MATH 112 NOTES

Abstract. Math 112 Real Analysis, taught by Salim Tayou. All virtues of these notes should be attributed to the instructor, and all vices of these notes should be attributed to the notetaker. The official textbook for the class is Principles of Mathematical Analysis by Walter Rudin. Take home midterm and final exam.

Contents

1. 9/6/23: introduction

The official textbook for the class is Principles of Mathematical Analysis by Walter Rudin. This class will require proof-writing. If you haven't seen before any proofs or way to write proofs, it might take a little more effort from you; in that

Date: Fall 2024.

case, you might take a look at the content of Math 101. Anyhow, today we'll see some proofs and you'll get a taste for what we might be doing.

There's a midterm and final exam, both take home. They'll consist of a set of exercises. The point of the take home exam is for you to spend more time thinking about the problems and digest the concepts. Weekly homework. Expected every Wednesday.

Real analysis will take a lot of practice. You'll see some formalism, and new ways of thinking, and you'll need to practice to let the concepts sink in.

Today, we'll be talking about the real numbers. Math in some ways is like a humanities class. Before you do computations and examples, we need to agree on our definitions. For example, what is a real number? We'll see limits, differentiation, and integration – all of these are concepts that will require us to understand and have a good definition of real numbers.

Take the integers $\mathbb{Z}$, whose elements consist of

$$\cdots - 1, -2, 0, 1, 2 \cdots$$

You can also consider rationals $\mathbb{Q}$

$$\frac{1}{2}, \frac{3}{5}, -\frac{101}{79}, \cdots.$$

In general, the rationals are fractions $\frac{a}{b}$ where a, b are integers and $b \neq 0$. But there are other real numbers: like $\sqrt{2}$, $\sqrt{3}$, π and e.

But these are not rational numbers! Here's one proof. We'll use proof by contradiction, i.e assume that the claim is false, then perform a series of deductions that leads to a contradictory conclusion.

Proposition 1.1. *$\sqrt{2}$ is not a rational number*

Proof. We'll prove this by contradiction. Assume that $\sqrt{2}$ is a rational number. Then there exists integers $a, b \in \mathbb{Z}$ such that $\frac{a}{b} = \sqrt{2}$.

Assume a, b are coprime, i.e the only common divisor of a and b is 1. If there was a common divisor of a and b, then we could reduce this fraction. More formally, if $a = a'd$ and $b = b'd$ where d is an integer greater than 1, then $\frac{a}{b} = \frac{a'}{b'}$.

Now assuming a, b are coprime, we have $a^2 = 2b^2$. This implies that 2 divides a^2. So a^2 is even. This implies that a is even (an odd number times an odd number remains odd, so for a^2 to be even, we must have a is even). If a is even, then by definition of being even, a is divisible by 2. In other words, $a = 2a'$ for some integer $a' \in \mathbb{Z}$. Then

$$a^2 = 2b^2 \implies 4a'^2 = 2b^2 \implies 2a^2 = b^2.$$

But this implies that 2 divides b^2, thus b^2 is even and thus b is even. But a and b both being even violates them being coprime. Contradiction. Thus, $\sqrt{2}$ is not rational. $\square$

So $\sqrt{2}$ is not in the rationals. More formally, $\sqrt{2} \notin \mathbb{Q}$. If you tried to compute $\sqrt{2}$ in decimal form,

$$\sqrt{2} = 1.41421356 \cdots$$

You could consider a sequence of rational numbers (each of these is expressible as a fraction)

$$1, 1.4, 1.41, 1.4142, \cdots$$

and we see that this sequence of rationals get closer and closer to $\sqrt{2}$, but we have already proved that these rationals will never touch $\sqrt{2}$, since $\sqrt{2}$ cannot be written as a fraction.

This example illustrates an important point. The problem of rational numbers is that "there are gaps." We can take a sequence of rational numbers such as

$$1, 1.4, 1.41, 14142, \cdots$$

and their "limit" (we have to define limits precisely later) will *not* be a rational number. We can analyze this situation more closely.

Proposition 1.2. *Let*

$$A = \{p \in \mathbb{Q} | p^2 < 2, p > 0\}$$

and

$$B = \{p \in \mathbb{Q} | p^2 > 2, p > 0\}.$$

Then A contains no largest number, and B contains no smallest number.

Proof. Let $p > 0$ be a rational number. Then define

$$q = \frac{2p+2}{p+2} = p - \frac{p^2-2}{p+2}.$$

Then

$$q^2 - 2 = \frac{(2p+2)^2}{(p+2)^2} - 2 = \frac{4p^2 + 8p + 4 - 2(p^2 + p + 4)}{(p+2)^2} = \frac{2(p^2-2)}{(p+2)^2}.$$

If $p \in A$, then by definition $p^2 < 2$. Then $q^2 < 2$ and $p < q$. So if $p \in A$, we've found a rational in A that is larger than p, namely q.

