimal expansions $0.d_1\ldots$ and $0.e_1\ldots$ give the same real number, then we have

$$\sum_{k=1}^{\infty} \frac{d_k - e_k}{10^k} = 0.$$

If these aren't identically equal, then say that the differ first at N. Then

$$\frac{e_N - d_N}{10^N} = \sum_{k>N} \frac{d_k - e_k}{10^k},$$

but this is very restrictive: the sum on the right is at most $\sum_{k>N} \frac{9}{10^k} = \frac{1}{10^N}$ to begin with, so the only way for this to occur is if $e_N = d_N + 1$ (where $e_N > d_N$) without loss of generality, and $d_k - e_k = 0$ for $k > N$.

We can also show that rational numbers have eventually repeating decimal expansions. For example, suppose we are looking at

$$\alpha = 0.d_1\ldots d_{k-1}\overline{d_k\ldots d_n}.$$

Then

$$\beta := 10^{k-1}\alpha = 0.\overline{d_k\ldots d_n}.$$

However, $10^{n-k}\beta - \beta$ is equal to $d_k\ldots d_n$ as an integer, so β is a rational, so α is a rational.

In the other direction, suppose we have a ratio of integers a/b, and we take $\gcd(b, 10) = 1$, for otherwise we can shift over the decimal expansion by multiplying the fraction by a sufficiently large power of 10. Now study the sequence of remainders

$$10^k \pmod b$$

for $k \in \mathbb{Z}$. This must repeat eventually because there are infinitely many integers and finitely many residues, so find some $k > \ell$ for which $10^k \equiv 10^\ell \pmod b$ so that $n := k - \ell$ has $10^n \equiv 1 \pmod b$. But then

$$\frac{a}{b} = \frac{a \cdot \frac{10^n - 1}{b}}{10^n -}$$

is a repeating decimal.

And here is an exercise to round our our discussion.

Exercise 2.126 (Ross 16.13)**.** Suppose that $\sum a_k$ and $\sum b_k$ both converge with $a_k \le b_k$ for each k and $a_k < b_k$ for at least one k. Then

$$\sum_{k=1}^\infty a_k < \sum_{k=1}^\infty b_k.$$

Proof. We show a kind of contrapositive. Suppose that $a_k \le b_k$ and

$$L := \sum_{k=1}^\infty a_k = \sum_{k=1}^\infty b_k.$$

Combining the convergences, we see that, for any $\varepsilon > 0$, there exists some N such that $n > N$ implies

$$\left| L - \sum_{k=1}^n a_k \right|, \left| L - \sum_{k=1}^n b_k \right| < \varepsilon,$$

which gives

$$\left| \sum_{k=1}^n (b_k - a_k) \right| < 2\varepsilon$$

by the triangle inequality. But $b_k \ge a_k$ for each k, so we see

$$\sum_{k=1}^n (b_k - a_k) < 2\varepsilon.$$

Now, for any particular $m \in \mathbb{N}$, we note that, for any $\varepsilon > 0$, we have some N such that $n > \max\{m, N\}$ implies

$$0 \le (b_m - a_m) \le \sum_{k=1}^n (b_k - a_k) < 2\varepsilon.$$

So we must have $b_m = a_m$ because their difference is smaller than any positive real number 2ε. ∎

THEME 3: CONTINUITY

3.1 October 13

So a midterm ocurred. I think it went ok.

3.1.1 Midterm Notes

Here are some general comments about the content.

- On the first question, many people tried to have $\lim s_n \neq -\infty$ imply $\lim s_n \in \mathbb{R} \cup \{+\infty\}$, but this is false. Namely, in all interesting cases the limit should not exist at all.

- The easiest way to solve the first problem is to note

$$-\infty = \limsup s_n \geq \liminf s_n$$

  and finish because $\liminf s_n = \limsup s_n = -\infty$. I didn't know if this was something we could cite.

- We have to be somewhat careful with what $\lim s_n \neq -\infty$ means. It tells us that there exists some M for which no N has $n > N$ implies $s_n > M$. This is not the same thing as saying there exists some M for which $s_n > M$ for each n.

- On the second question, we have to be careful with what "A is not open" means. Pushing quantifiers through, all we get is that there is some $a \in A$ for which all $r > 0$ has

$$\{x \in X : d(a,x) < r\} \not\subseteq A.$$

  Here, $\not\subseteq$ is an annoying symbol to work with.

- The class has the hardest time with number 2.

- For the third question, some people wrote down

$$\int_1^\infty \frac{1}{x^x}\,dx = \sum_{n=1}^\infty \frac{1}{x^x},$$

  which is untrue. What we know is that they converge together or diverge together.

- For the fourth question, the class did best. There was some discussion about "solving" for the inductive step.

- For the fifth question, the best solution is to make one pair of distinct points have distance 1.

- For the last question, people again did well.

Here are some notes on statistics.

- Class average was about 96 points out of 120, which is about 80%.

- In particular, the curve will likely be minor.

3.1.2 Continuity, Advertisement

As a warning, metric spaces, which people seem to understand only mediocrely, will appear in the future though not in a major sense. For example, we will be talking about continuity for a little while, and perhaps this should change along with our metric. So as our metric changes, we might want to be careful with how our notion of continuity changes.

Anyways, we will work with the normal distance metric on $\mathbb{R}$. Let's imagine we're trying to graph some function. When things are continuous, we can just guess a few points and connect the dots. For discontinuities, these tend to stand out in the graph: they might look like jumps or asymptotes or oscillations or similar.

To test for continuity, here is our definition.

Continuity, I

Definition 3.1 (Continuity, I). Fix a function $f : S \to \mathbb{R}$ for some $S \subseteq \mathbb{R}$. Then we say that f is *continuous at a* if and only if

$$\lim_{x\to a} f(x) = f(a).$$

We will skirt around what $\lim$ means for real numbers. In practice most of our examples are for piecewise functions or just actually continuous, so these limits are fine to evaluate. Namely, for most functions we care about—polynomials, $\exp, \sin,$ etc.—are all continuous.

But sometimes life is not so good.

Example 3.2. The function

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

is continuous at no $a \in \mathbb{R}$.

We will be able to prove this shortly. But the point here is that proving this is somewhat obnoxious without rigorous definitions of our terms.

3.1.3 Continuity, Rigorously

So here is our real definition of continuity.

Continuity, II

Definition 3.3 (Continuity, II). We say that a function $f : \mathbb{R} \to \mathbb{R}$ is *continuous at $x = a$* if and only if, for each $\varepsilon > 0,$ there exists $\delta > 0$ such that

$$|x - a| < \partial \implies |f(x) - f(a)| < \varepsilon.$$

Geometrically, are imagining that we have a given error bound of ε and an open interval $(f(a) - \varepsilon, f(a) + \varepsilon)$. Then we want to find some small open interval given by δ for which $(a - \delta, a + \delta)$ will go into the error bound interval.

As an aside, this definition is potentially more annoying to work with than it was for convergence in sequences, but this is because we are now upgrading to reals. Things will get harder; so it goes.

Here is one way to access continuity.

Proposition 3.4. Suppose that $f(x)$ is continuous at $x = a$. Then a sequence $\{a_k\}_{k\in\mathbb{N}}$ converging to a will have $\{f(a_k)\}_{k\in\mathbb{N}}$ converge to a.

Proof. Fix any $\varepsilon > 0$. By continuity of f at a, there is some $\delta > 0$ for which $|x-a| < \delta$ implies $|f(x)-f(a)| < \varepsilon$. But because $a_\bullet \to a,$ there exists N for which

$$n > N \implies |a_n - a| < \delta \implies |f(a_n) - f(a)| < \varepsilon.$$

So we see that $f(a_n) \to f(a)$. ∎

In fact, the converse is also true.

Proposition 3.5. Fix $a \in \mathbb{R}$. Suppose that $f : \mathbb{R} \to \mathbb{R}$ has the property that, for each sequence $\{a_k\}_{k\in\mathbb{N}}$ with $a_k \to a$, we have $f(a_k) \to a$. Then f is continuous at a.

Proof. We proceed by contraposition. Suppose that f is not continuous at a, and we will exhibit a sequence $a_k \to a$ for which $f(a_k)$ does not converge to $f(a)$. We know there exists $\varepsilon > 0$ for which no δ has

$$|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon.$$

In particular, for each $n \in \mathbb{N}$, we can find some a_n for which $|a_n - a| < \frac{1}{n}$ while $|f(a_n) - f(a)| \geq \varepsilon$.

We claim that this is the sequence that we want. Indeed, we have that $a_n \to a$ because, for any $\varepsilon_0 > 0$, we can set $N := 1/\varepsilon_0$ so that $n > N$ implies

$$|a_n - a| < \frac{1}{n} < \frac{1}{N} = \varepsilon_0.$$

However, $f(a_n)$ does not converge to $f(a)$. Indeed, for the given $\varepsilon > 0$, we know there is no N for which $n > N$ implies $|f(a_n) - f(a)| < \varepsilon$ because $|f(a_n) - f(a)| \geq \varepsilon$ for each $n \in \mathbb{N}$. ■

So Proposition 3.5 tells us that we can reduce study of continuity to study of sequences, but there is still some amount of cumbersome work because we would have to account for all such sequences; doing single sequences is not enough.

Example 3.6. The function $\sin\left(\frac{1}{x}\right)$ is not continuous at $x = 0$, but the sequence $\left\{\frac{1}{2\pi k}\right\}_{k\in\mathbb{N}}$ has values of only 0s, which do converge.

We remark that the definition of continuity we are working with focuses on differences of absolute values, which is really just the standard metric in disguise. So let's try changing the metric.

Example 3.7. In the telemetric (distances between distinct points is always 1), sequences converges if and only if it is eventually constant. Namely, if $a_k \to a$, then take $\varepsilon = \frac{1}{2}$ so that there is some N for which $n > N$ has

$$d_{\text{tele}}(a, a_k) < \frac{1}{2},$$

but this forces $a_k = a$. And conversely, if a sequence is eventually constant, then of course it converges.

But now, all functions are continuous everywhere! Indeed, for any sequence $a_k \to a$, we have $f(a_k)$ will eventually be exclusively terms of $f(a)$, which converges to $f(a)$, as needed.

Things get more complicated, for example, if the domain and range have different metrics and notions of continuity.

3.2 October 18

We continue with our discussion on continuity.

3.2.1 Talking Continuity

People usually start with a graphical definition of continuity, but we will want something more general. For example, graphical intuition does not really help define functions which are continuous on $\mathbb{R} \setminus \mathbb{Q}$ but not in $\mathbb{Q}$.

We have given a precise definition of continuity, so we can start talking about some continuity laws. For example, is the sum of two continuous functions itself continuous?

Proposition 3.8. Fix functions $f, g : S \to \mathbb{R}$ are continuous at some $a \in S$. Then $f + g$ is continuous at a.

Proof. Fix any $\varepsilon > 0$. Then $\varepsilon/2 > 0$, so we have some $\delta_f > 0$ such that

$$|x - a| < \delta_f \implies |f(x) - f(a)| < \frac{1}{2}\varepsilon.$$

Similarly, we have some $\delta_g > 0$ such that

$$|x - a| < \delta_g \implies |g(x) - g(a)| < \frac{1}{2}\varepsilon.$$

But now set $\delta := \min\{\delta_f, \delta_g\} > 0$ so that $\delta > 0$ has $|x - a| < \delta$ implies both of the above, so we bound

$$|(f + g)(x) - (f + g)(a)| \le |f(x) - f(a)| + |g(x) - g(a)| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon,$$

which is what we need. ■

Note that we could have just proven the corresponding limit law here using the sequence definition instead and used the fact that we know limits of sequences decompose.

And we can also do $f \cdot g$.

Proposition 3.9. Fix functions $f, g : S \to \mathbb{R}$ are continuous at some $a \in S$. Then fg is continuous at a.

Proof. The key point is that

$$|f(x)g(x)-f(a)g(a)| \le |f(x)g(x)-f(x)g(a)|+|f(x)g(a)-f(a)g(a)| = |f(x)|\cdot|g(x)-g(a)|+|g(a)|\cdot|f(x)-f(a)|.$$

Now we can bound this quantity; fix any $\varepsilon > 0$. Something annoying here is that we need to bound the $|f(x)|$ on the outside. So start by fixing some ε_0 to get δ_0 so that

$$|x - a| < \delta_0 \implies |f(x) - f(a)| < \varepsilon_0.$$

In particular, $|f(x)| < |f(a)| + \varepsilon_0$ always, so we see that $|x - a| < \delta_0$ implies

$$|f(x)g(x) - f(a)g(a)| \le ||f(a)| + \varepsilon_0| \cdot |g(x) - g(a)| + |g(a)| \cdot |f(x) - f(a)|.$$

Now bounding $|f(x) - f(a)|$ and $|g(x) - g(a)|$ we can bound without tricks. Now we can find $\delta_f > 0$ such that

$$|x - a| < \delta_f \implies |f(x) - f(a)| < \frac{\varepsilon/2}{|g(a)| + 1},$$

and we can find some $\delta_g > 0$ such that

$$|x - a| < \delta_g \implies |g(x) - g(a)| < \frac{\varepsilon/2}{||f(a)| + \varepsilon_0| + 1}.$$

In particular, $\delta := \min\{\delta_0, \delta_f, \delta_g\} > 0$ will have $|x - a| < \delta$ imply all of the above so that

$$|f(x)g(x) - f(a)g(a)| \le ||f(a)| + \varepsilon_0| \cdot |g(x) - g(a)| + |g(a)| \cdot |f(x) - f(a)| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon < \varepsilon,$$

which is what we wanted. ■

Let's do some exercises.

Exercise 3.10 (Ross 17.5)**.** Fix some m a positive integer. We show that $f(x) = x^m$ is continuous for each real number x.

Proof. Because m is a positive integer, we can do this by induction. Our base case is $m = 1$, where we say that each $a \in \mathbb{R}$ with $\varepsilon > 0$ has $\delta = \varepsilon$ giving

$$|x - a| < \varepsilon \implies |f(x) - f(a)| = |x - a| < \delta = \varepsilon,$$

establishing our continuity. Now, we know that products of continuous functions are continuous, so

$$\underbrace{x \cdot \ldots \cdot x}_{m}$$

is also continuous. ■

Here is a harder example.

Exercise 3.11 (Ross 17.9(a))**.** We show that $f(x) = x^2$ is continuous at $a = 2$.

Proof. We use the ε-δ definition of continuity. Namely, for each $\varepsilon > 0$, we need to find $\delta > 0$ so that

$$|x - 2| < \delta \overset{?}{\implies} \left|x^2 - 4\right| < \varepsilon.$$

But now we see that

$$\left|x^2 - 4\right| = |x - 2| \cdot |x + 2|.$$

Provided $\delta \le 1$, then we see that $|x - 2| < \delta$ implies $|x + 2| \le |x - 2| + 4 \le 1 + 4 = 5$. So we set $\delta := \min\{1, \varepsilon/5\} > 0$ so that $|x - 2| < \delta$ implies

$$\left|x^2 - 4\right| = |x - 2| \cdot |x + 2| < \delta \cdot 5 \le \varepsilon,$$

which is what we wanted. ■

Notably, we don't care for the exact value of δ that will give the sharpest inequality. We need something to work; constant factors are inconsequential.

Exercise 3.12 (Ross 17.9(b))**.** We show that $f(x) = \sqrt{x}$ is continuous at $a = 0$.

Proof. For each $\varepsilon > 0$, we need to find δ such that

$$|x - 0| < \delta \overset{?}{\implies} \left|\sqrt{x}\right| < \varepsilon.$$

Well, we can just set $\delta = \varepsilon^2$ so that $|x| < \delta$ implies $|\sqrt{x}| < \varepsilon$, which we can show by contraposition: $|\sqrt{x}| \ge \varepsilon$ would imply that $|x| \ge \varepsilon^2$ by multiplication. ■

3.2.2 Extreme Value Theorem

So we're actually starting to do some analysis in this class. We recall the following statement.

Theorem 3.13 (Extreme value)**.** Fix f a continuous function on a closed interval. Then f is bounded and achieves a maximum and minimum.

Warning 3.14. We have that $\mathbb{R} = (-\infty, \infty)$ is closed and an interval, but it is not a closed interval because it does not contain its endpoints.

Proof. We start by showing that f is bounded. We will show that f is bounded above, and bounded below is a similar argument; suppose for the sake of contradiction that f is not bounded above, so that we can find a sequence $\{f(x_k)\}_{k\in\mathbb{N}}$ so that $f(x_k) \geq k$ for each $k \in \mathbb{N}$. However, the sequence $\{x_k\}_{k\in\mathbb{N}}$ is a sequence in a bounded set, so it must have a convergent subsequence $\{x_{\sigma k}\}_{k\in\mathbb{N}}$ given by $\sigma : \mathbb{N} \to \mathbb{N}$.

So now say that $x_{\sigma k} \to x$, and we know that x is in the domain of f because the domain of f is closed (!). But now $f(x_{\sigma k}) \to f(x)$, so there is some N for which $n > N$ implies $|f(x_{\sigma n}) - f(x)| < 1$ so that

$$f(x) \geq f(x_{\sigma n}) - 1 \geq \sigma k - 1 \geq k - 1$$

for each $f(x)$. But now we can set $x = \lfloor f(x) \rfloor + 10$ to get a contradiction, finishing this.

Now we show that f achieves its maximum, and the minimum is a similar argument. Fix

$$s := \sup\{f(x)\}.$$

Now, the supremum is in the set of subsequential limits, which means we can find $\{x_k\}_{k\in\mathbb{N}}$ such that $f(x_k) \to s$. But now the $\{x_k\}_{k\in\mathbb{N}}$ has a convergent subsequence called $\{x_{\sigma k}\}_{k\in\mathbb{N}}$ for some strictly increasing $\sigma : \mathbb{N} \to \mathbb{N}$.