If $p \in B$, then by definition $p^2 > 2$ and so $q^2 > 2$ so $q \in B$, but also $q < p$. Thus if $p \in B$, we've found a smaller rational in B. □

This was an enlightening series of example. But here's our overarching goal: we need to explain how to construct and define the real numbers. Firstly, we have been implicitly using the notion of "ordering" with $<$ and $>$. Here is some standard notion of a set.

Definition 1.3. Some set-theoretic conventions:

(1) A set is a collection of objects.
(2) If A is a set, we write $x \in A$ if x is a member of A. The set with no elements is called the empty set, whose symbol is $\emptyset$.
(3) If A has at least one element, then A is nonempty and thus $A \neq \emptyset$.
(4) If A, B are two sets, we write $A \subset B$ if every element of A is an element of B. For example, $\mathbb{Z} \subset \mathbb{Q}$. If there is some element of B not in A, i.e. there exists $x \in B$ such that $x \notin A$, then A is a *proper* subset of B.

Definition 1.4. Let S be a set. An order on S is a relation $<$ such that

- if $x, y \in S$, then one of the following is true:

$$x < y \text{ or } x > y \text{ or } x = y.$$

- If $x, y, z \in S$, and $x < y$ and $y < z$, then $x < z$.

Notationally, we write $x \leq y$ if either $x < y$ or $x = y$. We call $(S, <)$ an *ordered set.*

Example 1.5. $(\mathbb{Q}, <)$ is an ordered set.

When you have a set, you can have different orders. For example, we could take the set of all students in Math 112, and impose an order relation on this set via age. Or we could impose an order relation on this set via height. Both of these order relations are valid order relations because they satisfy the above axioms (exercise: prove this to yourself). But they provide different orders.

Definition 1.6. Let S be an ordered set. Let $E \subset S$. If there exists $\beta \in S$ such that $\forall x \in E$ we have $x \leq \beta$, then we say that E is bounded above by β, and β is an upper bound for E.

If there exists $\beta \in S$ such that $\forall x \in E$, we have $x \geq \beta$, then we say that E is bounded below by β, and β is a lower bound of E.

Definition 1.7 (supremum, infimum)**.** Let S be an ordered set, and $E \subset S$ a subset. Suppose there exists $\alpha \in S$ such that the following are true:

(1) α is an upper bound for E
(2) If $\gamma < \alpha$, then γ is not an upper bound for E.

Then α is the **least upper bound** for E, and we write $\alpha = sup(E)$, or the supremum of E.

If there exists $\alpha \in S$ such that the following are true:

(1) α is a lower bound for E
(2) If $\gamma > \alpha$ then γ is not a lower bound for E

Then α is the **greatest lower bound** for E, and we write $\alpha = inf(E)$, or the infimum of E.

Example 1.8. We have ordered set $(\mathbb{Q}, <)$ and
$$E_1 = \{q \in \mathbb{Q}, q \leq 0\}$$
and
$$E_2 = \{q \in \mathbb{Q}, q < 0\}.$$
Then $\sup(E_1) = 0 \in E_1$, and $\sup(E_2) = 0 \notin E_2$.

Example 1.9. Let
$$A = \{q \in \mathbb{Q} | q^2 < 2, q > 0\}$$
and
$$B = \{q | q^2 > 2, q > 0\}.$$
Then
$sup(A)$ does not exist, and $\inf(B)$ does not exist. If the reader doesn't see why, they should go back and understand why before moving on.

Definition 1.10. An ordered set S has the least upper bound property if the following is true: if $E \subseteq S$ is a nonempty subset of S that is bounded above, then $sup(E)$ exists in S.

2. 9/11/23: Least upper bound, ordered fields, the real numbers

Last time we ended on the definition of the least upper bound.

Definition 2.1 (least upper bound property (LUB))**.** Given an ordered set S, the least upper bound property S is the property that for every nonempty subset $E \subseteq S$ that is bounded above, then $sup(E) \in S$.

Theorem 2.2. *Let S be an ordered set with the LUB. Let $B \subseteq S$, $B \neq \emptyset$, bounded below. Then $inf(B)$ exists in S.*

Proof. Let $L = \{x \in S | \forall y \in B, x \leq y\}$. We need to show that L is nonempty, bounded above. Showing this we know then that the supremum exists. Then we will show it actually equals $inf(B)$.

Since B is bounded below, then $L \neq \emptyset$. Since $B \neq \emptyset$, if $y_0 \in B$ then for every $x \in L$, then $x \leq y_0$. This implies that L is bounded above. Since S has the least upper bound property, then $sup(L)$ exists.

Now we need to prove that $sup(L) = inf(B)$. Let $\alpha = Sup(L)$.

- First we will prove that $\alpha \leq \gamma, \forall \gamma \in B$. Assume for the sake of contradiction that there exists $\gamma \in B$ such that $\gamma < \alpha$. Then γ is not an upper bound for L. This means that there exists $\delta \in L$ suhch that $\gamma < \delta$. But if $\delta \in L, \gamma \in B$ this would imply $\delta \leq \gamma$ by definition of L, contradiction. So in fact, we must have for every $\gamma \in B$, $\alpha \leq \gamma$.
- Now we prove that α is the greatest lower bound. Assume for the sake of contradiction, assume there exists $\alpha' > \alpha$ and α' is still a lower bound for B. This implies that $\alpha' \in L$. Since $\alpha = sup(L)$, we must have $\alpha' \leq \alpha$.

Together, these show that $\alpha = \inf(B)$. □

Remark from a question in class: the supremum and infimum of E in S are unique. The definition of the supremum of E in S was an element α such that $\alpha \geq \gamma, \forall \gamma \in E$, and also α is the smallest upper bound for E. This second condition forces uniqueness. If both α_1, α_2 are supremums, then we must have $\alpha_1 \leq \alpha_2$ since α_1 is the supremum. But also $\alpha_2 \leq \alpha_1$ since α_2 is a supremum. Thus, $\alpha_1 = \alpha_2$.

The next notion that we'll introduce is that of a field. A field is a general definition for an algebraic structure of a set. The real numbers have the algebraic structure of a field. Since this class is concerned with the real numbers, we'll need to be familiar with what a field is.

Definition 2.3. A field F is a set with two operations addition and multiplication $(+, *)$. Addition satisfies the following axioms:

- (baseline validity) $\forall x, y \in F$, $x + y \in F$.
- (commutativity) For all $x, y \in F$, $x + y = y + x$
- (associativity) For all $x, y, z \in F$, $(x + y) + z = x + (y + z)$
- (additive identity) There $\exists 0 \in F$ such that $0 + x = x \ \forall x \in F$
- (additive inverse) For every $x \in F$, there exists $-x \in F$ such that $x+(-x) = 0$.

Multiplication satisfies the following axioms:

- (baseline validity) For every $x, y \in F$, then $x * y \in F$.
- (commutativity) For every $x, y \in F$, then $x * y = y * x$
- (associativity) For every $x, y, z \, inF$, then $(x * y) * z = x * (y * z)$.
- (multiplicative identity) There exists $1 \in F$ such that $1 \neq 0$ and for every $x \in F$ then $x * 1 = x$ and $1 * x = x$.
- (multiplicative inverse) For every nonzero $x \in F$ there exists $x^{-1} \in F$ such that $x * x^{-1} = 1 = x^{-1} * x$.

Finally, multiplication and addition satisfy a compatability axiom:

- (distributivity) For every $x, y, z \in F$ we have $x * (y + z) = x * y + x * z$ and $(x + y) * z = x * z + y * z$.

The reason why $1 \neq 0$ is because otherwise, we would have $F = 0$. Can you prove this?

Example 2.4. The rationals $\mathbb{Q}$ equipped with the usual addition and multiplication axioms form a field. Check this!

Definition 2.5. An ordered field is a field F which is also an ordered set, where the order relation is compatible with the field structure. To be precise, this means:

- For every $x, y, z \in F$, $x + y \leq x + z \iff y \leq z$.
- For every $x, y \in F$ such that $x, y \geq 0$, then $xy \geq 0$.

We say that x is positive if $x > 0$ and negative if $x < 0$.

Here are immediate facts that we can deduce from the axioms.

Lemma 2.6. *Let F be a field.*

- *For every $x \in F$, we have $x * 0 = 0$.*
- *The additive identity (zero element) is unique.*
- *The multiplicative identity (unit element) is unique.*
- *If $x + y = y + z$, then $y = z$.*
- *If $x \neq 0$ and $xy = xz \implies y = z$.*
- *If $x \neq 0$ and $xy = 1$, then $y = x^{-1}$.*
- *If $x + y = 0$, then $y = -x$.*

Proof. Exercise left to the reader. If you need help, see proposition 1.14, 1.15, and 1.16 in Rudin. You should try thinking about these for yourself, so to familiarize yourself with the axioms. □

Example 2.7. The rational numbers $\mathbb{Q}$ with the usual order, usual addition, and usual multiplication forms a set that is in fact an ordered field.