To finish, we again see that $x_{\sigma k} \to x$ for some x in the domain of f because the domain of f is closed. But now $f(x_{\sigma k}) \to f(x)$, and because $f(x_{\sigma k}) \to s$ because the limit of the subsequence is the limit of the sequence when the limit of the sequence is defined. Thus, we have $f(x) = s$ because the limit is unique. ■

This is an important idea for calculus when we're trying to find the maximums and minimums of various functions. Namely, it is not immediately obvious—as the above theorem shows—that these actually exist at all. Here we see that we get this for all closed intervals (in fact, compact domains in general), but, say, open intervals are harder.

Remark 3.15. The above proof also works for any closed and bounded set. So, for example, we have an extreme value theorem on $[0, 1] \cup [2, 3]$.

We also note that we have some of the tools and thought processes for more general settings. For example, what if instead of $\mathbb{R}$ as our ambient space, we use $\mathbb{Q}$? Well, the function

$$f(x) = \frac{1}{x - \sqrt{2}}$$

on the closed interval $[0, 5]$ is not bounded above or below even though it is perfectly well-formed.

As an aside, we note that closed sets in metric spaces and topologies in general have important implications. Namely, being closed was paramount to the above discussion: everything breaks down on open intervals because we lost closure. In fact, compactness is an important condition here.

Example 3.16. All functions on $\{1/n\}_{n\in\mathbb{N}}$ are continuous. Roughly speaking, this is due to "discreteness."

In spite of the above example, it really feels like there should be a constraint on 0, but functions on $\{1/n\}_{n\in\mathbb{N}}$ cannot "see" 0.

Example 3.17. Not all functions on $\{1/n\}_{n\in\mathbb{N}} \cup \{0\}$ are continuous. Namely, we also have to check that the subsequence $1/n \to 0$ is carried to $f(1/n) \to f(0)$.

3.2.3 Intermediate Value Theorem

Here is another important result. For example, we used this quite frequently in calculus when doing root-finding.

Theorem 3.18 (Intermediate value)**.** Fix f a continuous function on the closed interval $[a, b]$. Then if f achieves all values between $f(a)$ and $f(b)$.

Proof. We proceed by contradiction. Without loss of generality, we take $f(a) < f(b)$. Suppose that d between $f(a)$ and $f(b)$ goes is not achieved. However, the idea now is that we can decompose $[a, b]$ into two disjoint open sets, which is a problem because $[a, b]$ is connected. Indeed, we look at

$$S := \{x \in [a, b] : f(x) < d\}.$$

Now, there must be a sequence $\{x_k\}_{k\in\mathbb{N}}$ such that $x_k \to \sup S$, and so it follows $\sup S \in [a, b]$ by closure. But now each $f(x_k)$ is at most d, so $f(\sup S) \le d$.

We very quickly note that $\sup S \notin \{a, b\}$. Indeed, $f(a) < d < f(b)$, so certainly $f(\sup S) \le d < f(b)$ implies $\sup S \ne b$. On the other hand, $\sup S = a$ would force $S = \{a\}$, but then $x > a$ would have $f(x) \ge d > f(a)$, so $|f(x) - f(a)| \ge d$, which breaks continuity of f at a by checking $\varepsilon = d - f(a)$.

To finish, we note that the complement

$$S^c = \{x \in [a, b] : f(x) > d\}$$

will have a sequence converging down to $\sup S$ because all points bigger than $\sup S$ are in S^c, and such points exist because $\sup S \notin \{a, b\}$. Thus, this downward moving subsequence shows that $f(\sup S) \ge d$, so $f(\sup S) = d$, which is our contradiction as we have a point which goes to d. ■

We note that the Intermediate value theorem implies that the output of our continuous function f on the closed interval must itself be a closed interval.

Corollary 3.19. Suppose that f is a continuous function on the closed interval $[a, b]$. Then the image of f is a closed interval.

Proof. We know that f must achieve its maximum M and minimum m from the Extreme value theorem, say from x_M and x_m respectively. But now the Intermediate value theorem implies that

$$f([a, b]) \supseteq f([x_M, x_m] \cup [x_m, x_M]) \supseteq [m, M],$$

but$f([a, b]) \subseteq [m, M]$ by definition of the maximum and minimum, so we conclude $f([a, b]) = [m, M]$, which finishes. ■

To continue our story, we pick up the following definition.

Strictly increasing

Definition 3.20 (Strictly increasing)**.** We say that $f(x)$ is *strictly increasing* if and only if $a < b$ implies $f(a) < f(b)$.

This is the same idea we had as in sequences.

We have the following result, acting as a partial converse to the Intermediate value theorem.

Proposition 3.21. Fix f a strictly increasing function on some interval. Then if the image of f is also an interval, then f is continuous.

Proof. Fix $f : I_1 \to I_2$ our function, where f surjects. If I_2 is a point, then f is constant and hence continuous. Else points in I have nontrivial open neighborhoods.

Now, fix $a \in I_1$ that we want f to be continuous at. We proceed by force. Choose some $\varepsilon > 0$, and then we observe that

$$(f(a) - \varepsilon, f(a) + \varepsilon) \cap I_2$$

is some interval because it is the intersection of two intervals, and this interval is nonempty because it contains $f(a)$. So let this interval have $[y_1, y_2]$ as a closed sub-interval, and then we can pull y_1 and y_2 back to x_1 and x_2 by f. Now finding some open neighborhood around a inside of (x_1, x_2) gives us our δ so that

$$|x - a| < \delta \implies x \in (x_1, x_2) \implies f(x) \in (f(a) - \varepsilon, f(a) + \varepsilon).$$

This finishes the proof. ■

This helps us test for function continuity, for example by partitioning a big function into small intervals where it is strictly increasing or decreasing because we know that the function ought be continuous on each of those intervals by simply checking what their output through f is.

3.3 October 20

Here we go again.

3.3.1 More on Monotonic Functions

Last time we showed that if a strictly increasing function has image an interval on an interval, then the function is continuous on the domain interval. Of course, the use here is somewhat restricted because because strictly increasing or decreasing is quite restricted.

By doing some trick, however, sometimes we can get out of a lot of work. For example, if a function is *locally monotone*, we might be able to piece together continuity for the full function.

Example 3.22. We show that $f(x) = 1/x$ is continuous on $[1, 2]$. Well, the function is strictly decreasing, and it is not too hard to show that the range is $[1/2, 1]$, so the function is continuous.

Continuing our discussion, we have the following definition.

One-to-one

Definition 3.23 (One-to-one)**.** A function f is *one-to-one* if and only if $f(a) = f(b)$ implies $a = b$.

Non-Example 3.24. The function $f(x) = x^2$ is not one-to-one on $\mathbb{R}$. For example, $f(2) = f(-2)$.

It feels like functions which are one-to-one and continuous should be strictly increasing or decreasing. Let's show this.

Proposition 3.25. Suppose $f : I \to \mathbb{R}$ is continuous and injective, where I is some interval. Then f is strictly increasing or strictly decreasing.

Proof. Show that if f is continuous and neither strictly increasing nor strictly decreasing, then f is not injective. Indeed, to not be strictly increasing nor decreasing, there must be $a < b$ and $c < d$ with $f(a) < f(b)$ and $f(c) < f(d)$.

By picking up a suitable subset of our three elements, we can say that either $f(a) \geq f(b) \leq f(c)$ or $f(a) \geq f(b) \leq f(c)$ for $a < b < c$. If any of these are equalities, we are done already. So without loss of generality take $f(a) < f(b) > f(c)$, but then f must hit each value in $(\max\{f(a), f(c)\}, f(b))$ twice, once in (a, b) and once more in (b, c). So this violates injectivity. ■

Let's do some exercises.

Exercise 3.26. We show that $xe^x = 2$ for some $x \in (0, 1)$.

Proof. This is by the Intermediate value theorem. Note that xe^x is continuous on $(0, 1)$ because it is the product of two continuous functions. Then it suffices to note that

$$0 = 0e^0 < 2 < 1e^1 = e.$$

So the Intermediate value theorem finishes. ■

Exercise 3.27. Show that any polynomial of odd degree will have a root.

Proof. This is somewhat technical. The main idea is that $x \to \infty$ makes $f(x) \to \pm\infty$ and $x \to -\infty$ makes $f(x) \to \mp\infty$. So there is some a with $f(a) > 0$ and some b with $f(b) < 0$, and so there is some x between them with $f(x) = 0$ by the Intermediate value theorem.

The technicalities are somewhat annoying, so we will not give them. Roughly speaking, we would have to fix

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

where all but finitely many of the $a_\bullet$ are zero. Then we can find some positive and some negative outputs by hand. ■

The exercises in this section should not be too challenging.

3.3.2 Uniform Continuity

This is one of our first, new analysis concepts. Here is the motivating example.

Exercise 3.28. We show that $f(x) = x^2$ is continuous at $x = 2$.

Proof. For each $\varepsilon > 0$, we showed last time that $\delta := \frac{1}{5}\varepsilon$ was good enough. ■

Something funny about this proof is that the constant $\frac{1}{5}$ was generated off of $x = 2$. For example, with $\varepsilon = 1$, we can choose $\delta = \frac{1}{5}$ so that

$$|x - 2| < \frac{1}{5} \implies \left|x^2 - 4\right| < 1.$$

But now if we chose 100 as our point of interest, then $\delta = \frac{1}{5}$ will no longer work here: $\left(100 + \frac{1}{10}\right)^2 > 100 + 20 > 100 + 1$, so we are out of luck.

The issue we are running into here is that δ is highly dependent on our choice of point we are studying. It does not feel like we can make δ independent of this, no matter how small it goes.

Proposition 3.29. Fix $f(x) = x^2$ and $\varepsilon = 1$. Then there does not exist δ such that, for each $a \in \mathbb{R}$ and $x \in \mathbb{R}$,

$$|x - a| < \delta \implies \left|x^2 - a^2\right| < \varepsilon.$$

Proof. Suppose for the sake of contradiction that there is such a δ. Then, for any $x, y \in \mathbb{R}$, we have that

$$|x - y| < \delta \implies \left|x^2 - y^2\right| < \delta.$$

But now take $x = 2/\delta$ and $y = 2/\delta + \delta/2$. Then these have distance $\delta/2 < \delta$, but

$$y^2 - x^2 = \left(\frac{2}{\delta} + \frac{\delta}{2}\right)^2 - \frac{\delta^2}{4} = 1 + \frac{4}{\delta^2} > 1 = \varepsilon,$$

so we have hit our contradiction by the hypothesis on δ. ■

So we have the following definition.

Definition 3.30. Fix $f : S \to \mathbb{R}$ a function. Then we say that f is *uniformly continuous* on S if and only if, for each $\varepsilon > 0$, we have that there exists a single $\delta > 0$ such that

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

Example 3.31. Constant functions are uniformly continuous. For each $\varepsilon > 0$, take $\delta = 1$ so that

$$|x - y| < \delta \implies |f(x) - f(y)| = 0 < \varepsilon.$$

Example 3.32. The identity function is uniformly continuous. For each $\varepsilon > 0$, take $\delta = \varepsilon$ so that

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

Non-Example 3.33. The function $f(x) = x^2$ is not uniformly continuous, as we showed above.

Example 3.34. The function $f(x) = e^x$ seems to have the same problems due to it moving at increasing speeds.

Non-Example 3.35. The function $\sin\left(x^2\right)$ is bounded and continuous, but it is not uniformly continuous.

We note that this is a stronger condition than continuity.

Proposition 3.36. If $f : S \to \mathbb{R}$ is uniformly continuous, then f is continuous on all of S.

Proof. Fix any $a \in S$. Then for any $\varepsilon > 0$, we know there is a δ such that

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

Plugging in $y = a$ shows that

$$|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon,$$

so it follows that f is continuous at a. ■

But of course, we showed that the reverse implication is untrue with $f(x) = x^2$.

3.3.3 Properties of Uniform Continuity

We would like to have a more concrete way to test for uniform continuity because the given definition has quite a few quantifiers to digest. We have the following example.

Exercise 3.37. We show that $f(x) = x^2$ is uniformly continuous on $[-4, 10]$.

Proof. Here we are safe because our "speed" is now bounded, and in fact the condition seems to be the worst at 10. Fix any $\varepsilon > 0$. Then we need to find δ so that

$$|x - y| < \delta \overset{?}{\implies} \left|x^2 - y^2\right| < \varepsilon.$$

Now, we note that

$$\left|x^2-y^2\right| = |x-y|\cdot|x+y| \le |x-y|\cdot(|x|+|y|) \le |x-y|\cdot 20,$$

so we may choose $\delta := \varepsilon/20$ which gives

$$|x-y|<\delta \implies \left|x^2-y^2\right| \le |x-y|\cdot 20 < 20\delta = \varepsilon.$$

This is what we wanted. ■

Here is the more general result.

Theorem 3.38. Suppose that $f(x)$ is continuous on a closed interval I. Then f is uniformly continuous.

Proof. We go by contradiction, I guess. Then there is a $\varepsilon > 0$ such that each $\delta > 0$ has a pair (x,y) with $|x-y|<\delta$ even though $|f(x)-f(y)| \ge \varepsilon$.

We now attack the continuity. For each $k \in \mathbb{N}$, we may find a_k and b_k with $|a_k - b_k| < \frac{1}{n}$ even though $|f(a_k)-f(b_k)| \ge \varepsilon$. Now find a convergent subsequence (!) $a_{\sigma k}$, which converges to some ℓ in the closed interval. But now

$$|b_{\sigma k} - \ell| \le |b_{\sigma k} - a_{\sigma k}| + |a_{\sigma k} - \ell|$$

will go to 0 as $k \to \infty$, so it follows that $b_{\sigma k} \to \ell$.

So we have by continuity that

$$\lim_{k\to\infty}\big(f(a_{\sigma k}) - f(b_{\sigma k})\big) = f(\ell) - f(\ell) = 0,$$

but this contradicts the assertion that $|f(a_{\sigma k}) - f(b_{\sigma k})| \ge \varepsilon$ always, so we are done here. ■

Remark 3.39. The above proof does not work for open intervals because our convergent subsequence does not need to converge in the interval. As a concrete example, $f(x) = \frac{1}{x}$ on $(0,1)$ is not uniformly continuous because of the speediness at 0.

Here is another nice property.

Proposition 3.40. Fix $f : S \to \mathbb{R}$ a uniformly continuous. Then if $\{a_k\}_{k\in\mathbb{N}}$ is Cauchy, then $\{fa_k\}_{k\in\mathbb{N}}$ is Cauchy.

This is not true for general functions.

Example 3.41. The continuous function $f(x) = \frac{1}{x}$ on $(0,1)$ has the Cauchy sequence $\{1/n\}_{n\in\mathbb{N}}$, which is not outputted to a Cauchy sequence.

Proof of Proposition 3.40. We do this by hand. Fix some ε so that we want N for which

$$n,m > N \overset{?}{\implies} |fa_n - fa_m| < \varepsilon.$$

But now there exists a δ for which $|x-y| < \delta$ implies $|f(x)-f(y)| < \varepsilon$, so we choose N by the Cauchy condition such that

$$n,m > N \implies |a_n - a_m| < \delta \implies |f(x)-f(y)| < \varepsilon,$$

which finishes. ■

Remark 3.42. We are using Cauchy sequences in the above rather than convergent sequences because we don't have to worry about the limit being inside of the domain or not. Our example from earlier exemplified this.

We have the following quick thought.

Proposition 3.43. If $f : S \to \mathbb{R}$ is uniformly continuous and $T \subseteq S$, then $f|_T : T \to \mathbb{R}$ is uniformly continuous.

Proof. Fix $\varepsilon > 0$. Then on S there is some δ for which $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$, but this same δ also works on T. ■

Example 3.44. We have that x^2 is uniformly continuous on $(0, 1)$ because it is uniformly continuous on $[0, 1]$.

Non-Example 3.45. We cannot "fix" $f(x) = 1/x$ on $(0, 1)$ to be uniformly continuous by replacing this with $[0, 1]$ because there is no way to add a point at 0 to get a continuous function.

Synthesizing the above two examples gives the following result.

Theorem 3.46. Fix (a, b) an open interval and $f : (a, b) \to \mathbb{R}$. Then f is uniformly continuous on (a, b) if and only if f can be extended to be continuous on $[a, b]$.

Proof. One direction is not so bad: if we can extend f to be continuous on $[a, b]$, then it is uniformly continuous, so it is uniformly continuous on (a, b).

The other direction is a bit trickier, so we won't prove it explicitly; take f uniformly continuous. Then the point is that f takes Cauchy sequences to Cauchy sequences! So find some Cauchy sequence in (a, b) which converges to $\inf(a, b) = a$, and then we know that the outputs by f will be another Cauchy sequence. This converges in $\mathbb{R}$, so say it converges to $f(a)$. Doing similar for $f(b)$ gets us our necessary extension.

There is some trick that we need to do to make sure all sequences converging to a will converge to the required $f(a)$. Well, if we have another a'_n such that $a'_n \to a$ as $n \to \infty$, we need to show $f(a'_n) \to f(a)$ still. Well, if it converges to some value b, then the sequence

$$a_1, a'_1, a_2 a'_2, \ldots$$

will have output converging to both $f(a)$ and b, so $b = f(a)$ as needed. Regardless, something like this should finish the proof. ■

Here is another result.

Proposition 3.47. Fix f a function defined on an interval I. If f is differentiable with bounded derivative on the interior of I, then f is uniformly continuous on I.

Non-Example 3.48. The function $f(x) = \frac{1}{x}$ on $(0, 1)$ has derivative approaching $-\infty$ as $x \to 0$, and so it fails the above, as it should.

Proof. Essentially, the bounded derivative shows that the secant lines have bounded slopes by the Mean value theorem. We will not say more here. ■

3.3.4 Exercises

And let's close off with some exercises.

Exercise 3.49 (Ross 19.1(f))**.** The function $f(x) = \sin\left(\frac{1}{x^2}\right)$ is not uniformly continuous on $(0, 1]$.