Proposition 2.8. *Let F be an ordered field. Then*

- *For every $x \in F$, $x > 0 \iff -x < 0$.*
- *If $x \neq 0$, then $x^2 > 0$. In particular, $1 > 0$.*
- *If $0 < x < y$, then*

Proof. □

Theorem 2.9. *(1) There exists a set $\mathbb{R}$ with the structure of an ordered field.*
(2) It has the least upperbound property
(3) It contains $(\mathbb{Q}, +, *, <)$ with the usual order, addition, and multiplication relations as an ordered subfield.

Proof. Proof is in appendix in page 17 of Rudin. The proof uses a procedure called Dedekind cuts. □

Here are some useful properties of $\mathbb{R}$.

Theorem 2.10. *The field of real numbers $\mathbb{R}$ has the following properties:*

- *(archimedean) If $x, y \in \mathbb{R}$ and $x > 0$, then there exists an integer $n \geq 1$ such that $nx > y$.*
- *The rational numbers $\mathbb{Q}$ are dense in $\mathbb{R}$. In other words, for every distinct $x, y \in \mathbb{R}$, there $\exists p \in \mathbb{Q}$ such that*
$$x < p < y$$

Proof. Assume item 1 does not hold for some $x, y \in \mathbb{R}$ where $x > 0$. Let $A = \{nx | n \in \mathbb{Z}^+\}$. This is nonempty. For the sake of contradiction, suppose item 1 does not hold for A. Then this implies that A is bounded above by y. But $\mathbb{R}$ has the least upper bound property, so if A is bounded, then $\alpha = sup(A)$ exists.

We have $\alpha - x < \alpha$. Then $\alpha - x$ is not an upper bound for A. So there exists $n_0 \in \mathbb{Z}^+$ such that $\alpha - x < n_0 x$. This implies that $\alpha < x + n_0 x = (n_0 + 1) * x$. But this contradicts α being $sup(A)$ and in particular being an upper bound for A. Thus, item 1 must hold for all $x, y \in \mathbb{R}$ and $x > 0$.

Here is the proof of item 2. Take $y - x > 0$. by item 1, there exists an integer n such that $n(y - x) > 1$. Then there exists $m_1, m_2 \in \mathbb{Z}$ both ≥ 1 such that

$$m_1 > nx, m_2 > -nx.$$

Thus $-m_2 < nx < m_1$ and there exists m such that $m - 1 \leq nx < m$. Thus,

$$nx < m \leq 1 + nx < yn$$

which implies that

$$n < \frac{m}{n} < y.$$

□

3. 9/13/23: EXTENDED REAL, COMPLEX NUMBERS, ENDING CHAPTER 1

Today we'll wrap up chapter 1 of Rudin by introducing some relevant characters: extended reals and complex numbers.

We'll just state the following theorem from Rudin. It's something you intuitively know. We won't prove it completely rigorously for now, as we'll see a much more illuminating proof later on in the course.

We know how exponentiation works: x^n is just multiplying x by itself n times. But what if we went the other way? What does it mean to take the nth root of x? We have to be careful, but luckily we can prove this just from what we already know so far in the course.

Theorem 3.1 (Existence of nth root)**.** *For any real number $x > 0$, there exists a unique $y > 0$ such that $x = y^n$.*

Proof. The complete proof of existence can be found in Theorem 1.21 in Rudin. But as noted, we'll see a better proof later on in the class. We'll just sketch the main idea here. You can consider the set

$$E = \{t \in \mathbb{R} | t > 0, t^n < \alpha\}$$

and we know that $sup(E)$ exists by proposition 2.2, since this set is bounded and $\mathbb{R}$ has the least upper bound property. One need to shows that $sup(E)$ is the number we're looking for, such that $sup(E)^n = x$.

We can, however, easily prove uniqueness. Assume for the sake of contradiction that there exists distinct $y_1, y_2 > 0$ such that $y_1^n = y_2^n = x$. But if $y_1 \neq y_2$, then we must have, say, $0 < y_1 < y_2$. But by the field axioms, this would imply $0 < y_1^n < y_2^n$, which cannot happen. Thus we must have uniquess, i.e only one $y > 0$ such that $y^n = x$. □

A corollary of this fact is justification for something you already know. Distributing a fractional exponent! For example, we already know that $(ab)^2 = (ab)(ab)$ by definition, and by commutatvity, we have $(ab)^2 = (ab)(ab) = aabb = a^2b^2$. But what about a fractional exponent like $(2*3)^{\frac{1}{99}}$? How do we know we can distribute it?

Lemma 3.2. *For any real numbers $a, b > 0$, then $(ab)^{\frac{1}{n}} = a^{\frac{1}{n}}b^{\frac{1}{n}}$.*

Proof. We have $(a^{\frac{1}{n}}b^{\frac{1}{n}})^n = (a^{\frac{1}{n}})^n(b^{\frac{1}{n}})^n = ab$. Since $a^{\frac{1}{n}}b^{\frac{1}{n}}$ satisfies the equation for the nth root of ab, then by the uniqueness part of theorem 3.1, this is the nth root. □

There is nothing actually deep going on here. You know all of these facts already. But the point is that we want to place ourselves in the shoes of our ancestors hundreds of years ago; and rediscover for ourselves what we take for granted today.

Now that we have the real numbers, for completeness we should also introduce two other sets of numbers that are important, and depend on the notion of real numbers: the extended real line, and the complex numbers.

Definition 3.3. The set of extended real numbers is

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}.$$

$\overline{\mathbb{R}}$ is an ordered field, given the usual ordering of $\mathbb{R}$ for non-infinite elements, and for every $x \in \mathbb{R}$ we assert that $-\infty < x < \infty$.

Note that $\overline{\mathbb{R}}$ is not an ordered field, as $\overline{\mathbb{R}}$ is not even a field. In fact, ∞ and $-\infty$ do not even have additive inverses, let alone multiplicative inverses.

The only operations you are allowed to do with ∞ is: for any $x \in \mathbb{R}$ is $x + \infty = \infty$. Note that since ∞ has no additive inverse, you cannot cancel ∞ on both sides (if ∞ did have an additive inverse, this would be nonsensical, because then $x = 0$). Here's a complete list of operations we can extend.

$$\begin{aligned}
a \pm \infty = \pm\infty + a &= \pm\infty, && a \neq \mp\infty \\
a \cdot (\pm\infty) = \pm\infty \cdot a &= \pm\infty, && a \in (0, +\infty] \\
a \cdot (\pm\infty) = \pm\infty \cdot a &= \mp\infty, && a \in [-\infty, 0) \\
\frac{a}{\pm\infty} &= 0, && a \in \mathbb{R} \\
\frac{\pm\infty}{a} &= \pm\infty, && a \in (0, +\infty) \\
\frac{\pm\infty}{a} &= \mp\infty, && a \in (-\infty, 0)
\end{aligned}$$

We'll be revisiting the extended real line later on in the course. Now we'll discuss the field of complex numbers.

Proposition 3.4. *The set of complex numbers is the set*

$$\mathbb{C} = \mathbb{R}^2 = \{(a, b) | a, b \in \mathbb{R}\}.$$

The set of complex numbers is a field under the following operations: let $x, y \in \mathbb{C}$. So we can write $x = (a, b)$ and $y = (c, d)$, where $a, b, c, d \in \mathbb{R}$. Then

$$x + y = (a + c, b + d) \text{ and } x * y = (ac - bd, ad + bc).$$

This may seem strange at first. Where's i? In general, $(a, 0) = a$, and $(0, b) = bi$. So in this case, "i" here is $(0, 1)$. You can check that $i^2 = (0, 1) * (0, 1) = (1, 0) = 1$. As you can see, really (a, b) is another way of saying $a + bi$.

You can check that we have defined is indeed how addition and multiplication really works in the complex numbers. You know that $(a + bi) + (c + di) = (a + c) + (b + d)i$. Furthermore, $(a+bi)(c+di) = ac-bd+(ad+bc)i$. From now on, we'll just refer to a complex number as $(a+bi)$, which secretly from our definition/proposition is (a, b).

If $z = a + bi \in \mathbb{C}$, then the real part is $a = Re(z)$, and the imaginary part is $Im(z) = b$. The complex conjugate is $\bar{z} = a - bi$. The norm is $|z| = (z\bar{z})^{\frac{1}{2}} = \sqrt{a^2 + b^2}$.

Proposition 3.5. *For $z, w \in \mathbb{C}$, we have*

(1) $\overline{z + w} = \overline{z} + \overline{w}$
(2) $\overline{zw} = \overline{zw}$
(3) For all $z \in \mathbb{R}$, $z\overline{z} \in \mathbb{R}$ and $z\overline{z} \geq 0$ and $z\overline{z} = 0 \iff z = 0$.
(4) $|zw| = |z||w|$
(5) (triangle inequality) $|z + w| \leq |z| + |w|$

We can also define higher euclidean spaces.

Definition 3.6. For each $k \geq 1$, let the set of k-th dimensional real euclidean space is

$$\mathbb{R}^k = \{(x_1, \cdots, x_k) | \forall x_i \in \mathbb{R}\}.$$

This is a k-dimensional vector space. We have vector addition and scalar multiplication: letting $\vec{x} = (x_1, \cdots, x_k), \vec{y} = (y_1, \cdots, y_k)$ we have

$$\vec{x} + \vec{y} = (x_1 + y_1, \cdots, x_k + y_k)$$

and

$$\lambda\vec{x} = (\lambda x_1, \cdots, \lambda x_k)$$

The dot product is defined as

$$\vec{x} * \vec{y} = \sum_{i=1}^{k} x_i y_i$$

where $||\vec{x}|| = (\vec{x} * \vec{x})^{\frac{1}{2}} = (\sum_{i=1}^{k} x_i^2)^{\frac{1}{k}}$.