Proof. The point here is that we cannot extend f to a continuous function on $[0, 1]$. Namely, we can find sequences $\{a_k\}_{k\in\mathbb{N}}$ and $\{b_k\}_{k\to\infty}$ with $a_k \to 0$ and $b_k \to 0$ with $f(a_k) \equiv 0$ and $f(b_k) \equiv 1$ (say), so there is no way to assign some $f(0)$ continuously. ■

Exercise 3.50 (Ross 19.1(g))**.** The function $g(x) = x^2 \sin\left(\frac{1}{x}\right)$ on $(0, 1]$.

Proof. We can extend g to a continuous function on $[0, 1]$ by setting

$$g(0) = \lim_{x\to 0} g(x) = \lim_{x\to 0} x^2 \sin\left(\frac{1}{x}\right),$$

which is 0 by the Squeeze theorem. ■

Exercise 3.51 (Ross 19.3(a))**.** We show that $f(x) = \frac{x}{x+2}$ on $[0, 2]$ by hand.

Proof. For each $\varepsilon > 0$, we need to find $\delta > 0$ such that

$$|x - y| < \delta \implies \left|\frac{x}{x+2} - \frac{y}{y+2}\right| < \varepsilon.$$

Well, we can massage a bit by writing

$$\left|\frac{x}{x+2} - \frac{y}{y+2}\right| = \left|\frac{2(x-y)}{(x+2)(y+2)}\right| \le \frac{2}{(0+2)(0+2)}|x-y| = \frac{1}{2}|x-y|.$$

So we see $\delta = 2\varepsilon$ so that

$$|x - y| < \delta \implies \left|\frac{x}{x+2} - \frac{y}{y+2}\right| < \frac{1}{2}|x-y| < \frac{1}{2}\delta = \varepsilon,$$

which is what we wanted. ■

Exercise 3.52. Suppose that f is continuous on $[0, \infty)$ an uniformly continuous on some $[k, \infty)$ for some $k > 0$. Then f is uniformly continuous on $[0, \infty)$.

Proof. Fix $\varepsilon > 0$. We can find δ_2 so that, for $x, y \in [k, \infty)$, we have

$$|x - y| < \delta_2 \implies |f(x) - f(y)| < \varepsilon.$$

Similarly, we can find δ_1 so that, for $x, y \in [0, k + \delta]$, we have

$$|x - y| < \delta_1 \implies |f(x) - f(y)| < \varepsilon$$

because f is continuous and hence uniformly continuous on $[0, k + \delta]$. So taking $\delta := \min\{\delta_1, \delta_2\}$ has

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

in all cases, after some care. Indeed, if $|x - y| < \delta$, then either $x, y \in [0, k + \delta]$ or $x, y \in [k, \infty)$. ■

3.4 October 25

Let's have some fun today.

3.4.1 Limits

We're going to talk about limits, with many of the same ideas from calculus or sequences. For example, they prove in the book the limit laws and so on, and we will not do all of these formally. Regardless, here is our definition.

Limits

Definition 3.53 (Limits). Fix $f : S \to \mathbb{R}$. We say that the *limit of* $f(x)$ *as* $x \to a$ *is* L *along* S, notated

$$\lim_{x \to a^S} f(x) = L,$$

if and only if for every sequence $\{a_n\}_{n\in\mathbb{N}} \subseteq S$ converging to a has $f(a_n)$ converge to L.

Note that we have added a subset S to our definition. One reason this is a good thing to do is that it lets us talk about limits of functions which are not defined over $\mathbb{R}$. For example, it's not that the limit

$$\lim_{x \to -\infty} \sqrt{x}$$

does not exist—the limit does not even make sense because the interesting values with $x \to -\infty$ aren't defined for $\sqrt{x}$. If we let $S := \mathbb{R}_{\geq 0}$, then

$$\lim_{x \to -\infty^S} \sqrt{x}$$

now at least will compile, though $-\infty$ perhaps doesn't make sense with this S. Similarly,

$$\lim_{x \to \infty^S} \frac{\tan x}{x(|\tan x| + 1)}$$

where S is the domain of $\tan x$ will at least make sense and equal 0, though without the S here, things make less sense.

Remark 3.54. We have that

$$\lim_{x \to a^+} f(x) = \lim_{x \to a^{(a,\infty)}} f(x),$$

so our limits generalize left and right limits. This also explains our notation.

As usual, we can modify our sequences definition of the limit to an ε-δ definition.

Proposition 3.55. Fix $f : S \to \mathbb{R}$. We have that

$$\lim_{x \to a^S} f(x) = L$$

if and only if, for each $\varepsilon > 0$, there exists $\delta > 0$ such that each $x \in S$ satisfying $|x - a| < \delta$ have $|f(x) - L| < \varepsilon$.

Proof. We have two implications.

- Suppose that, for each $\varepsilon > 0$, there exists $\delta > 0$ such that $x \in S$ has $|x - a| < \delta$ implying $|f(x) - L| < \varepsilon$. Now, take any sequence $\{a_n\}_{n\in\mathbb{N}} \subseteq S$ which converges to $a \in S$.

  Then, for any $\varepsilon > 0$, there exists δ such that $x \in S$ has $|x - a| < \delta$ implying $|f(x) - L| < \varepsilon$. However, there is an N for which $n > N$ implies $|a_n - a| < \delta$, so this N has $n > N$ implies $|f(a_n) - f(a)| < \varepsilon$ as well.

- Conversely, suppose that there is $\varepsilon > 0$ for which no $\delta > 0$ has $|x-a| < \delta$ implies $|f(x)-f(a)| < \varepsilon$. Then, for each $n \in \mathbb{N}$, there is some x_n with $|x_n - a| < 1/n$ and $|f(x_n) - f(a)| \geq \varepsilon$.

  But now the sequence x_n converges to a while $f(x_n)$ never goes within ε of $f(a)$, so $f(x_n)$ does not converge to $f(a)$. So it follows $\lim_{x\to a^S} f(x)$ is not L. ■

As an example, let's show the following.

Proposition 3.56. Fix $f : \mathbb{R} \to \mathbb{R}$. We have that $\lim_{x\to a} f(x) = L$ if and only if

$$\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = L.$$

Proof. We have two implications here.

- Suppose that $\lim_{x\to a} f(x) = L$. We will show $\lim_{x\to a^+} f(x) = L$, and the other limit is similar. Now, for any $\varepsilon > 0$, there exists $\delta > 0$ such that $x \in \mathbb{R}$ has

$$|x-a| < \delta \implies |f(x) - f(a)| < \varepsilon.$$

  Then for any $a \in (a, \infty)$, we have $|x-a| < \delta$ implies $|f(x)-f(a)| < \varepsilon$ still, finishing.

- Suppose that $\lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = L$. Then, for any $\varepsilon > 0$, there exists $\delta^+ > 0$ such that $x \in (a, \infty)$ has

$$|x-a| < \delta^+ \implies |f(x) - f(a)| < \varepsilon.$$

  Similarly, there exists $\delta^- > 0$ such that $x \in (-\infty, a)$ has the same. Then set $\delta := \min\{\delta^+, \delta^-\}$. Then because $x \in \mathbb{R}$ implies $x = a$ or $x \in (a, \infty)$ or $x \in (-\infty, a)$, we have in all cases that

$$|x-a| < \delta \implies |f(x) - f(a)| < \varepsilon.$$

  This is what we wanted. ■

Remark 3.57. It is possible to do the first implication with ideas from sequences: any sequence approaching a from the left will be some sequence and hence have the output converging to L.

To finish, let's do an exercise.

Exercise 3.58. Suppose $f_1, f_2, f_3 \in (a, b)$ with $f_1(x) \leq f_2(x) \leq f_3(x)$ on the domain. Then

$$\lim_{x\to a^+} f_1(x) = \lim_{x\to a^+} f_3(x) = L$$

implies

$$\lim_{x\to a^+} f_2(x) = L.$$

Proof. There might be ways to do this with sequences, but we can do this with ε-δ style ideas. Fix $\varepsilon > 0$ so that there exists $\delta_1 > 0$ such that $x \in (a, b)$ has

$$|x-a| < \delta_1 \implies |f_1(x) - L| < \varepsilon.$$

Similarly, there exists $\delta_3 > 0$ such that $x \in (a, b)$ has

$$|x-a| < \delta_2 \implies |f_3(x) - L| < \varepsilon.$$

Now, as usual, take $\delta := \min\{\delta_1, \delta_2\}$. Then any $x \in (a, b)$ will have

$$-\varepsilon < f_1(x) - L \leq f_2(x) - L \leq f_3(x) - L < \varepsilon$$

by using our definitions of δ_1 and δ_3 on the left and right. This finishes. ■

3.4.2 Upgrading to Metric Spaces

Let's start by moving continuity up to a metric space. The main point is that the condition

$$|x-a|<\delta \implies |f(x)-f(a)|<\varepsilon$$

can be moved into a continuity condition by noting that $|x-a|$ is really our distance metric. So here is our definition.

Continuity for metric spaces

Definition 3.59 (Continuity for metric spaces)**.** Fix (X, d_X) and (Y, d_Y) metric spaces. Then a function $f : X \to Y$ is *continuous* at $a \in X$ if and only if, for each $\varepsilon > 0$, we have some $\delta > 0$ such that $x \in X$ has

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

Example 3.60. Using the telemetric on $\mathbb{R}$, all functions are continuous. The main point is that for any $\varepsilon > 0$, we can take $\delta = 1/2$ so that $d_X(x, a) < \delta$ implies $x = a$. Alternatively, any sequence converging to a must be eventually constant, so it of course lifts to a sequence converging to $f(a)$ upon pushing through f.

Example 3.61. The taxicab metric on $\mathbb{R}^2$, continuity actually looks the same as for Euclidean continuity $\mathbb{R}^2 \to \mathbb{R}^2$! Namely, the metrics induce the same topology on $\mathbb{R}^2$, though the taxicab metric does look different.

And we can go to uniformly continuous by moving around our quantifiers, as we did in $\mathbb{R}$.

Uniform continuity for metric spaces

Definition 3.62 (Uniform continuity for metric spaces)**.** Fix (X, d_X) and (Y, d_Y) metric spaces. Then a function $f : X \to Y$ is *uniformly continuous* if and only if, for each $\varepsilon > 0$, we have $\delta > 0$ such that $x, y \in X$ has

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

Let's build towards a more topological definition of continuity.

Pre-image

Definition 3.63 (Pre-image)**.** Let $f : S \to T$ be a function. Then we define the *pre-image*, for $B \subseteq T$,

$$f^{-1}(B) := \{s \in S : f(S) \in B\}.$$

Warning 3.64. This f^{-1} is not the same as an inverse function! Here, $f^{-1} : \mathcal{P}(T) \to \mathcal{P}(S)$ is defined for arbitrary functions (not necessarily one-to-one).

3.4.3 A Better Continuity

And here is our topological definition of continuity.

Theorem 3.65. Fix (X, d_X) and (Y, d_Y) metric spaces. We have that $f : X \to Y$ is continuous if and only if, f^{-1} sends open sets of Y to open sets of X.

Proof. We have two implications.

- Let's start by taking f continuous. Fix $U \subseteq Y$ an open set so that we want to show $f^{-1}(U)$ is open. Well, find any $a \in f^{-1}(U)$, and we want to show that a is in the interior of $f^{-1}(U)$.

Now, $f(a) \in U$ is in an open set, so there is a ε such that

$$\{y \in Y : d_Y(y, f(a)) < \varepsilon\} \subseteq U.$$

By the continuity of f, we are promised $\delta > 0$ such that $d_X(x, a) < \delta$ implies $d_Y(f(x), f(a)) < \varepsilon$. It follows that

$$\{x \in X : d_X(x, a) < \delta\} \subseteq \{x \in X : d_Y(f(x), f(a)) < \varepsilon\} \subseteq f^{-1}(U)$$

because $d_Y(f(x), f(a)) < \varepsilon$ implies $f(x) \in U$. Thus, a is indeed in the interior of $f^{-1}(U)$, finishing.

- Now suppose that f takes open sets to open sets. Take $a \in X$ so that we want to show f is continuous at a. Now, for any $\varepsilon > 0$, we note that

$$B := \{y \in Y : d_Y(a, y) < \varepsilon\}$$

is an open set of Y, so it follows that

$$f^{-1}(B)$$

is also open. But $a \in X$ has $f(a) \in B$, so $a \in f^{-1}(B)$ is in the interior of $f^{-1}(B)$, so there is $\delta > 0$ such that

$$\{x \in X : d_X(a, x) < \delta\} \subseteq f^{-1}(B).$$

It follows that $d_X(a, x) < \delta$ implies $f(x) \in B$ implies $d_Y(f(a), f(x)) < \varepsilon$, which shows that f is indeed continuous at $x = a$. ■

Remark 3.66. What is good about this definition is that it works nicely and is quite simple for more general topological spaces.

Remark 3.67. I think we can actually strengthen the above statement to say that f is continuous at $x = a$ if and only if each open set $U_Y \subseteq Y$ containing $f(a)$ has a pre-image $f^{-1}(U)$ containing some open subset $U_X \subseteq f^{-1}(U)$ containing $a \in U_X$. The point is that the above proof is very "local" at a.

Here is an example of something that we get from this characterization of continuity.

Proposition 3.68. Fix $f : X \to Y$ a continuous map of metric spaces. Then if X is compact, then $f(X)$ is compact.

Proof. The idea is that any open cover of $f(X)$ can be pulled back along f to an open cover of X. Then the open cover of X has a finite subcover, which tells us which sets of the open cover of $f(X)$ we "really need" to cover.

Formally, suppose that $\{U_\alpha\}_{\alpha \in \lambda}$ is an open cover of $f(X)$. Then, for each $x \in X$, we have that $f(x) \in U_\alpha$ for some $\alpha \in \lambda$, so each $x \in X$ belongs to some $f^{-1}(U_\alpha)$ for some α. It follows that

$$\mathcal{U}_X := \left\{f^{-1}(U_\alpha) : U_\alpha \in \lambda\right\}$$

will fully cover X, and this is in fact an open cover because f is continuous. Now, $\mathcal{U}_X$ is an open cover of X, so compactness of X promises us some sequence $\{\alpha_k\}_{k=1}^n$ which is yields a finite subcover of X.

To finish, we claim that the $\{U_{\alpha_k}\}_{k=1}^n$ fully covers $f(X)$. Indeed, for any $y \in f(X)$, there is x such that $f(x) = y$. Now, $x \in f^{-1}(U_{\alpha_k})$ for some k because the $f^{-1}(U_{\alpha_\bullet})$ fully cover X. Thus, $y \in f(x) \in U_{\alpha_k}$, so indeed our finite subcover does cover. ■

And as usual, we can show the following.

Proposition 3.69. Fix (X, d_X) and (Y, d_Y) metric spaces. If X is compact, and if $f : X \to Y$ is continuous, then f is uniformly continuous.

Proof. Fix $\varepsilon > 0$. Now, for each $a \in X$, there is some δ_a such that

$$\{f(x) \in Y : d_X(a, x) < \delta_a\} \subseteq \{y \in Y : d_Y(y, a) < \varepsilon/2\}.$$

Now set $U_a := \{a \in X : d_X(a, x) < \delta_a/2\}$, which is open in X. Then we see that $a \in U_a$, so

$$X = \bigcup_{x \in X} \{x\} \subseteq \bigcup_{x \in X} U_x,$$

so the $U_\bullet$ provide an open cover of X. But now compactness of X implies that we may choose some finite subcover $\{U_{x_k}\}_{k=1}^n$ of X, and because this is finite, we may fix

$$\delta := \frac{1}{2} \min_{1 \le k \le n} \delta_k.$$

Fix $a, b \in X$ with $d_X(a, b) < \delta$. Because of our open cover, we can place $x_1 \in U_k$, but now

$$d_X(b, x_k) \le d(a, b) + d(b, x_k) < \delta + \frac{1}{2}\delta_k < \delta_{x_k}.$$

But also $d_X(a, x_k) < \delta_{x_k}$, so we see that $d_Y(f(a), f(x_k)) < \varepsilon/2$ and $d_Y(f(b), x_k) < \varepsilon/2$ by construction of the $U_\bullet$. It follows $d_Y(f(a), f(b)) < \varepsilon$. ∎

Remark 3.70. This is essentially the generalization of the statement that any uniformly continuous function on a closed interval is uniformly continuous.

Let's close with some exercises.

Exercise 3.71. Fix (X, d) a metric space, and fix $x_0 \in X$. Then show $f : X \to \mathbb{R}$ defined by

$$f(x) := d(x, x_0)$$

is uniformly continuous.

Proof. Fix $\varepsilon > 0$. Then we want to define $\delta > 0$ such that

$$d(x, y) < \delta \overset{?}{\implies} |d(x, x_0) - d(y, x_0)| < \varepsilon.$$

Well, the point is that

$$d(x, x_0) \le d(x, y) + d(y, x_0),$$

so $d(x, x_0) - d(y, y_0) \le d(x, y)$. Similarly, we see that

$$d(y, x_0) \le d(y, x) + d(x, x_0),$$

so $d(y, y_0) - d(x, x_0) \le d(x, y)$. It follows that we can take $\delta := \varepsilon$ so that

$$d(x, y) < \delta \implies |d(x, x_0) - d(y, y_0)| \le d(x, y) < \delta < \varepsilon,$$

which is exactly what we wanted. ∎

Exercise 3.72 (Ross 21.9)**.** Find a continuous, surjective function $[0,1]^2 \twoheadrightarrow [0,1]$.

Proof. Consider the function $\pi : [0,1]^2 \to [0,1]$ defined by

$$\pi(x,y) := x.$$

This is well-defined because the first coordinate does live in $[0,1]$. It remains to check continuity; we show that π is uniformly continuous. Fix $\varepsilon > 0$. Now, for any two points $(x_1,y_1),(x_2,y_2) \in [0,1]^2$ take $\delta := \varepsilon$ so that

$$\sqrt{(x_1-x_2)^2+(y_1-y_2)^2} < \delta = \varepsilon$$

implies that $(x_1-x_2)^2 < \varepsilon - (y_1-y_2)^2 \le \varepsilon^2$ implies that $|x_1-x_2| < \varepsilon$. ■

3.5 October 27

Here we go again.

3.5.1 Connectedness

Let's talk about connectedness. It matters later in life. We have the following definition.

Connected, I

Definition 3.73 (Connected, I)**.** A metric space (X,d) is *connected* if and only if has no proper nonempty subset which is both open and closed in X. If (X,d) is not connected, then we say that it is *disconnected*.

Remark 3.74. Of course, both $\varnothing$ and X are

There are some equivalent variations of this definitions. Here is an example.

Connected, II

Definition 3.75 (Connected, II)**.** A metric space (X,d) is *disconnected* if and only if we can write $X := U_1 \sqcup U_2$ for nonempty subsets U_1 and U_2. If this is impossible, we say that X is connected.

With this in mind, we have the following.

Definition 3.76. If some proper nonempty open subset $U \subseteq X$ with U both open and closed, then we say that U *disconnects* X.

Remark 3.77. This is best seen geometrically: usually what is happening with disconnections is seen with $X := [0,1] \cup [2,3]$, where the subset $[0,1]$ is both open and closed.

Example 3.78. In $\mathbb{Z}$, any nonempty proper subset will disconnect $\mathbb{Z}$. Indeed, points are open in $\mathbb{Z}$, so all subsets are open.

Example 3.79. In $\mathbb{Q}$, the subset

$$U := \{x \in \mathbb{Q} : x > \sqrt{2}\}$$

is open, but its complement is

$$U^c = \{x \in \mathbb{Q} : x < \sqrt{2}\}$$

because $\sqrt{2} \notin \mathbb{Q}$, so we find that U^c is also open, making U closed.

In general, it is much harder to prove that a space is connected than disconnected because disconnection merely requires us to exhibit the disconnecting subset.

Let's use this definition for something.

Proposition 3.80. Fix $f : X \to Y$ a continuous function between the metric spaces (X, d_X) and (Y, d_Y). If X is connected, then $f(X)$ is connected.

Proof. We can think about this topologically, similar to the proof that continuity preserves compactness. Anyways, we show this by contraposition: take $f(X)$ disconnected, and we show that X is disconnected. This means that we have a proper nonempty subset $U \subseteq f(X)$ which is both open and closed so that

$$f(X) = U \cup U^c$$

shows that $f(X)$ can be disjointedly decomposed into two nonempty open sets. By continuity, we see that $f^{-1}(U)$ and $f^{-1}(U^c)$ are both open. Further, they union to X because each $x \in X$ has $f(x)$ in one of U and U^c, so x lives in one of $f^{-1}(U)$ and $f^{-1}(U^c)$.

So we see that $f^{-1}(U)$ is both open and closed; it remains to show that $f^{-1}(U)$ is nonempty and proper. Well, U was nonempty, so take $y \in U \subseteq f(X)$ so that there exists $x \in X$ such that $f(x) = y \in U$. It follows that $x \in f^{-1}(U)$, so $f^{-1}(U)$ is nonempty. Then because U^c was nonempty, it follows $f^{-1}(U^c)$ is nonempty as well, so $f^{-1}(U)$ is proper. This finishes. ■

Remark 3.81. We can use this as something to determine what continuous functions might look like. For example, we know immediately that there is no continuous surjection from $[0, 1] \to [0, 1] \cup [2, 3]$ once we know that $[0, 1]$ is connected while $[0, 1] \cup [2, 3]$ is not.

Anyways, we should probably give an example of a nontrivial connected set.

Proposition 3.82. A subset $S \subseteq \mathbb{R}$ is connected if and only if S is an interval.

Proof. We have two implications.

- Suppose that S is not an interval. This means that there exists $t \in \mathbb{R}$ with $t \notin S$ such that $S \cap (t, \infty) \neq \varnothing$ and $S \cap (-\infty, t) \neq \varnothing$.[1] But now these exact two sets will disconnect S! Indeed, we set

$$U_1 := S \cap (t, \infty) \qquad \text{and} \qquad U_2 := S \cap (-\infty, t).$$

  We see that $U_1 \cup U_2 = S$ because $t \notin S$, and we know they are both nonempty, so it remains to show that they are open. This is immediate by the induced topology, but we can use sphere arguments here: each point $s \in U_1$ will live in the interior of U_1 by using $r := s - t > 0$ as our radius.

Remark 3.83. This is more or less generalizing the proof that $\mathbb{Q}$ is disconnected.

- It remains to show that intervals are connected. We start by reducing to the case where interval has the form $[a, b]$. Fix $I \subseteq \mathbb{R}$ an interval, and suppose for the sake of contradiction we can write

$$I := U_1 \cup U_2,$$

  where U_1 and U_2 are disjoint, nonempty open subsets. Now find $a \in U_1$ and $b \in U_2$, and without loss of generality we may take $a < b$. Now we see that

$$(U_1 \cap [a, b]) \cup (U_2 \cap [a, b]) = I \cap [a, b] = [a, b].$$

[1] This seems annoying to show, but I don't want to think about it.

Here both of these are open by the indued topology, they are nonempty because $a \in U_1 \cap [a,b]$ and $b \in U_2 \cap [a,b]$, and these are disjoint because U_1 and U_2 are in fact disjoint. Thus, $[a,b]$ is also disconnected.

We now forget what we were doing and show that closed intervals $[a,b]$ are disconnected. Indeed, suppose that we can write

$$[a,b] = U_1 \sqcup U_2,$$

where U_1 and U_2 are disjoint open sets. Without loss of generality, $a \in U_1$ and $b \in U_2$.

Now we use the order topology on $\mathbb{R}$ by applying the pushing trick from showing $[0,1]$ is compact. We see that $\sup U_1 > a$ because there is a sphere around a inside of U_1, and we see that $\sup U_2 < b$ because there is a sphere around b inside of U_2.

But now we notice that $\sup U_1 \notin U_1$ because if so then we could place an open ball around $\sup U_1$ to get larger. On the other hand, $\sup U_1 \notin U_2$ because if so we could place an open ball around $\sup U_1$ in U_2 to force the elements of U_1 to be smaller. ■

We remark that the Intermediate value theorem follows from this statement: if a continuous function $f : S \to \mathbb{R}$ has domain an interval, then its domain is connected, so its image is connected, so its image is an interval.

Remark 3.84. There is also a notion of path-connectedness mentioned in the book, but we will not care about it in this class.

Let's do an exercise and then move on.

Exercise 3.85 (Ross 22.3)**.** Fix $E \subseteq (X,d)$ a connected subset of a metric space. Show that the closure $\overline{E}$ is also connected.

Proof. We show the contraposition: suppose $\overline{E}$ is disconnected, and we show E is disconnected. Then fix disjoint nonempty open sets $U_1, U_2 \subseteq E$ which union to E. But now we see that

$$(U_1 \cap E) \cup (U_2 \cap E) = E$$

is a disjoint union of open sets in E which union to E.

It remains to show that U_1 and U_2 are nonempty. Well, suppose for the sake of contradiction that (say) $U_1 \cap E$ contains E so that U_1 contains E. But now this implies that U_1 is a proper closed subset of $\overline{E}$ which contains E, which violates the fact that $\overline{E}$ is the smallest closed set. Rigorizing this would be somewhat painful (we have to write out $\overline{E} = \bigcap_{V \supseteq E} V$ and talk about that definition). ■

3.5.2 Power Series

Let's preview some of chapter 4. Some of this will be review from calculus.

Power series

Definition 3.86 (Power series)**.** Given a sequence of real numbers $\{a_n\}_{n=0}^{\infty} \subseteq \mathbb{R}$, we define the (formal) *power series*

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

to more or less represents a function $\mathbb{R} \to$
RR.

Remark 3.87. I have defined the above formally so that we can plug in values of x and then ask if the series converges. This prevents us from wondering if the series "exists" at all.

Intuitively, it might feel like if

$$\sum_{n=0}^{\infty} a_k \cdot (-100)^k$$

converges then

$$\sum_{k=0}^{\infty} a_k \cdot 20^k$$

should also converge because the second series seems "uniformly" smaller in magnitude, even though perhaps the alternating series has an effect. But of course such intuition requires care because these series potentially feel very different.

In general, a series has a few possibilities.

- The series can converge at $x = 0$ only. All series must converge here because the series just looks like a_0 here.

- The series converges for all $x \in \mathbb{R}$.

- The series converges for each $|x| < R$ for some finite $R \in \mathbb{R}$, but the series diverges for $|x| > R$. (The behavior at $x = \mp R$ is intentionally unspecified.)

And here are some examples.

Example 3.88. The series

$$\exp(z) = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$$

will converge everywhere. We can kind of feel that the coefficients get really small (smaller than exponential) fast, so the series ought converge.

Example 3.89. The series

$$\sum_{k=0}^{\infty} 2^{k^2} x^k$$

will converge nowhere outside of 0. We can show this using the root test because

$$\sqrt[n]{|2^{n^2} x^n|} = |x| \cdot \sqrt[n]{2^{n^2}} = |x| \cdot 2^n \to \infty,$$

so this always diverges for $x \neq 0$.

Example 3.90. The series

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k$$

will converge for $|x| < 1$ and diverge for $|x| > 1$.

To evaluate the radius of convergence, as in the third case, it is most helpful to use the Root test. The ratio test is potentially helpful for particular series, but if the $a_\bullet$ are poorly behaved locally, then the Ratio test is also potentially poorly behaved. With this in mind, we have the following definition.

Radius of convergence

Definition 3.91 (Radius of convergence)**.** Given a power series

$$\sum_{k=0}^{\infty} a_k x^k,$$

we define the *radius of convergence* to be

$$R := \left(\limsup_{n\to\infty} \sqrt[n]{|a_n|}\right)^{-1}.$$

By convention, if the $\limsup$ is 0, we set $R := \infty$.

This behaves like the radius of convergence essentially by the Root test. We will not write out the proof here, but we will say that the main idea is that

$$\limsup_{n\to\infty} \sqrt[n]{|a_n x^n|} = |x| \cdot \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{|x|}{R},$$

so we converge (absolutely!) for $\frac{|x|}{R} < 1$ and diverge for $\frac{|x|}{R} > 1$. Namely, the fact that the Root test is inconclusive for ± 1 turns into the fact that we are unsure what happens for $x = \pm R$.

3.5.3 Uniform Continuity: A Prelude

We might want to formally integrate and differentiate a power series. For example, with $f(x) := \arctan x$, we have

$$f'(x) = \frac{1}{1+x^2} = \frac{1}{1-(-x^2)} = \sum_{k=0}^{\infty} (-1)^k x^{2k}.$$

Integrating would tell us that

$$f(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}.$$

These sorts of tricks are nice; for example, we are able to compute $\arctan$ from this, but the series for $\arcsin$ is quite worse.

However, these sorts of ideas require some care. For example, the series

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1} x^k}{k}$$

will converge at $x = 1$ because it is the alternating harmonic series, and in fact this converges to $\ln 2$. But when we differentiate this series, we see

$$\frac{1}{1+x} = \sum_{k=0}^{\infty} (-1)^k x^k,$$

which diverges at $x = 1$. So we do not get what we want without paying attention. What happened here?

On one hand, this problem will not occur worse: if we consider the series

$$f(x) = \sum_{k=0}^{\infty} a_k x^k,$$

then we see that

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \limsup_{n\to\infty} \sqrt[n]{|(n+1)a_{n+1}|}$$

because $\sqrt[n]{n+1} \to 1$. Thus, the radii of convergence between $f(x)$ and $f'(x)$ are the same. The point of this is to say that strengthening our convergence (to absolute) helps here, but absolute convergence is not actually necessary to make this work. The correct notion is uniform continuity.

Let's do some exercises before we talk more about uniform continuity.

Exercise 3.92. We find the interval of convergence for

$$\sum_{n=1}^{\infty} n^2 x^n.$$

Proof. The main point is to compute

$$\limsup_{n\to\infty} \sqrt[n]{n^2}.$$

Well, we see that $\log \sqrt[n]{n^2} = \frac{2}{n}\log n$, which goes to 0 as $n \to \infty$. So we have that

$$\lim_{n\to\infty} \sqrt[n]{n^2} = e^1 = 1,$$

so our radius of convergence is $R = 1^{-1} = 1$. Alternatively, the book proves $n^{1/n} \to 1$ somewhere, so it follows $n^{2/n} \to 1$ follows.

It remains to deal with the endpoints. Well, $\sum(-1)^n n^2$ will always diverge by the Divergence test because $\left|n^2\right| = n^2 \to \infty$. So our interval of convergence is $\boxed{(-1,1)}$. ■

Exercise 3.93. We find the interval of convergence of

$$\sum_{n=1}^{\infty} \left(\frac{x}{n}\right)^n.$$

Proof. The main point is to compute

$$\limsup_{n\to\infty} \sqrt[n]{\left(\frac{1}{n}\right)^n} = \limsup_{n\to\infty} \frac{1}{n} = 0,$$

so the radius of convergence is $+\infty$. Thus, our series converges $\boxed{\text{everywhere}}$. ■

Remark 3.94. This makes sense because n^{-n} gets small very fast, faster than for $\frac{1}{n!}$.

Exercise 3.95. We compute the interval of convergence of

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

Proof. The main point is to compute

$$\limsup_{n\to\infty} \sqrt[n]{\frac{2^n}{n^2}} = 2 \cdot \limsup_{n\to\infty} n^{-2/n} = 2,$$

so our radius of convergence is $1/2$. As for the endpoints, we find that

$$\sum_{n=1}^{\infty} \left|\frac{2^n}{n^2}\left(\frac{1}{2}\right)^n\right| = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

converges, so the series converge absolutely. It follows that our interval of convergence is $\boxed{[-1/2, 1/2]}$. ■

Exercise 3.96. We compute the interval of convergence of

$$\sum_{n=1}^{\infty} \frac{n^3}{3^n} x^n.$$

Proof. We can see that

$$\limsup_{n\to\infty} \sqrt[n]{\frac{n^3}{3^n}} = \frac{1}{3},$$

so it follows that our radius of convergence is is 3. The endpoints give $\sum(\pm 1)^n n^3$, which fail the divergence test, so our interval of convergence is $\boxed{(-3,3)}$. ■

3.5.4 Uniform Continuity: Another Prelude

Let's see another reason we might want stronger continuity. Fix

$$f_n(x) := x^n$$

for $f_n : [0,1] \to [0,1]$ and $n \ge 0$. But now we see that, for $x \in [0,1)$, we have that

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} x^n = 0$$

while

$$\lim_{n\to\infty} f_n(1) = \lim_{n\to\infty} 1 = 1.$$

So we let

$$f(x) := \begin{cases} 1 & x = 1, \\ 0 & x \neq 1, \end{cases}$$

so that $f_n \to f$ as $n \to \infty$. But now the f_n are continuous while converging to a discontinuous function, which is sad. The problem, again, is that our convergence is somehow "too weak," and the stronger form of convergence–"uniform convergence"–is what we want.

3.6 November 1

Finally done with October. One last push I guess.

3.6.1 Uniform Convergence

Last class we were looking at the sequence of functions

$$f_n(x) := x^n$$

for $f_n : [0,1] \to [0,1]$. Then we say that, for $x \in [0,1)$, we see that

$$\lim_{n\to\infty} f_n(x) = 0,$$

but

$$\lim_{n\to\infty} f_n(1) = 1.$$

With this in mind, we define

$$f(x) := \begin{cases} 0 & 0 \le x < 1, \\ 1 & x = 1. \end{cases}$$

We would like to say that $f_n \to f$ as $n \to \infty$. This notion is more rigorously called "pointwise" convergence.

Pointwise convergence

Definition 3.97 (Pointwise convergence)**.** Given a sequence of functions $f_n : S \to \mathbb{R}$, we say that f_n *converges pointwise to* to some $f : S \to \mathbb{R}$ if and only if, for each $x \in S$, we have that $f_n(x) \to f(x)$ as $n \to \infty$.

This is perhaps the weakest and worst form of convergence. For example, it has the defect that continuous functions can pointwise converge to a discontinuous function.

Remark 3.98. The high-level reason why this is failing is that our convergence is not very uniform: points x close to 1 will have $f_n(x) \to 0$ very slowly.

So we need to strengthen our notion of convergence. This gives uniform convergence.

Uniform convergence

Definition 3.99 (Uniform convergence)**.** Fix a sequence of functions $f_n : S \to \mathbb{R}$. Then $f_n \to f$ *uniformly converges to* $f : S \to \mathbb{R}$ if and only if, for each $\varepsilon > 0$, there is N such that $n > N$ implies

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

for each $x \in \mathbb{R}$.

This is different from pointwise convergence because now ε is not allowed to vary with x. This is similar to uniform continuity in that, again, ε was not allowed to depend on which point whose continuity we were looking at.

Non-Example 3.100. It is not the case that $f_n(x) := x^n$ converges to $f(x) = 1_{x=1}$ uniformly. Well, take $\varepsilon := 1/2$. Then there is no N such that $n > N$ implies

$$|x^n| < 1/2$$

for each $x \in [0, 1)$. Indeed, for any N, choose any $n > N$ and then take $x := \sqrt[n]{2/3}$ so that $|x^n| = 2/3 > 1/2$.

Anyways, let's see uniform convergence do something useful.

Proposition 3.101. Suppose that $f_n : S \to \mathbb{R}$ is a sequence of continuous functions converging uniformly to $f : S \to \mathbb{R}$. Then f is continuous.

Proof. Fix any $a \in S$ so that we want to show f is continuous at a. Well, fix any $\varepsilon > 0$. Then there exists N such that $n > N$ has

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3}$$

for each $x \in S$. Now, fix any $n > N$. Because f_n is continuous, there exists δ such that

$$|x - a| < \delta \implies |f_n(x) - f_n(a)| < \frac{\varepsilon}{3}.$$

The point of these estimates is the following manipulation: for $x \in S$ with $|x - a| < \delta$, we see

$$\begin{aligned} |f(x) - f(a)| &\le |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\ &= \varepsilon, \end{aligned}$$

which establishes the needed continuity. ■

Remark 3.102. This sort of "double triangle inequality" idea is fairly common. For example, a similar argument shows that uniformly continuous functions that converge uniformly will converge to a uniformly convergent function. (Dropping the "uniformly" in either hypothesis makes this false.)