Compare these with the complex numbers. These k-dimensional euclidean spaces are vector spaces, but they are not fields. The dot product is some kind of mutliplication, but it is not a multiplication that gives $\mathbb{R}^k$ a field structure. There is really something special about $\mathbb{C}$ and its multiplication that endows it with special properties that so much of mathematics depends on.

Theorem 3.7. *Let $\vec{x}, \vec{y} \in \mathbb{R}^k$, $\lambda \in \mathbb{R}$.*

(1) $||\vec{x}|| \geq 0$, $||\vec{x}|| = 0 \iff \vec{x} = 0$
(2) $||\lambda\vec{x}|| = \lambda||\vec{x}||$
(3) $||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$.
(4) $||\vec{x} * \vec{y}|| \leq ||\vec{x}||||\vec{y}||$.

There's a famous inequality that appears quite frequently in mathematics.

Theorem 3.8 (Cauchy-Schwarz Inequality)**.** *If $a_1, \cdots, a_n, b_1, \cdots, b_n \in \mathbb{C}$, then*

$$|\sum_{i=1}^n a_i\overline{b_i}|^2 \leq (\sum_{i=1}^n |a_i|^2)(\sum_{i=1}^n |b_i|^2)$$

Proof. There are more conceptal proofs, but here we'll just do the direct, formal proof. We'll take the right hand side and subtract by the left hand side, and try to show that this is equal nonnegative. Doing this, and noting that $|z|^2 = z\overline{z}$, we have

$$(\sum_{i=1}^n a_i\overline{a_i})(\sum_{i=1}^n b_i\overline{b_i}) - (\sum_{i=1}^n a_i\overline{b_i})(\sum_{i=1}^n \overline{a_i}b_i) = \sum_{i,j=1}^n a_i\overline{a_i}b_i\overline{b_i} - \sum_{i,j=1}^n a_i\overline{b_i}\overline{a_j}b_j.$$

This simplifies to

$$\sum_{i\neq j} a_i\overline{a_i}b_i\overline{b_i} - \sum_{i<j} a_i\overline{b_i}\overline{a_j}b_j - \sum_{j<i} a_i\overline{a_j}\overline{b_i}b_j$$

which simplifies to

$$\sum_{i<j}[a_i\overline{a_i}b_j\overline{b_j} + a_j\overline{a_j}b_i\overline{b_i} - a_i\overline{b_i}\overline{a_j}b_j - a_j\overline{a_i}\overline{b_j}b_i] = \sum_{i<j}|a_i\overline{b_j} - a_j\overline{b_i}|^2 \geq 0$$

□

This is the end of the first chapter of Rudin. Now we'll introduce chapter 2, which is about topology.

Definition 3.9. Let A, B be two sets. A function $f : A \to B$ is the association of every element $x \in A$ to some element $f(x) \in B$.

If $E \subseteq A$, then

$$f(E) = \{f(x)|x \in E\} \subseteq B.$$

If $E \subseteq B$, then

$$f^{-1}(E) = \{x \in A|f(x) \in E\} \subseteq A.$$

If $y \in B$, then clearly

$$f^{-1}(\{y\}) = \{x \in A|f(x) = y\} \subseteq A.$$

Definition 3.10. Let $f : A \to B$ be a function. We say f is surjective if $f(A) = B$. In other words, for every element of $b \in B$, there is some element of $a \in A$ such that $f(a) = b$.

We say f is injective if every element $x \in A$ maps to a unique $f(x) \in B$. In other words, for $x, x' \in A$, $f(x) = f(x') \iff x = x'$.

4. 9/18/23: COUNTABILITY, CANTOR'S DIAGONAL ARGUMENT

Definition 4.1. Let $f : A \to B$ where A, B are sets. Then f is a bijection if

(1) f is one-to-one, or *injective*
(2) f is onto, or *surjective*

Definition 4.2. Let A, B be two sets. We say that A and B have the same cardinality if there exists a bijection $f : A \to B$.

The notion of sets having the same cardinality is a concrete instance of an abstract notion called an equivalence relation. Call sets A and B equivalent, written $A \sim B$, if they have the same cardinality, i.e there exists a bijection $f : A \to B$.

The following proposition verifies that the notion of sets having the same cardinality is indeed an equivalence relation.

Proposition 4.3. *(1) A has the same cardinality as A, i.e $A \sim A$.*
(2) If $A \sim B$, then $B \sim A$.
(3) If $A \sim B$ and $B \sim C$, then $A \sim C$.