We quickly remark that we also have the following equivalent definition of uniform convergence.

Proposition 3.103. Fix $f_n : S \to \mathbb{R}$ and $f : S \to \mathbb{R}$. Then $f_n \to f$ converges uniformly if and only if

$$\lim_{n\to\infty} \sup |f(x) - f_n(x)| = 0.$$

Proof. Omitted; see the book. ■

Let's do some exercises.

Exercise 3.104. We study the sequence of functions $f_n(x) := \frac{1}{1+x^n}$ on $[0, \infty)$ for $n \in \mathbb{N}$.

Proof. We start by evaluating the pointwise limit.

- For $0 \le x < 1$, we see that $x^n \to 0$, so $1 + x^n \to 1$, so $f_n(x) \to 1$.
- For $x = 1$, we see that this is constantly $f_n(x) = \frac{1}{2}$.
- For $x > 1$, we see that $x^n \to \infty$, so $f_n(x) \to 0$.

So our convergent function is

$$f(x) = \begin{cases} 1 & 0 \le x < 1, \\ 1/2 & x = 1, \\ 0 & x > 1. \end{cases}$$

For example, this implies that $f_n \to f$ is not uniform on $[0, 1]$ because f is not continuous on $[0, 1]$.

Explicitly, we fix $x \in [0, 1]$, and we can evaluate

$$f_n(x) - f(x) = \frac{-x^n}{1 + x^n},$$

so we see that

$$\sup_{x\in[0,1,)} |f_n(x) - f(x)| = \frac{1}{2}$$

after doing a bit of legwork. So indeed, $f_n \to f$ is not uniform.

How about $[0, 1/2]$? Here we can evaluate

$$\sup_{x\in[0,1)} |f_n(x) - f(x)| = \frac{(1/2)^n}{1 + (1/2)^n}$$

because this difference is increasing. So this vanishes as $n \to \infty$, and we are safe. ■

Remark 3.105. Uniform continuity does not really care for individual points because it is a global concept. For example, $f_n \to f$ is not uniform even on $[0, 1)$, using the same function.

3.6.2 Being Integrable

It feels like there ought to be some property that pointwise convergence preserves. For example, does pointwise convergence preserve being integrable?

Example 3.106. Consider

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

on $(0,1)$. These functions will converge to $\frac{1}{1-x}$ as $n \to \infty$, which is not integrable on $(0,1)$ (even though f is not continuous).

Well, what about uniform continuity?

Proposition 3.107. Fix $f_n : S \to \mathbb{R}$ a sequence of functions converging uniformly to a function $f : S \to \mathbb{R}$. Then f is integrable on bounded domains.

Proof. We take a few properties of integration on faith because we have not defined what it means to be integrable. Then, if we are integrating of over a bounded domain $T \subseteq S \cap [-M, M]$, then, for any $\varepsilon > 0$, there exists N such that $n > N$ has

$$\left| \int_T f(x)\,dx - \int_T f_n(x)\,dx \right| \le \int_T |f(x) - f_n(x)|\,dx \le \int_{[-M,M]} \varepsilon\,dx = 2M\varepsilon.$$

Sending $\varepsilon \to 0$ shows that $\int_T f(x)$ is well-defined as a real number. ■

However, we need to take some care for unbounded domains. I'm honestly not sure what is the case.

3.6.3 Uniformly Cauchy

Here is another notion.

Definition 3.108. A sequence of functions $f_n : S \to \mathbb{R}$ is *uniformly Cauchy* if and only if there exists $\varepsilon > 0$ such that $N \in \mathbb{N}$ has $n, m > N$ implies

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

for each $x \in S$.

Remark 3.109. This is essentially asserting that $\{f_n\}_{n\in\mathbb{N}}$ converges in the space of functions $\mathbb{R} \to \mathbb{R}$.

We would like uniformly Cauchy sequences of functions to converge uniformly to some function. Indeed this is the case.

Proposition 3.110. A uniformly Cauchy sequence $f_n : S \to \mathbb{R}$ converges uniformly to a function $f : S \to \mathbb{R}$.

Proof. We start by exhibiting our function $f : S \to \mathbb{R}$. We do this pointwise: for each $x \in S$, we see that $\{f_n(x)\}_{n\in\mathbb{N}}$ is a Cauchy sequence and hence converges to some $f(x)$.

Now we show that $f_n \to f$ uniformly. Well, for any $\varepsilon > 0$, there exists N so that $n, m > N$ has

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

for each $x \in S$. Now we notice that each $f_n(x)$ is inside of $\left(f_m(x) - \frac{\varepsilon}{2}, f_m(x) + \frac{\varepsilon}{2}\right)$, so the limit of the f_n must live inside of this interval as well. In particular, have that our N has $m > N$ implying

$$|f(x) - f_m(x)| \le \frac{\varepsilon}{2} < \varepsilon$$

for each $x \in S$. ■

3.6.4 Series of Functions

Of course, we are really interested in power series. So, for example, does

$$\sum_{k=0}^{n} \frac{1}{k!} x^k \to e^x$$

converge uniformly? More explicitly, we define

$$f_n(x) := \sum_{k=0}^{n} \frac{1}{k!} x^k.$$

We would like this to converge to a real function, and we would also like to know if this converge is uniform. The answer is no: very negative values can go over the rails, even for large n. But we can save ourselves by looking at a bounded interval, where our convergence should be better-behaved. We won't write out the details here.

It turns out that restricting to a bounded interval is also not good enough, however.

Example 3.111. The series

$$-\log(1-x) = \sum_{k=1}^{\infty} \frac{x^k}{k}$$

will have the partial sums converge but not uniformly.

What about preserving differentiability?

Example 3.112. The series

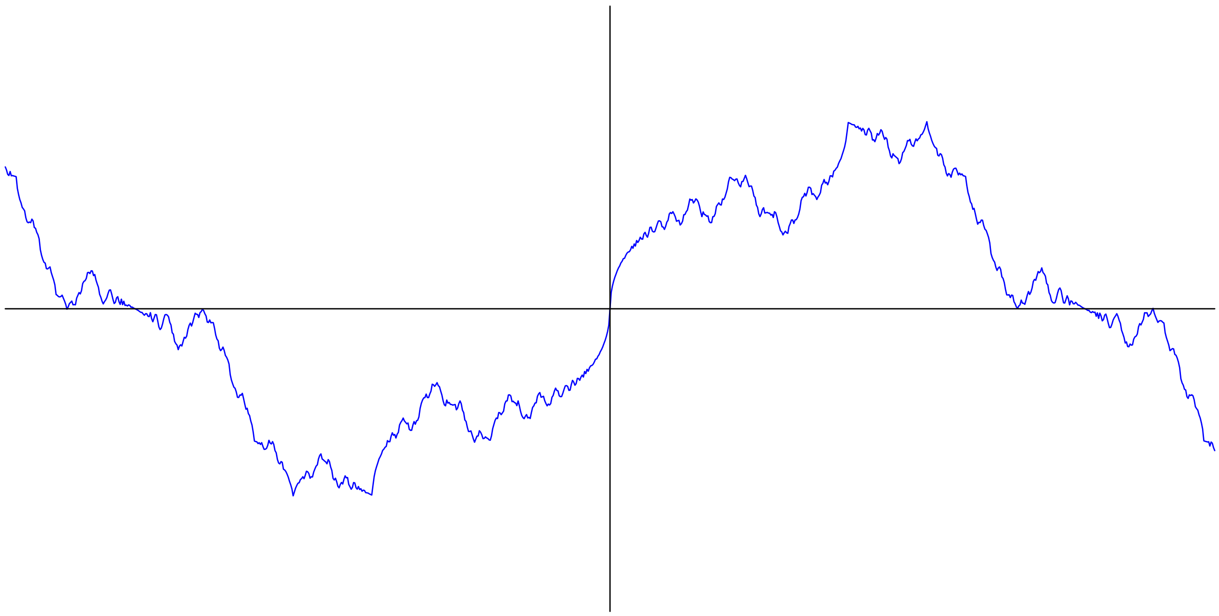
$$f(x) := \sum_{k=1}^{\infty} \frac{\sin\left(n^2 x\right)}{n^2}$$

has its partial sums converge uniformly to a continuous function. However, its derivative, taken pointwise, is

$$f'(x) = \sum_{k=1}^{\infty} \cos\left(n^2 x\right),$$

which does not seem to converge anywhere. So this function seems to be continuous everywhere and differentiable nowhere.

For completeness, here is the graph of the above function.



Anyways, we would like to talk about series of functions. Here is a good test for this purpose.

Theorem 3.113 (Weierstrass M-test)**.** Fix a sum

$$\sum_{k=1}^{\infty} g_k(x).$$

Suppose that there exists $M_k \geq 0$ such that $|g_k(x)| \leq M_k$ for each k while $\sum_{k=1}^{\infty} M_k$ converges uniformly. Then the partial sums uniformly converge.

Example 3.114. We can use the above test to show that e^x converges uniformly on any bounded interval. Essentially, even though $\frac{x^k}{k!}$ might not be bounded on all of $\mathbb{R}$, it is at least bounded on the bounded interval, which is good enough for the above test.

And here is an exercise to close us out.

Exercise 3.115. Suppose that $f_n : S \to \mathbb{R}$ is a bounded function converging uniformly to some $f : S \to \mathbb{R}$. Then f is bounded.

Proof. Fix $\varepsilon = 1$, and then we are promised N such that $n > N$ has

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

for each $x \in S$. Now fix some $n > N$. We see that f_n is bounded so that $f_n(x) \in [-M, M]$ for some $M \in \mathbb{R}$. But now $f(x) \in [-M-1, M+1]$ for each $x \in S$, so we get that f is bounded. ∎

3.7 November 3

So we have a midterm next class.

3.7.1 Integrating and Differentiating Power Series

Quickly, we recall that, if $f_n : [a,b] \to \mathbb{R}$ is a sequence of continuous functions converging uniformly to some continuous function f, then

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

Note that the integral on the right-hand side integrand $f(x)$ is integrable because f is continuous.

Remark 3.116. We can imagine changing the closed interval $[a,b]$ to something weaker, but bad things happen.

Indeed, the above is true essentially because, for any $\varepsilon > 0$, we can find N such that $n > N$ has

$$\left|\int_a^b [f_n(x) - f(x)]\,dx\right| \le \int_a^b |f_n(x) - f(x)|\,dx < \varepsilon(b-a),$$

so taking $\varepsilon \to 0$ gives what we want.

Now, for our discussion of power series, we start by fixing

$$\sum_{k=0}^{\infty} a_k x^k$$

some power series with finite radius of convergence $R > 0$. Let's say that this function converges to $f(x)$ pointwise on our interval of convergence. Of course, this need not always uniformly converge.

Example 3.117. The power series

$$\sum_{k=0}^{\infty} x^k$$

does not uniformly converge to $\frac{1}{1-x}$ on its interval of convergence. Intuitively, the problem is that the power series explodes next to $x = 1$.

However, we can almost get this.

Proposition 3.118. Fix

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

some power series converging pointwise to $f(x)$ with finite radius of convergence $R > 0$. Then for any $r > 0$ with $r < R$, we have that the power series converges uniformly to $f(x)$ on $[-r,r]$.

Proof. We mostly omit this proof. Essentially the point is that the power series geometrically vanishes, we is fast enough to get our uniform convergence. For concreteness, we use the Weierstrass M-test. The point is that each summand is bounded above by $a_k r^k$, and we know that

$$\sum_{k=0}^{\infty} a_k r^k$$

converges because $r < R$. ■

For now we are interested in differentiating and integrating a power series. Though we have not formally defined the derivative nor integration, so we will just do this term by term, taking

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \overset{d/dx}{\longmapsto} f'(x) := \sum_{k=1}^{\infty} k a_k x^{k-1} = \sum_{k=0}^{\infty} (k+1) a_{k+1} x^k$$

as an example of our differentiation.

For example, we have the following statement.

Proposition 3.119. If the power series f has radius of convergence R, then f' also has the same radius of convergence.

Proof. The point is that, when we write

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \overset{d/dx}{\longmapsto} f'(x) := \sum_{k=0}^{\infty} (k+1)a_{k+1}x^k,$$

we are interested in evaluating

$$\beta := \limsup_{n\to\infty} \sqrt[n]{|(n+1)a_{k+1}|}.$$

However, we can remove the $\sqrt[n]{n+1}$, which converges to 1. So it follows that

$$\limsup_{n\to\infty} \sqrt[n]{|(n+1)a_{k+1}|} = \limsup_{n\to\infty} \sqrt[n]{|a_{k+1}|} = \limsup_{n\to\infty} \sqrt[n]{|a_k|},$$

so the reciprocals here give the radius of convergence equal. ■

Remark 3.120. Something similar works for integration; we won't show this explicitly here.

Essentially this means that the "worst-case" scenario is that the endpoints change when differentiating. This can indeed happen.

Example 3.121. The power series

$$\sum_{k=0}^{\infty} \frac{x^k}{k}$$

has interval of convergence $[-1, 1)$, but the derivative

$$\sum_{k=0}^{\infty} x^k$$

has interval of convergence $(-1, 1)$.

Of course, we are not totally sure that this kind of term-by-term integration is legal. For example, we might ask if

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

has

$$\sum_{k=0}^{\infty} a_k \cdot \frac{r^{k+1}}{k+1} = \int_0^r f(x)\,dx.$$

Well, we do: truncate the power series to the partial sums $f_n \to f$, and then both sides are uniformly converging to the same value, using our discussion from earlier.

3.7.2 Abel's Theorem

There is a last theorem in this section, which is Abel's theorem, but its proof is not very instructive. Regardless, here is the statement.

Theorem 3.122. Fix

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

a power series with finite radius of convergence $R > 0$. If f exists at $x = R$, then f is continuous at $x = R$. Similar holds for $x = -R$.

Example 3.123. Because we know that

$$-\log(1-x) = \sum_{k=1}^{\infty} \frac{x^k}{k}$$

converges at $x = -1$ even though it has radius of convergence 1, we are still able to conclude

$$-\log 2 = \sum_{k=1}^{\infty} \frac{(-1)^k}{k}.$$

Anyways, let's do some examples.

Exercise 3.124. We compute

$$\sum_{n=0}^{\infty} n^2 x^n.$$

Proof. The main idea is to differentiate the series

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

which gives

$$\frac{1}{(1-x)^2} = \sum_{n=1}^{\infty} n x^{n-1}.$$

Now we multiply both sides by x and differentiate again, which gives

$$\sum_{n=1}^{\infty} n^2 x^{n-1} = \frac{x+1}{(1-x)^3},$$

where I have omitted the computation with the quotient rule. Multiplying through by x once more tells us that

$$\sum_{n=1}^{\infty} n^2 x^n = \frac{x(x+1)}{(1-x)^3}.$$

We can also see that our radius of convergence remains 1, and of course we do not converge at the endpoints (say, by the divergence test), so our interval of convergence is $(-1, 1)$. ■

Exercise 3.125. We cannot create a power series for $|x|$.

Proof. The point here is that any power series with positive radius of convergence, we were able to take the derivative termwise, so our power series will be infinitely differentiable (strictly) inside of its radius of convergence. However, $|x|$ is not even differentiable at $x = 0$, so we cannot create a power series from this. ■

3.7.3 Some Closing Remarks

The last section in this chapter shows that any continuous function on a closed interval has a sequence of polynomials uniformly converge to it. This is fun but not central to the story in this course.

One can ask, if $f_n \to f$ is a sequence of integrable functions converging uniformly, then do we have f integrable? The answer is no.

Exercise 3.126. We exhibit $f_n \to f$ converging uniformly on $[1, \infty)$ such that the f_n are integrable while f is not.

Proof. Take the functions $f_n : [1, \infty)$ defined by

$$f_n(x) := \frac{1}{x^{1+1/n}}.$$

This sequence of functions is integrable all over $[1, \infty)$, but the limit function $f(x) := 1/x$ is not. It remains to show that $f_n \to f$ uniformly. Recall that it suffices to look at

$$\lim_{n\to\infty} \sup\{|f_n(x) - f(x)| : x \in [1, \infty)\},$$

so we are interested in the difference

$$f_n(x) - f(x) = \frac{1}{x^{1+1/n}} - \frac{1}{x}.$$

To bound this, we want the maximums and minimums, so we differentiate it, getting

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

We are interested in where this vanishes, which is at $x = \left(1 + \frac{1}{n}\right)^n$. Now note the difference vanishes at $x = 1$ and as $x \to \infty$, so our only candidate to worry about is our critical point, which gives

$$\frac{1}{\left(1 + \frac{1}{n}\right)^{n+1}} - \frac{1}{\left(1 + \frac{1}{n}\right)^n} = \frac{1}{\left(1 + \frac{1}{n}\right)^n} \left(\frac{1}{1 + \frac{1}{n}} - 1 \right).$$

So as $n \to \infty$, this approaches 0 because of the second term in the product. In particular, it follows

$$\lim_{n\to\infty} \sup\{|f_n(x) - f(x)| : x \in [1, \infty)\} = 0,$$

which finishes. ■

THEME 4: DIFFERENTIATION

4.1 November 10

So a midterm ocurred.

4.1.1 Midterm Housekeeping

Here are some notes.