Proof. (1) (reflexitivity) There is a bijection $A \to A$ via the identity map. Thus, $A \sim A$.
(2) (symmetry) If $A \sim B$, then there is a bijection $f : A \to B$. Then because it is a bijection, there is a bijective inverse $f^{-1} : B \to A$. Thus, $B \sim A$.
(3) (transitivity) If $A \sim B$, $B \sim C$, then there are bijections $f : A \to B, g : B \to C$. Then take the bijections $g \circ f : A \to B \to C$. This gives a bijection from $A \to C$. Thus, $A \sim C$.

$\square$

For $n \geq 1$, let $J_n = \{1, \cdots, n\}$, and let $J_0 = \emptyset$, the empty set.

Definition 4.4. Let A be a set.

(1) We say that A is finite if there exists $n \geq 1$ such that $A \sim I_n$ or A is empty.
(2) We define A being infinite to mean A is not finite.
(3) A is countably infinite if $A \sim \mathbb{N}$. A is countably finite if A is finite. We say A is countable if it is either countably infinite or countably finite.
(4) A is uncountable if A is not countable.

The notion of being countable is quite intuitive. A set is countable if you can keep taking out elements and count $1, 2, 3, \cdots$. There are some sets, however, that are uncountable. Let's see some examples of countable and uncountable.

Example 4.5. Immediately, we see that $\mathbb{N} = \{0, 1, 2, 3 \cdots\}$ is countable.

Furthermore, $\mathbb{Z}$ is countable. We count in this way: $0, 1, -1, 2, -2, 3, -3, 4, -4 \cdots$ — can you explicitly construct the bijection yourself, i.e. write a function $f : \mathbb{N} \to \mathbb{Z}$ giving a bijection ?

If A is a finite set and $B \not\subset A$ (not a subset), then B and A don't have the same cardinality.

Definition 4.6. Let A be a set. A sequence in A is a function $f : \mathbb{N} \to A$. For every $n \geq 0$, we have $f(n) := x_n \in A$. So you have a sequence

$$x_0, x_1, x_2, \cdots .$$

Note that you can replace a sequence $f : \mathbb{N} \to A$ with a sequence $\{1, 2, 3, \cdots\} \to A$, i.e. start the indexing at 1 instead of 0. So instead of $x_0, x_1, x_2, \cdots$, you do $y_1, y_2, \cdots$. Note that these definitions of a sequence are the same, since $\mathbb{N}$ (nonnegative integers) is equivalent to (in bijection with) the positive integers. $\mathbb{Z}^+$.

Theorem 4.7. *Let A be a countable set and $B \subseteq A$ a subset. Then B is countable.*

Proof. If B is finite, then we're done. Assume B is infinite. Then A is also infinite. Since A is countable, there exists a bijection $f : \mathbb{N} \to A$. Then we can write the elements of A as

$$\{x_1, x_2, x_3, \cdots\}.$$

Let n_1 be the smallest integer such that $x_{n_1} \in B$. Assume that we have constructed $x_{n_1}, \cdots, x_{n_k}$. Let n_{k+1} be the smallest integer n_{k+1} such that $n_{k+1} > n_k$ and $x_{n_{k+1}} \in B$. Then we can define a map

$$g : \mathbb{N} \to B$$

where $k \mapsto x_{n_k}$. This map is injective and onto. Can you finish this part? $\square$

Formally, here is what it means to have a family of sets indexed by some arbitrary set A.

Definition 4.8 (A family of sets indexed by A)**.** Let Ω be a set. Let A be a set, such that for any $\alpha \in A$, we have $E_\alpha \subseteq \Omega$. In other words, we have a map

$$E : A \to P(\Omega)$$

where E maps $\alpha \mapsto E_\alpha \in P(\Omega)$, where $P(\Omega)$ denotes the set of all subsets of Ω.

Often we want to do operations on this family such as taking intersections and unions.

Definition 4.9. Define

$$\bigcup_{\alpha \in A} E_\alpha = \{x \in \Omega | \exists \alpha \in A, x \in E_\alpha\}.$$

Define

$$\bigcap_{\alpha \in A} E_\alpha = \{x \in \Omega | \forall x \in A, x \in E_\alpha\}.$$

Example 4.10. $\bigcup_{n \geq 1} [\frac{1}{n}, 1] = (0, 1]$. A 100 percent airtight proof of this requires the archmidean principle. In particular, the key step is to demonstrate that for every positive number close to 0, there is a $\frac{1}{n}$ that is even closer to 0.