- Some people tried to use the Ratio test on the first question. It is difficult to make such an argument rigorous because the
- The check of the endpoint $x = -2$ on the first question is a bit tricky. The point here is that the subsequence
$$\sum_{n \text{ odd}} \frac{(-2)^n}{n2^n} = \sum_{n \text{ odd}} \frac{(-1)^n}{n}$$
actually diverges even though it looks like it converges: the sum is only over a single sign. Additionally, one must keep track of the positive even terms to make sure they do not counteract the negative divergence.
- People did pretty well on 2(a). On 2(b), many people looked at
$$\lim_{n\to\infty} \sup \left\{ \left| \sin\left(\frac{x}{n}\right) \right| : x \in [0,1] \right\},$$
which is correct. However, some people tried to reverse the limit and the supremum to say that this goes to 0 for free, which is not correct: such an argument would work for $x \in \mathbb{R}$, where the convergence is not uniform.
- For the later questions, some people had trouble keeping track of quantifiers, especially upon negation. For example, the difference between continuity and uniform continuity is subtle but important.
- Professor Sharma expected more people to use sequences on number 3 because this turns discontinuity into a tangible object.
- Professor Sharma is surprised that people approached (a) and (b) different even though they are the same question: the telemetric on $\mathbb{Z}$ induces the same topology on $\mathbb{Z}$ as the usual one.
- The common idea for number 5 was to do something like
$$U = \left\{ (x,y) \in \mathbb{Q}^2 : x^2 < 2 \right\}$$
to disconnect $\mathbb{Q}^2$. However, some people did not prove that U and $\mathbb{Q}^2 \setminus U$ are open, which is not completely obvious. In general, proving a disconnection requires many checks.
- Some people tried to use the fact that the image of a connected set is connected for number 5, which is difficult to rigorize into a proof.
- Keeping track of symbols for number 6 was a bit nontrivial. Namely, we need a third symbol in this proof to make things function, which was not present in all submissions.

This exam was harder than the first one, on average. People had the most trouble with number 5, probably because connectedness is a weird concept. Last time the average was about 80%; this time it was about 75% (about 89/120). It is likely that there will be some minor curve, though how much there is remains unclear.

4.1.2 Derivatives

We are talking about derivatives, for now. We'll do integrals as the last part of this class.

Example 4.1. Fix $1_{\mathbb{Q}} : \mathbb{R} \to \mathbb{Q}$ the $\mathbb{Q}$-indicator. It's not continuous anywhere, which we had a few ways of showing. But can we still compute

$$\int_3^4 f(x)\,dx?$$

The answer will be no, but it's not obvious why or why not because typical calculus (such as the Fundamental theorem of calculus) ideas do not apply.

To attack these kinds of questions, we will need rigor. So that is where we are going.

The first section of chapter 5 reviews derivatives, proving the various derivative rules. The proofs are not terribly interesting, but they are good to "better understand" the derivative. At the very least they require technical skill.

Derivative

Definition 4.2 (Derivative)**.** Fix $f : S \to \mathbb{R}$ a real-valued function. Then we define the *derivative*

$$f'(x) := \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}.$$

If, given $a \in S$, we have $f'(a) \in \mathbb{R}$, then we say that f is *differentiable* at a.

In practice, the limit definition might be annoying to use, or even infeasible. For example, it is reasonable to expect

$$f(x) := \begin{cases} \sqrt{x^4+\sin^4 x}/x & x \neq 0, \\ 0 & x = 0. \end{cases}$$

We expect that $f'(0) = 0$, but this is not obvious. If we could show that the derivative is continuous, we could use the quotient rule and go to 0, but it is not obvious that the derivative is in fact continuous. For a worse example,

$$f(x) = \begin{cases} x^2\sin(1/x) & x \neq 0, \\ 0 & x = 0, \end{cases}$$

is differentiable on all of $\mathbb{R}$ but not continuously differentiable at $x = 0$.

Anyways, let's prove a fairly basic result.

Proposition 4.3. Fix $f : S \to \mathbb{R}$ a function which is differentiable at some point $a \in S$. Then f is continuous at $a \in S$.

Proof. We are given that the limit

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = f'(a) \in \mathbb{R}$$

exists and is finite. We want to show that

$$\lim_{a\to 0} f(x) = f(a).$$

Intuitively, if we lose continuity, then the numerator in the limit for $f'(a)$ will not go to 0 even though the denominator does, which will imply that $f'(a) \notin \mathbb{R}$. We will not make this more rigorous here. ■

Of course, there are continuous functions which are not differentiable.

Example 4.4. The function $f(x) := |x|$ is continuous but not differentiable at 0 because the corresponding limit does not exist.

Example 4.5. The function $f(x) := \sqrt[3]{x}$ is continuous but not differentiable at 0 because the corresponding limit is infinite.

4.1.3 Derivative Rules

Here are some derivative rules.

Proposition 4.6. Fix $f, g : S \to \mathbb{R}$, and suppose that f and g are both differentiable at $a \in S$. Then $(f+g)'(a)$ exists and is equal to $f'(a) + g'(a)$.

Proof. We see that

$$\begin{aligned}(f+g)'(a) &= \lim_{h\to 0} \frac{(f+g)(a+h) - (f+g)(a)}{h} \\ &= \lim_{h\to 0} \frac{f(a+h) - f(x)}{h} + \lim_{h\to 0} \frac{g(a+h) - g(a)}{h} \\ &= f'(a) + g'(a),\end{aligned}$$

where the key step was to split up the sum of limits into the individual limits. ■

Proposition 4.7. Fix $f, g : S \to \mathbb{R}$, and suppose that f and g are both differentiable at $a \in S$. Then $(f+g)'(a)$ exists and is equal to $f'(a)g(a) + f(a)g'(a)$.

Proof. We see that

$$(fg)'(a) = \lim_{h\to 0} \frac{(fg)(a+h) - (fg)a}{h} = \lim_{h\to 0} \frac{f(a+h)g(a+h) - f(a)g(a)}{h}.$$

Now, the key point is to add and subtract $f(a+h)g(a)$ in the numerator, as we did for the corresponding product rule for limits. We get

$$\begin{aligned}(fg)'(a) &= \lim_{h\to 0} \frac{f(a+h)g(a+h) - f(a+h)g(a) + f(a+h)g(a) - f(a)g(a)}{h} \\ &= \lim_{h\to 0} \frac{f(a+h)g(a+h) - f(a+h)g(a)}{h} + \lim_{h\to 0} \frac{f(a+h)g(a) - f(a)g(a)}{h} \\ &= \lim_{h\to 0} f(a+h) \cdot \lim_{h\to 0} \frac{g(a+h) - g(a)}{h} + g(a) \cdot \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} \\ &= f(a)f'(a) + g(a)f'(a).\end{aligned}$$

Importantly, in the last equality we have used the fact that f is continuous at a, which is true because f is differentiable at a. ■

And here is an exercise.

Exercise 4.8. Fix f a function on an open interval continuous on some $a \in \mathbb{R}$. Further, take g defined on an open interval containing $f(a)$. Then we show $g \circ f$ is defined on some open interval containing a.

Proof. Without loss of generality, we assert that g is defined on $(f(a) - \varepsilon, f(a) + \varepsilon)$, for some $\varepsilon > 0$. Then, because f is continuous at a, there exists $\delta > 0$ such that each $x \in \operatorname{dom} f$ has

$$|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon.$$

So we see that $x \in \operatorname{dom} f$ with $x \in (a - \delta, a + \delta)$ implies $f(x) \in (f(a) - \varepsilon, f(a) + \varepsilon)$, which implies that $(g \circ f)(x)$ is well-defined.

To finish, we know there exists δ' such that f is defined on $(a - \delta', a + \delta')$. Then $\delta'' := \min\{\delta, \delta'\}$ will have $x \in (a - \delta'', a + \delta'') \subseteq \operatorname{dom} f$ and $x \in (a - \delta, a + \delta)$, so indeed $(a - \delta'', a + \delta'') \in \operatorname{dom}(g \circ f)$. ■

Remark 4.9. Keeping track of the semantics is annoying but important: it is easy to miss (as I did) the last step of intersecting the interval inside of $\operatorname{dom} f$ with our $(a - \delta, a + \delta)$. This is exacerbated by the fact we usually write continuity omitting the condition $x \in \operatorname{dom} f$.

Next time we will talk about the Mean value theorem.

4.2 November 15

I was not present for class due to the Serge Lang lecture and dinner. I am told we covered the Mean value theorem section and the L'Hôpital's rule section.

4.3 November 17

So I am present for class this time.

4.3.1 L'Hôpital's Rule Warnings

Let's talk a little more about L'Hôpital's rule. Quickly, we recall our indeterminate forms.

Indeterminate forms

Definition 4.10 (Indeterminate forms)**.** Any expression/limit of the form

$$\frac{0}{0}, \frac{\infty}{\infty}, \quad -\cdot\infty, \quad 0^0, \infty^0, 1^\infty, \quad \infty - \infty$$

are called *indeterminate forms.*

Each of the indeterminate forms has a different way of applying L'Hôpital's rule.

- For $\frac{0}{0}$ or $\frac{\infty}{\infty}$, we can apply directly.
- For $0 \cdot \infty$, we rewrite this as $\frac{0}{1/\infty}$, which is now $\frac{0}{0}$.
- For $0^0, \infty^0, 1^\infty$, we call the limit L, and then we are able to default to one of the previous cases with $\log L$.
- For $\infty - \infty$, we call the limit L, and then we are able to go to $\frac{0}{0}$ by looking at $\exp L$.

Example 4.11. We have a variety of ways to compute

$$\lim_{x\to 1} \frac{\sqrt{x}-1}{x-1}.$$

We could multiply the top and bottom by $\sqrt{x}+1$. But if we change the numerator to something more complicated, such tricks become difficult; L'Hôpital's rule is the way to finish.

Here are some (non-)examples.

Exercise 4.12. We compute the limit

$$\lim_{x\to\infty} \frac{x-\sin x}{x}.$$

Proof. This limit directly goes to $\frac{\infty}{\infty}$, so we could try to apply L'Hôpital's rule, but we get

$$\lim_{x\to\infty} \frac{1-\cos x}{1},$$

which does not exist. However, the limit does actually exist: we can split up the limit as

$$\lim_{x\to\infty} \left(1-\frac{\sin x}{x}\right),$$

and now this goes to $1-0=\boxed{1}$. ■

The point of the above discussion is to remind ourselves that L'Hôpital's rule only applies when the latter limit does not exist.

Exercise 4.13. We compute the limit

$$\lim_{x\to\infty} \frac{e^{4x}-e^{-x}}{e^{3x}+e^{2x}}.$$

Proof. We could apply L'Hôpital's rule, but it will never terminate. To make this actually possible, we write

$$\frac{e^{4x}-e^{-x}}{e^{3x}+e^{2x}} = \frac{e^{x}-e^{-4x}}{1+e^{-x}},$$

which we can see goes to $\boxed{+\infty}$ as $x\to\infty$. ■

The point of the above exercise is that sometimes L'Hôpital's rule will not always directly apply.

Exercise 4.14. We compute

$$\lim_{x\to\infty} x^{\sin(1/x)}.$$

Proof. Our indeterminate form is ∞^0. So we set the limit equal to L, and we find

$$\log L = \lim_{x\to\infty} \sin\left(\frac{1}{x}\right)\log x.$$

We would like to differentiate $\log x$, so we keep it quarantined and move the $\sin\left(\frac{1}{x}\right)$ to the denominator, giving

$$\log L = \lim_{x\to\infty} \frac{\log x}{1/\sin\left(\frac{1}{x}\right)}.$$

Now the form is $\frac{\infty}{\infty}$, so we use L'Hôpital's rule, which gives

$$\log L = \lim_{x\to\infty} \frac{1/x}{-1/\sin\left(\frac{1}{x}\right)^2 \cdot \cos\left(\frac{1}{x}\right) \cdot -1/x^2}.$$

This collapses down to

$$\log L = \lim_{x\to\infty} x \sin\left(\frac{1}{x}\right) \tan\left(\frac{1}{x}\right).$$

We could do this by brute force, but it is a bit more efficient to write this as

$$\log L = \lim_{x\to\infty} \frac{\sin(1/x)}{1/x} \cdot \tan\left(\frac{1}{x}\right),$$

which will got $1 \cdot 0 = 0$, so $L = \boxed{1}$. ■

Exercise 4.15. We compute

$$\lim_{x\to 0}(1+2x)^{1/x}.$$

Proof. The form here is 1^∞, so we set the limit equal to L and compute

$$\log L = \lim_{x\to 0} \frac{\log(1+2x)}{x}.$$

Applying L'Hôpital's rule, we get

$$\log L = \lim_{x\to 0} \frac{\frac{2}{1+2x}}{1} = \lim_{x\to 0} \frac{2}{1+2x} = 2.$$

So it follows $L = \boxed{e^2}$. ■

As a further remark, we note that applying L'Hôpital's rule to

$$\lim_{x\to a} \frac{f(x)}{g(x)}$$

needs f and g to be differentiable in a neighborhood around a and g' to be nonzero in a neighborhood around a. This last condition might appear awkward, but it is necessary. Normally this isn't a problem due to, say, smoothness, but it does prevent us from applying the rule for, say,

$$\lim_{x\to\infty} \frac{x - \sin x}{x}$$

as we saw earlier. Essentially these conditions go into the assertion that

$$\lim_{x\to a} \frac{f'(x)}{g'(x)}$$

actually exists.

Exercise 4.16. We set $f(x) = x + \cos x \cdot \sin x$ and $g(x) = e^{\sin x} f(x)$. We compute

$$\lim_{x\to\infty} \frac{f(x)}{g(x)}.$$

Proof. Each of f and g are nonzero in a neighborhood around 0 by, say, continuity. We can also compute the derivatives as

$$f'(x) = x + \cos x \cdot \cos x - \sin x \cdot \sin x = 2\cos^2 x$$
$$g'(x) = e^{\sin x} \cdot \cos x[2\cos x + f(x)]$$

after applying some elbow grease. In particular, we see

$$\frac{f'(x)}{g'(x)} = \frac{2e^{-\sin x} \cdot \cos x}{2\cos x + f(x)}$$

so that the numerator is bounded and the denominator goes to infinity, giving 0 as $x \to \infty$.

However, the original limit is of $\frac{f}{g} = e^{-\sin x}$, so this limit does not actually exist. What went wrong in the above example is that the limit of $\frac{f'}{g'}$ does not actually exist because g' vanishes arbitrarily close to $+\infty$. So it goes. ■

4.3.2 Taylor's Theorem

Our last topic on derivatives is Taylor's theorem. Again there will be strange convergence issues to keep track of, so we will want to study this a bit closer. Here is our definition.

Taylor series

Definition 4.17 (Taylor series)**.** Fix $f : \mathbb{R} \to \mathbb{R}$ infinitely differentiable at $x = a$. Then we define the infinite *Taylor series* (formally, say) by

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)(a)}{a!}(x-a)^k.$$

We also define our remainder term by

$$R_n(x) = f(x) - \sum_{k=0}^{n-1} \frac{f^{(k)}(c)}{k!}(x-c)^k.$$

Exercise 4.18. We check the remainder for

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k.$$

Proof. We note that for $|x| < 1$, we have

$$R_n(x) = \sum_{k=n}^{\infty} x^k = x^n \cdot \frac{1}{1-x},$$

which will go to 0 as $n \to \infty$, as it should ■

Most functions that we like will have the Taylor remainder term approach 0 as long as the series converges. The typical way to check this is as follows.

Proposition 4.19. Suppose that $f : \mathbb{R} \to \mathbb{R}$ is nth differentiable. Then, fixing some x and c, there exists d between x and c such that

$$R_n(x) = \frac{(x-c)^n}{n!} f^{(n)}(d).$$

The point here is that oftentimes we can bound $f^{(n)}(d)$ in such a way that we can promise $n!$ will dominate.

Example 4.20. For $f(x) = e^{2x}$ take $c = 0$ and $x = 1$. Then $f^{(n)}(d) \le 2^n$ for all d between x and c, so $R_n(x) \to 0$ as $n \to \infty$, and our Taylor series converges correctly.

And here is the warning to correspond to our hopes.

Exercise 4.21. We consider the Taylor series for

$$f(x) = \begin{cases} e^{-1/x^2} & x \neq 0, \\ 0 & x = 0. \end{cases}$$

In particular, the Taylor series does not converge to f for any $x \neq 0$.

Proof. We can check that $f : \mathbb{R} \to \mathbb{R}$ is continuous, and it is in fact infinitely differentiable, even at 0. The main point to differentiability at 0 is that, for each $n \in \mathbb{N}$, there exists $p_n \in \mathbb{Z}[x]$ such that

$$f^{(n)}(x) = p_n(1/x)e^{-1/x^2},$$

which we can show by an induction. But this vanishes as $x \to 0$, so $f^{(n)}(0) = 0$ for each $n \in \mathbb{N}$. It follows that the Taylor series is the zero series, which is not equal to f for any $x \neq 0$. ■

Remark 4.22. More viscerally, we can see that the derivatives of f at particular values x around 0 are growing at a factorial rate. Essentially, the $p_n(x)$ will accumulate leading coefficients at a combinatorial speed, which is something we can see from direct expansion.

4.3.3 The Binomial Theorem

Let's spend a moment discussing the Binomial theorem. We have the following.

Proposition 4.23. Fix $\alpha \in \mathbb{R}$. We claim that, for $|x| < 1$,

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^\alpha,$$

where

$$\binom{\alpha}{n} := \prod_{k=0}^{n-1} \frac{\alpha - k}{n - k}.$$

Remark 4.24. It is not hard to verify by hand that $\alpha \in \mathbb{N}$ causes $\binom{\alpha}{n}$ to match what we want it to. I won't write this our because it is just a matter of writing down the formula and staring.

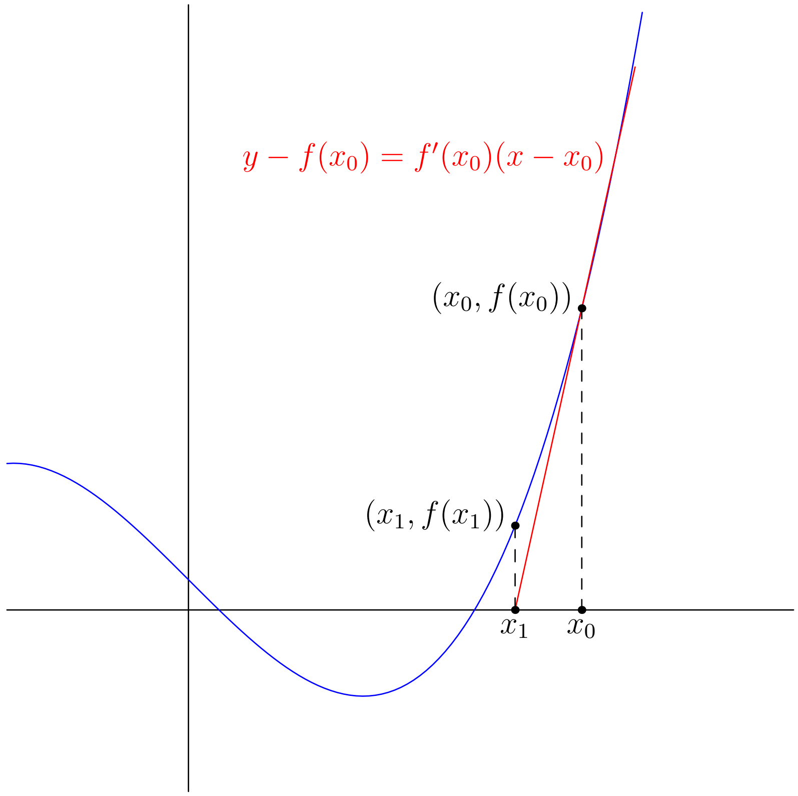
Proof. We can check the convergence of this by the Ratio test. We find that

$$\frac{\binom{\alpha}{k+1}x^{k+1}}{\binom{\alpha}{k}x^k} = \frac{\alpha - k}{k} \cdot x,$$

so we get convergence for $|x| < 1$. In fact, we observe that this sort of argument can be used to verify that the remainder term from $(1+x)^\alpha$ vanishes as the number of terms in the series goes to infinity. ■

4.3.4 Newton's Method

The idea of Newton's method is to find roots of a differentiable function f, and we want to find a root. So we start with a guess x_0, and to get closer, we draw the tangent line at 0 and find the root. Namely, f is hopefully locally linear, so we can hope to get close to the root by a linear approximation. Here is the image.



Expanding this out by hand, our recursion is

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

This is a recursion we could hopefully do by computer. The rest of §31 can be omitted. It is quite technical and a lot about approximations, which we tend to not care about in this class.

Anyways, let's do an exercise.

Exercise 4.25 (Ross 31.4(a)). Fix $a, b \in \mathbb{R}$ such that $a < b$. Then find f a function infinitely differentiable so that $f(x) = 0$ for $x \leq 0$ and $f(x) > 0$ for $x > 0$.

Proof. This is somewhat subtle. For example,

$$f(x) \stackrel{?}{=} \begin{cases} x^2 & x \geq 0, \\ 0 & x \leq 0, \end{cases}$$

will not work because this function is not twice-differentiable at 0. So we need a function with lots of vanishing derivatives at 0 even though it is not zero, so borrow our sad example from the Taylor series section: we define

$$f(x) \stackrel{?}{=} \begin{cases} e^{-1/x} & x > 0, \\ 0 & x \leq 0. \end{cases}$$

The point is that we computed $f^{(n)}(0) = 0$ for all $n \in \mathbb{N}$, even as we approach from the right, so this function will be differentiable at 0 and hence infinitely differentiable everywhere. ■

THEME 5: INTEGRATION

5.1 November 22

The fun continues.

5.1.1 Darboux Sums

We are going to talk about integration now.

Example 5.1. It is difficult to compute the integral

$$\int_2^3 1_{\mathbb{Q}}(x)\,dx.$$

In fact, it does not exist, but we need to know what the integral means.

The point is that are going to more rigorously define what integration is for these questions to be tractable. For now, we don't even know what being integrable means, so let's move towards that.

Most of our work will be for bounded functions f defined on a closed interval $[a, b]$. We have the following convenient definitions.

Maximum and minimum

Definition 5.2 (Maximum and minimum)**.** Fix $f : [a, b] \to \mathbb{R}$ a bounded function. Then we define

$$M(f, S) := \sup\{f(x) : x \in S\} \qquad \text{and} \qquad m(f, S) := \inf\{f(x) : x \in S\}$$

for any $S \subseteq [a, b]$.

Note that these exist because f is bounded. Now, the idea is to "break" $[a, b]$ into n pieces, for some $n \in \mathbb{Z}^+$.

Partition

Definition 5.3 (Partition)**.** Fix $a, b \in \mathbb{R}$ with $a < b$. Then we define a *partition* of $[a, b]$ into n pieces to be any increasing sequence $\mathcal{P} = \{t_k\}_{k=0}^n$ such that $t_0 = a$ and $t_n = b$ so that

$$a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b.$$

In particular, a partition $\{t_k\}_{k=0}^n$ of $[a, b]$ creates the union

$$[a, b] = \bigcup_{k=0}^{n-1} [t_k, t_{k+1}].$$

This gives us the following definition.

Darboux sums

Definition 5.4 (Darboux sums)**.** Fix $f : [a,b] \to \mathbb{R}$ a bounded function and $\mathcal{P} = \{t_k\}_{k=0}^n$ some partition of $[a,b]$. Then we define the following.

- We define the *upper Darboux sum* by

$$\mathcal{U}(f,\mathcal{P}) = \sum_{k=1}^n M(f,[t_{k-1},t_k]) \cdot (t_k - t_{k-1}).$$

- We define the *lower Darboux sum* by

$$\mathcal{L}(f,\mathcal{P}) = \sum_{k=1}^n m(f,[t_{k-1},t_k]) \cdot (t_k - t_{k-1}).$$

Remark 5.5. These are in spirit a generalization of Riemann sums to arbitrary partitions: partition and then average over values.

Intuitively, the upper Darboux sum is the sum of the "largest possible rectangle" of a corresponding Riemann sum. Note that $\mathcal{U}(f,\mathcal{P})$ is upper-bounded by

$$\mathcal{U}(f,\mathcal{P}) \le \sum_{k=1}^n M(f,[a,b]) \cdot (t_k - t_{k-1}) = M(f,[a,b]) \cdot (b-a).$$

In fact, we can lower-bound this stupidly by

$$\mathcal{U}(f,\mathcal{P}) \ge \sum_{k=1}^n m(f,[a,b]) \cdot (t_k - t_{k-1}) = m(f,[a,b]) \cdot (b-a).$$

Similarly, $\mathcal{L}(f,\mathcal{P})$ is lower-bounded by

$$\mathcal{L}(f,\mathcal{P}) \ge \sum_{k=1}^n m(f,[a,b]) \cdot (t_k - t_{k-1}) = m(f,[a,b]) \cdot (b-a),$$

and is upper-bounded by

$$\mathcal{L}(f,\mathcal{P}) \le \sum_{k=1}^n M(f,[a,b]) \cdot (t_k - t_{k-1}) = M(f,[a,b]) \cdot (b-a).$$

This gives us the following definition.

Darboux integrals

Definition 5.6 (Darboux integrals)**.** Fix $f : [a,b] \to \mathbb{R}$ a bounded function. Then we define the following.

- The *upper Darboux integral* is

$$\mathcal{U}(f) := \inf\{\mathcal{U}(f,\mathcal{P}) : \mathcal{P} \text{ partitions } [a,b]\}.$$

- The *lower Darboux integral* is

$$\mathcal{L}(f) := \sup\{\mathcal{L}(f,\mathcal{P}) : \mathcal{P} \text{ partitions } [a,b]\}.$$

Note that these numbers do exist because we showed that $\mathcal{U}(f,\mathcal{P})$ and $\mathcal{L}(f,\mathcal{P})$ always live in the interval $[m(f,[a,b])(b-a), M(f,[a-b])(b-a)]$.

5.1.2 Darboux Integrability

We would hope that $\mathcal{U}(f) \geq \mathcal{L}(f)$. Certainly, because a supremum will be at least an infimum, we can say that $\mathcal{U}(f,\mathcal{P}) \geq \mathcal{L}(f,\mathcal{P})$, but this does not actually imply $\mathcal{U}(f) \geq \mathcal{L}(f)$. This is not obvious. To help with this, we pick up the following technical lemma.

Lemma 5.7. Fix $f : [a,b] \to \mathbb{R}$ a bounded function. Suppose $\mathcal{P} \subseteq \mathcal{Q}$ are partitions of $[a,b]$. Then

$$\mathcal{L}(f,\mathcal{P}) \leq \mathcal{L}(f,\mathcal{Q}) \qquad \text{and} \qquad \mathcal{U}(f,\mathcal{Q}) \leq \mathcal{U}(f,\mathcal{P}).$$

Proof. Only the inequalities on the ends require comment, and we will only discuss one on the left because the right inequality is similar. Intuitively, this is because further subdividing an interval lets us increase an infimum on one side.

To be explicit, say $Q = \{t_k\}_{k=0}^n$ so that $P = \{t_{\sigma\ell}\}_{\ell=0}^m$ for some strictly increasing σ with $\sigma(0) = 0$ and $\sigma(m) = n$. Then the point is that

$$\sum_{k=\sigma(\ell)}^{\sigma(\ell+1)} m(f,[t_{\sigma\ell},t_{\sigma(\ell+1)}])(t_{\sigma(\ell+1)} - t_{\sigma\ell}) \geq \sum_{k=\sigma(\ell)}^{\sigma(\ell+1)} m(f,[t_{\sigma\ell},t_{\sigma(\ell+1)}])(t_{\sigma(\ell+1)} - t_{\sigma\ell})$$

by the same lower Darboux lower-bounding as before. Summing this over all ℓ gives the result. ■

And here is our result.

Proposition 5.8. We have that $\mathcal{L}(f) \leq \mathcal{U}(f)$.

Proof. We show that any lower Darboux sum is less or equal to any upper Darboux sum. Indeed, if P and Q are partitions of $[a,b]$, we get

$$\mathcal{L}(f,\mathcal{P}) \leq L(f,\mathcal{P}\cup\mathcal{Q}) \leq \mathcal{U}(f,\mathcal{P}\cup\mathcal{Q}) \leq \mathcal{U}(f,\mathcal{Q})$$

by simply applying the above lemma repeatedly. ■

So what about equality?

Integrable

Definition 5.9 (Integrable)**.** Fix $f : [a,b] \to \mathbb{R}$ a bounded function. We say that f is *Darboux integrable* if and only if $\mathcal{L}(f) = \mathcal{U}(f)$.

This condition is a bit unwieldy: it's got partitions and supremum and infimum of lots of sets floating around. Let's try to reduce the number of quantifiers.

Proposition 5.10. Fix $f : [a,b] \to \mathbb{R}$ a bounded function. Then f is Darboux integrable if and only if for all $\varepsilon > 0$ there exists a partition P such that

$$\mathcal{U}(f,\mathcal{P}) - \mathcal{L}(f,\mathcal{P}) < \varepsilon.$$

Proof. We show the directions one at a time.

- Suppose the conclusion. Then, for any $\varepsilon > 0$, we can find a partition $\mathcal{P}$ such that

  $$\mathcal{L}(f,\mathcal{P}) \leq \mathcal{L}(f,\mathcal{P}) \leq \mathcal{U}(f) \leq \mathcal{U}(f,\mathcal{P}).$$

  Sending $\varepsilon \to 0$ shows that f is integrable.

- Otherwise take f integrable. In order to not break the supremum and infimum things, there must be partitions $\mathcal{P}$ and $\mathcal{Q}$ such that
$$0 \le \mathcal{L}(f) - \mathcal{L}(f, \mathcal{P}) < \frac{\varepsilon}{2} \qquad \text{and} \qquad 0 \le \mathcal{U}(f, \mathcal{Q}) - \mathcal{U}(f) < \varepsilon 2.$$
Taking $P \cup Q$, we see, using the technical lemma again,
$$0 \le \mathcal{L}(f) - \mathcal{L}(f, \mathcal{P} \cup \mathcal{Q}) < \frac{\varepsilon}{2} \qquad \text{and} \qquad 0 \le \mathcal{U}(f, \mathcal{P} \cup \mathcal{Q}) - \mathcal{U}(f) < \varepsilon 2,$$
so $\mathcal{P} \cup \mathcal{Q}$ will be the needed partition. ■

We remark that it is general somewhat hard to actually test if a function is Darboux integrable. We have a definition, but actually computing these various values is difficult because there are so many quantifiers to keep track of. This might be technically easier to work with, but it is difficult to work with.

Let's try to make this more computationally feasible.

Definition 5.11. For a partition $\{t_k\}_{k=0}^n \subseteq [a, b]$, then we define the *mesh* of $\{t_k\}_{k=0}^n$ to be the maximum of $t_{k+1} - t_k$.

We have the following result.

Proposition 5.12. Fix $f : [a, b] \to \mathbb{R}$ a bounded function. Then f is integrable if and only if, for each $\varepsilon > 0$, there exists a $\delta > 0$ such that any partition $\mathcal{P}$ with mesh less than δ with
$$\mathcal{U}(f, \mathcal{P}) - \mathcal{L}(f, \mathcal{P}) < \varepsilon.$$

Proof. This is technically hard to prove, so we won't give it here. The main point is that, once our partitions are close enough together, we can make the mesh smaller artificially to get the condition. In the reverse condition, we can do this by hand. ■

5.1.3 Riemann Integrability

Because we should, let's talk about Riemann sums.

Definition 5.13. Fix $f : [a, b] \to \mathbb{R}$ a bounded function. Then, given a partition $\{t_k\}_{k=0}^n$ of $[a, b]$, we define the *Riemann sum* to be
$$\sum_{k=1}^n f(x_k)(t_k - t_{k-1}),$$
where $x_k \in [t_{k-1}, t_k]$ is any point.

Usually we place some conditions on x_k and t_k, but we don't have to. For example, we might require the $[t_{k-1}, t_k]$ to have equal length, but there is no immediate reason to do this. We note quickly that the Riemann sum will have
$$m(f, [t_k, t_{k-1}]) \le f(x_k) \le M(f, [t_k, t_{k-1}]),$$
so the Riemann sum will be upper-bounded by $\mathcal{U}(f)$ and lower-bounded by $\mathcal{L}(f)$. So intuitively, it looks like Darboux integrable will easily imply Riemann integrable.

So let's define Riemann integrable.

Riemann integrable

Definition 5.14 (Riemann integrable)**.** Fix $f : [a, b] \to \mathbb{R}$ a bounded function. Then f is *Riemann integrable* if and only if, for each $\varepsilon > 0$, there exists $\delta > 0$ such that each partition $\mathcal{P}$ with mesh less than δ has any Riemann sum ε-close to some $I \in \mathbb{R}$. Note that $I \in \mathbb{R}$ depends only on f and $[a, b]$.

This happens to be equivalent to Darboux integrable, but we will not show it here in detail.

Theorem 5.15. Fix $f : [a,b] \to \mathbb{R}$ a bounded function. Then f is Darboux integrable if and only if it is Riemann integrable.

Proof. We talk about these one direction at a time.

- Fix f Darboux integrable. Then we claim that the integral needed is $I := \mathcal{U}(f) = \mathcal{L}(f)$. Then, for any Riemann sum
$$\mathcal{L}(f,\mathcal{P}) \le \sum_{k=1}^{n} f(x_k)(t_k - t_{k-1}) \le \mathcal{U}(f,\mathcal{P}),$$
so sending our mesh to 0 takes $\mathcal{L}(f,\mathcal{P}),\mathcal{U}(f,\mathcal{P}) \to I$.
- The reverse direction is annoying; essentially the point is that we can get a Riemann sum arbitrarily close to any particular $\mathcal{L}(f,\mathcal{P})$ or $\mathcal{U}(f,\mathcal{P})$ by choosing the x_ks to be close to the supremum and infimum of its interval. However, this is annoying because it is possible that the supremum and infimum are not actually achieved, but it is possible. ■

Anyways, let's do an example.

Exercise 5.16. We show that the function
$$f(x) = \begin{cases} x^2 & x \in \mathbb{Q}, \\ 0 & x \notin \mathbb{Q}, \end{cases}$$
is not integrable on $[0,1]$ by computing $\mathcal{L}(f)$ and $\mathcal{U}(f)$.

Proof. Fix $\mathcal{P} = \{t_k\}_{k=0}^n$ to be some partition of $[0,1]$. Then we see that
$$\mathcal{L}(f,\mathcal{P}) = \sum_{k=1}^{n} m(f,[t_{k-1},t_k])(t_k - t_{k-1}) = 0$$
because each interval will have some irrational in it.

The upper Darboux integral is more annoying. We find that
$$\mathcal{U}(f,\mathcal{P}) = \sum_{k=1}^{n} M(f,[t_{k-1},t_k])(t_k - t_{k-1}) = \sum_{k=1}^{n} t_k^2(t_k - t_{k-1}).$$
Indeed, that supremum is t_k^2 because this certainly upper-bounds, and it is the least upper bound by taking some sequence of rationals in $[t_{k-1},t_k]$ approaching t_k.

The main idea here is that $\mathcal{U}(f,\mathcal{P})$ is going to have the same upper Darboux sum as $g(x) = x^2$, but now g is continuous and hence integrable, so we can compute it directly as, say, $\int_0^1 x^2\,dx = \frac{1}{3}$. Because $\mathcal{U}(f) \ne \mathcal{L}(f)$, this is not integrable. So we are done. ■

Remark 5.17. One could also compute the integral of g by hand using some summation.

5.2 November 29

Here we go again.

5.2.1 Integrability Conditions

We're still studying integrability. Last time we defined integrability and gave some easier conditions for being integrable without actually computing Darboux integrals, and we will continue with that story today.

Proposition 5.18. Fix $f : [a, b] \to \mathbb{R}$ is a monotonic function. Then f is integrable on $[a, b]$.

Before we go into proving this, we should point out that monotonic functions are pretty nice. Granted, they aren't terrible—a monotonic function will have only many countably many discontinuities. Though "countably many" is a bit weak of an assertion, for any given countable set has a corresponding monotonic function with discontinuities on that set.

Proof. We are going to need to somewhat start from scratch. Take f increasing so that our Darboux sums are, for a given partition $P = \{t_k\}_{k=0}^n \subseteq [a, b]$,

$$U(f, P) = \sum_{k=1}^{n} M(f, [t_{k-1}, t_k])(t_k - t_{k-1}) = \sum_{k=1}^{n} f(t_k)(t_k - t_{k-1}),$$

and

$$L(f, P) = \sum_{k=1}^{n} m(f, [t_{k-1}, t_k])(t_k - t_{k-1}) = \sum_{k=1}^{n} f(t_{k-1})(t_k - t_{k-1}).$$

Now we notice that

$$U(f, P) - L(f, P) = \sum_{k=1}^{n} \big(f(t_k) - f(t_{k-1})\big)(t_k - t_{k-1}).$$

Fix $\varepsilon > 0$, and set $\delta := \frac{\varepsilon}{f(b)-f(a)}$ so that, if the mesh of P is less than δ, we have

$$U(f, P) - L(f, P) < \sum_{k=1}^{n} \big(f(t_k) - f(t_{k-1})\big) \cdot \frac{\varepsilon}{f(b) - f(a)} = \varepsilon,$$

where we have telescoped to evaluate the sum. This shows that f is integrable. ■

Proposition 5.19. Fix $f : [a, b] \to \mathbb{R}$ is a continuous function. Then f is integrable on $[a, b]$.

Proof. Take f continuous so that our Darboux sums are, for a given partition $P = \{t_k\}_{k=0}^n \subseteq [a, b]$,

$$U(f, P) = \sum_{k=1}^{n} M(f, [t_{k-1}, t_k])(t_k - t_{k-1}),$$

and

$$L(f, P) = \sum_{k=1}^{n} m(f, [t_{k-1}, t_k])(t_k - t_{k-1}).$$

Now we notice that

$$U(f, P) - L(f, P) = \sum_{k=1}^{n} \big(M(f, [t_{k-1}, t_k]) - m(f, [t_{k-1}, t_k])\big)(t_k - t_{k-1}).$$

Fix $\varepsilon > 0$, and we would like small mesh to make the above difference less than ε. Now, f is continuous on $[a, b]$, so it is uniformly continuous (this is the key trick!), so we may find $\delta > 0$ such that

$$|x_1 - x_2| < \delta \implies |f(x_1) - f(x_2)| < \frac{\varepsilon}{b - a}.$$

In particular, if the mesh of P is less than our δ, we find

$$M(f,[t_{k-1},t_k]) - m(f,[t_{k-1},t_k] < \frac{\varepsilon}{b-a}$$

because $|t_k - t_{k-1}| < \delta$, where we are using the fact that f achieves its maximum and minimum at some x_1 and x_2, so $|x_1 - x_2| < \delta$ gives the result. Thus,

$$U(f,P) - L(f,P) < \sum_{k=1}^{n} \frac{\varepsilon}{2(b-a)} \cdot (t_k - t_{k-1}) = \varepsilon,$$

where we have telescoped to evaluate the sum. This shows that f is integrable. ■

Remark 5.20. Intuitively, thinking of continuous as "locally monotone" and building up continuity from the monotone result is probably fine. Of course, this is not technically fine because

Remark 5.21. It is somewhat impressive that the notion of uniform continuity came up crucially in the above proof. Namely, it is explicitly needed because we need the same δ to work over the entire interval.

5.2.2 Integral Rules

Let's talk through some integral rules. Here are the main statements.

Proposition 5.22. Fix f and g integrable functions $[a,b] \to \mathbb{R}$ with $c\mathbb{R}$.

(a) $\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx$.

(b) $\int_a^b (f+g)(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$.

Proof. Omitted. The main idea is to imitate the proof of the corresponding limit laws. ■

Similarly, here is a bounding result.

Proposition 5.23. Fix integrable functions $f,g:[a,b]\to\mathbb{R}$ such that $f(x) \ge g(x)$ for each $x \in [a,b]$. Then

$$\int_a^b f(x)\,dx \ge \int_a^b g(x)\,dx.$$

Proof. The main idea is to compare the Darboux sums by hand. To make this technically easier, we show that

$$\int_a^b (f-g) \ge 0.$$

Namely, writing down a Darboux sum, we find that, for a given partition $P \subseteq [a,b]$, we find that

$$U(f-g,P) = \sum_{k=1}^{n} \underbrace{M(f-g,[t_{k-1},t_k])}_{\ge 0}(t_k - t_{k-1}) \ge 0,$$

so it follows $U(f) = \inf_P U(f-g,P) \ge 0$, which is what we need because we know $f-g$ is integrable. ■

5.2.3 Pushing Integrability

Integrable functions tend to be pretty well-behaved. We have the following definition.

Piecewise continuous

Definition 5.24 (Piecewise continuous)**.** A function $f : [a,b] \to \mathbb{R}$ is said to be *piecewise continuous* if and only if there is a partition $\{t_k\}_{k=0}^n$ of $[a,b]$ such that f is uniformly continuous (!) on each (t_{k-1}, t_k).

In theory, we'd like to say that f is continuous on each $[t_{k-1}, t_k]$ (giving uniform continuity), but continuity on two sets implies continuity on their union, so this would make piecewise continuity imply continuity. More concretely, we would like functions like

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

for $f : [0,1] \to \mathbb{R}$ to be piecewise continuous, but there is no partition of $[0,1]$ will fix the problem at 0.

As for why we want uniform continuity, the intuition is that it imitates continuity. More formally, we want f restricted to (t_{k-1}, t_k) to extend to a continuous function on $[t_{k-1}, t_k]$ to really earn the name "piecewise." More concretely, we are trying to prevent infinite discontinuities.

Similarly, we have the following result.

Piecewise montonic

Definition 5.25 (Piecewise montonic)**.** A function $f : [a,b] \to \mathbb{R}$ is said to be *piecewise monotonic* if and only if there is a partition $\{t_k\}_{k=0}^n$ of $[a,b]$ such that f is monotonic on each (t_{k-1}, t_k).

Again, we want the open intervals here for approximately the same reason as before: functions such as

$$f(x) = \begin{cases} x & x < 0, \\ x+1 & x \geq 0, \end{cases}$$

with bad jumps would not be piecewise monotonic.

Warning 5.26. Piecewise monotonic does not imply bounded. For example, a patched version of $\tan$.

These covers most functions we care about. Of course, it does not cover all of them.

Non-Example 5.27. The function

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

is continuous on $\mathbb{R}$ but not locally monotone at 0. The function $f(x) + \frac{x}{2}$ even has $f'(x) > 0$. The function $\int_0^x f(t)\,dt$ is even continuously differentiable at 0 but not locally monotone.

However, it is true that, if f is continuously differentiable, then $f'(x) > 0$ ($f'(x) < 0$) implies that f is locally increasing (decreasing) around x. Importantly, this assertion is agnostic about $f'(x) = 0$, as discussed in the above example.

Anyways, we have the following result.

Proposition 5.28. Fix $f : [a,b] \to \mathbb{R}$ which is piecewise continuous or bounded and piecewise monotonic. Then f is integrable.

Proof. We take these one at a time.

- Because f is piecewise continuous, find our partition $\{t_k\}_{k=0}^n$ such that f is uniformly continuous on each (t_{k-1}, t_k). Then extend f to a continuous function $[t_{k-1}, t_k]$, and we note that f is bounded and hence integrable here automatically. Then taking the union of these integrals will finish.

- Because f is piecewise monotonic, find a partition $\{t_k\}_{k=0}^n$ such that f is monotonic on each (t_{k-1}, t_k). Then extend f to a monotonic function $[t_{k-1}, t_k]$ by taking

$$f(t_{k-1}) := \lim_{t \to t_{k-1}^+} f(t) \qquad \text{and} \qquad f(t_{k-1}) := \lim_{t \to t_k^-} f(t)$$

  Now, we note that f is bounded and hence integrable here automatically. Then taking the union of these integrals will finish. ∎

5.2.4 Some Extra Bits

This is a result we might care about.

Theorem 5.29 (Intermediate value for integrals)**.** Fix f continuous on $[a, b]$. Then there exists $c \in (a, b)$ such that

$$f(c) = \frac{1}{b-a} \int_a^b f(t)\, dt.$$

Proof. Roughly speaking, this comes from using the intermediate value theorem on

$$F(x) := \int_a^x f(t)\, dt$$

to finish. ∎

Ross also mentions the Dominated convergence theorem, but we don't care.

Anyways, let's jump into some exercises.

Step **Definition 5.30** (Step)**.** A function $f : [a, b] \to \mathbb{R}$ is said to be *step* if and only if there is a partition $\{t_k\}_{k=0}^n$ of $[a, b]$ such that f is constant on each (t_{k-1}, t_k).

We note that step functions are piecewise monotonic and hence integrable.

5.2.5 Fundamental Theorem of Calculus

Here is the first part.

Theorem 5.31 (Fundamental theorem of calculus, I)**.** Fix f a continuous function on $[a, b]$ which is differentiable on (a, b) such that g' is integrable on $[a, b]$. Then

$$\int_a^b g'(x)\, dx = g(b) - g(a).$$

We remark that the typical requirement is that g' is continuous, but we have weakened it to integrable. Granted, I am not sure how to create a function whose derivative is bad enough to not be integrable.

Remark 5.32. It is true that an integrable function must be continuous somewhere.

5.3 December 1

5.3.1 Fundamental Theorem of Calculus

Last time we were discussing the Fundamental theorem of calculus; here is the first part.

Theorem 5.33 (Fundamental theorem of calculus, I)**.** Fix f a continuous function on $[a,b]$ which is differentiable on (a,b) such that g' is integrable on $[a,b]$. Then

$$\int_a^b g'(x)\,dx = g(b) - g(a).$$

Proof. At the very least we know that

$$I := \int_a^b g'(x)\,dx \in \mathbb{R}.$$

In particular, for any fixed partition $P = \{t_k\}_{k=0}^n \subseteq [a,b]$ with sufficiently small mesh $\delta > 0$, we have that $U(g',P) - L(g',P) < \varepsilon$.

Now, the trick is to apply the Mean value theorem on each $[t_{k-1}, t_k]$ so that there exists $x_k \in (t_{k-1}, t_k)$ such that

$$\frac{g(t_k) - g(t_{k-1})}{t_k - t_{k-1}} = g'(x_k).$$

This is going to be our rigorization of g being "locally linear." So we see

$$g(b) - g(a) = \sum_{k=1}^n (g(t_k) - g(t_{k-1})) = \sum_{k=1}^n g(x_k)(t_k - t_{k-1})$$

which is a Riemann sum for g. In particular, we see that

$$L(g',P) \le g(b) - g(a) \le U(g',P).$$

So we find that

$$\varepsilon < L(g',P) - U(g',P) \le L(g',P) - I \le g(b) - g(a) - I \le U(g',P) - I \le U(g',P) < \varepsilon,$$

so sending ε to 0 will show the desired equality. ∎

Remark 5.34. It might appear strange that we are assuming g' is integrable on $[a,b]$, but this is indeed necessary because there exists functions with derivatives which are not integrable. Concrete examples are difficult.

And here is the second part of the Fundamental theorem of calculus.

Theorem 5.35 (Fundamental theorem of calculus, II)**.** Fix $f : [a,b] \to \mathbb{R}$ an integrable function. Then

$$F(x) := \int_a^x f(t)\,dt$$

is a continuous function $[a,b] \to \mathbb{R}$. If f is continuous at some $x \in (a,b)$, then F is differentiable at $x \in (a,b)$ with derivative $F'(x) = f(x)$.

Proof. We show the claims one at a time.

- Fix some $x_0 \in [a,b]$ and some $\varepsilon > 0$. The key is that f is bounded, say by some $M \in \mathbb{R}$ so that $|f(x)| < M$ for each $x \in [a,b]$. In particular, we let $\delta > 0$ be some variable to be set later, and we note that $|x - x_0| < \delta$ implies

$$|F(x) - F(x_0)| = \left|\int_a^x f(t)\,dt - \int_a^{x_0} f(t)\,dt\right| = \left|\int_{x_0}^x f(t)\,dt\right| < |x - x_0| \cdot M < \delta M$$

by bounding directly. (We are blatantly ignoring some details in these inequalities.) But now we can set $\delta := \frac{\varepsilon}{M}$ to get the result.

- We omit this proof. ■

Here is an exercise.

Exercise 5.36 (Ross 34.12)**.** Fix f continuous on $[a,b]$ such that

$$\int_a^b f(x)g(x)\,dx = 0$$

for each continuous function $g : [a,b] \to \mathbb{R}$. Then f is zero.

Proof. Fix f a continuous function on $[a,b]$, and we suppose there exists some $c \in [a,b]$ for which $f(c) \neq 0$, and we will show there exists a $g : [a,b] \to \mathbb{R}$ such that

$$\int_a^b f(x)g(x)\,dx \neq 0.$$

Take $f(c) > 0$ without loss of generality. Now, by continuity, there exists $\delta > 0$ such that

$$|x - c| < \delta \implies |f(x) - f(c)| < \frac{f(c)}{2}$$

so that $|x-c| < \delta$ implies $f(x) \geq f(c)/2$. To make life easy, we fix an interval $[a_0, b_0] \subseteq [c-\delta, c+\delta] \cap [a,b]$.

The idea, now, is to make g "bump" on $[a_0, b_0]$ and then be zero everywhere else. We won't write this out rigorously, but the point is that we will have

$$\int_a^b f(x)g(x)\,dx = \int_{a_0}^{b_0} f(x)g(x)\,dx \geq \frac{f(c)}{2}\int_{a_0}^{b_0} g(x)\,dx,$$

which we can make sure is positive. This finishes. ■

5.3.2 Improper Integrals

We are going to skip over the discussion of the Riemann-Stieltjes integral and go straight into the improper integral.

Our setting for normal integrals was for bounded functions on closed intervals. For improper integrals, we will relax this condition to functions on open intervals. As an example, we might be tempted to fix

$$I := \int_0^1 \sin\left(\frac{1}{x}\right)\,dx$$

into a proper integral by just setting $\sin\left(\frac{1}{x}\right)$ to be something arbitrary at 0, but it is probably more productive to set $u := 1/x$ to make this into an improper integral

$$\int_1^\infty \frac{\sin u}{u^2}\,du.$$

This integral is more tractable because its absolute value has

$$\int_1^\infty \left|\frac{\sin u}{u^2}\right|\,du \leq \int_1^\infty \frac{1}{u^2}\,du < \infty,$$

so we have that I will converge.

Another approach to I is to look at

$$\lim_{a\to 0^+}\int_a^1 \sin\left(\frac{1}{x}\right)\,dx,$$

which has the benefit of us not having to do a substitution or even touch 0 directly at all. So here is our actual definition.

Improper integral

Definition 5.37 (Improper integral). Fix $f : [a,b) \to \mathbb{R}$, where $b \in \mathbb{R} \cup \{\infty\}$. Then, if f is integrable on each $[c,d] \subseteq [a,b)$, we define
$$\int_a^b f(t)\,dt := \lim_{x \to b^-} \int_a^x f(t)\,dt.$$
There is an analogous definition for $f : (a,b] \to \mathbb{R}$ (where $a \in \mathbb{R} \cup \{-\infty\}$) and even $f : (a,b) \to \mathbb{R}$ (where $a \in \mathbb{R} \cup \{-\infty\}$ and $b \in \mathbb{R} \cup \{\infty\}$).

Let's do some exercises; there honestly is not much theory here.

Exercise 5.38 (Ross 36.3). Fix $p > 0$. Then
$$\int_0^1 x^{-p}\,dx$$
converges if and only if $p < 1$.

Proof. We simply do our cases.

- If $p \neq 1$, then we see
$$\int_0^1 x^{-p}\,dx = \lim_{a \to 0^+} \int_a^1 x^{-p}\,dx = \lim_{a \to 0)} \frac{x^{1-p}}{1-p}\bigg|_a^1 = \lim_{a \to 0^+} \frac{1-a^{1-p}}{1-p}.$$
If $p > 1$, then $1 - p < 0$, so $a^{1-p} \to \infty$ as $a \to 0^+$. If $p > 1$, then $1 - p > 0$, so $a^{1-p} \to 0$ as $a \to 0^+$.
- If $p = 1$, then we see
$$\int_0^1 x^{-1}\,dx = \lim_{a \to 0^+} \int_a^1 x^{-1}\,dx = \lim_{a \to 0^+} \log 1 - \log a,$$
which diverges. ∎

Exercise 5.39. Fix $p > 0$. Then we claim
$$\int_0^\infty x^{-p}\,dx$$
always diverges.

Proof. We have problems at 0 and ∞, so we are interested in computing
$$\int_0^\infty x^{-p}\,dx = \int_0^1 x^{-p}\,dx + \int_1^\infty x^{-p}\,dx.$$
We know the original integral converges if and only if $p < 1$. The second integral is
$$\lim_{b \to \infty} \int_1^b x^{-p}\,dx = \lim_{b \to \infty} \frac{b^{1-p} - 1}{1-p}.$$
We are only interested in the case where $p < 1$, where the term $b^{1-p} \to \infty$ as $b \to \infty$. So this diverges always. ∎

Exercise 5.40. We show that
$$\int_{-\infty}^{\infty} e^{-t^2}\,dt < \infty.$$

Proof. It suffices to study
$$\int_{1}^{\infty} e^{-t^2}\,dt = \int_{-\infty}^{-1} e^{-t^2}\,dt.$$
But now we can say that
$$\int_{1}^{\infty} e^{-t^2}\,dt \le \int_{1}^{\infty} e^{-x}\,dt = e^{-1} < \infty,$$
which finishes. ■

Remark 5.41. Note that we in fact know that this integral exists because the function
$$F(x) := \int_{0}^{\infty} e^{-t^2}\,dt$$
is increasing and bounded above as $x \to \infty$. So F will actually have a limit, though this is somewhat annoying to prove.

Remark 5.42. One can actually show that
$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$
To see this, take a course in multi-variable calculus.

And that covers the material of the course.