Theorem 4.11. *Let $\{E_n\}_{n \geq 1}$ be a countable family of sets, such that E_n is countable for every $n \geq 1$. Then*

$$S = \bigcup_{i=1}^{n} \text{ is also countable.}$$

Proof. Let's represent the elements in a table $\begin{pmatrix} E_1 = & \{x_{11}, x_{12}, x_{13}, \cdots\} \\ E_2 = & \{x_{21}, x_{22}, x_{23}, \cdots\} \\ E_3 = & \{x_{31}, x_{32}, x_{33}, \cdots\} \\ \vdots & \vdots \end{pmatrix}$ How do we count all these elements x_{jk}? Count them along the antidiagonal, i.e the first antidiagonal is

$$x_{11},$$

the second antidiagonal is

$$x_{12}, x_{21},$$

the third antidiagonal is

$$x_{13}, x_{22}, x_{31},$$

and the fourth antidiagonal is

$$x_{14}, x_{23}, x_{32}, x_{41},$$

and so on. Note that the n-th antidiagonal is comprised of the elements x_{ij} such that $i + j = n + 1$. We can count along these diagonals, while eliminating repetition. $\square$

A corollary of this is that, if A is countable, and for every $\alpha \in A$, E_α is countable, then

$$\bigcup_{\alpha \in A} E_\alpha$$

is countable.

Theorem 4.12. *If A is countable, then*

$$A^n = \{(a_1, \cdots, a_n) | a_i \in A\}$$

is countable.

Proof. Proceed by induction. The $n = 1$ case is already true. Assume it holds true for n. Now we show it holds true for $n + 1$. Note that if A and B are countable, then

$$A \times B = \bigcup_{b \in B} A \times \{b\}$$

is countable. Thus, $A^{n+1} = A^n \times A$ is countable. $\square$

Then we have the following corollary:

Proposition 4.13. *The rationals $\mathbb{Q}$ are countable.*

Proof. We have

$$\mathbb{Q} = \{\frac{a}{b} | \gcd(a, b) = 1, a \in \mathbb{Z}, b \geq 1\}.$$

Furthermore,

$$A = \{(a, b) | \gcd(a, b) = 1\} \subseteq \mathbb{Z} \times \mathbb{Z}^+ \to \mathbb{Q}$$

where the map $\mathbb{Z} \times \mathbb{Z}^+ \to \mathbb{Q}$ sends $(a, b) \mapsto \frac{a}{b}$. The map restricted to domain A is a bijection to $\mathbb{Q}$. But note $\mathbb{Z} \times \mathbb{Z}^+$ is countable, thus A is countable. Thus $\mathbb{Q}$ is countable. $\square$

Now we arrive here to an important example of an uncountable set. The following is Georg Cantor's famous "diagonal argument." Georg Cantor played a major role in the discovery of set theory and the cementing of it as a foundational theory of mathematics. If you understand it, then congratulations, you have undergone a rite of passage of every young mathematician.

Theorem 4.14. *Let A be the set of sequences $\mathbb{N} \to \{0, 1\}$, i.e*

$$A = \{f : \mathbb{N} \to \{0, 1\} | \ f \text{ is a function } \}.$$

Then A is uncountable.

Proof. Assume for the sake of contradiction that A is countable. Then we could write $A = \{f_1, f_2, f_3, \cdots\}$. And we can write their evaluations at numbers in $\mathbb{N}$ in a table: $\begin{pmatrix} f_1(1) & f_1(2) & f_1(3) & \cdots \\ f_2(1) = & f_2(2) & f_2(3) & \cdots \\ f_3(1) = & f_3(2) & f_3(3) & \cdots \\ \vdots & \vdots & & \end{pmatrix}$. Consider the sequence $g(n) = 1 - f_n(n)$. So $g(n) = 0$ if $f_n(n) = 1$, and $g(n) = 1$ if $f_n(n) = 0$. Then $g : \mathbb{N} \to \{0, 1\} \implies g \in A$. But $g \neq f_k$ for every k. This is because $g(k) = 1 - f_k(k) \neq f_k(k)$. Thus, for every n, there is some value, namely n, such that g and f_n do not agree at n. $\square$

We can use this to show that $\mathbb{R}$ is uncountable!

Proposition 4.15. *The real numbers $\mathbb{R}$ are uncountable.*

Proof. We can write every element $x \in \mathbb{R}$ in binary expansion (including decimals):

$$x = \sum_{i \in \mathbb{Z}} \frac{a_i}{2^i}$$

and this establishes that $R \sim$ the set of functions $\mathbb{N} \to \{0, 1\}$. We'll do this proof in more detail next time. $\square$

5. 9/20/23: METRIC SPACES

Today we're going to talk about metric spaces. These are spaces with a notion of distance.

Definition 5.1 (Metric space)**.** Let X be a set. A metric (or distance function) on X is a function $d : X \times X \to \mathbb{R}_{\geq 0}$ satisfying

(1) $d(x, y) \geq 0 \forall x, y$ and $d(x, y) = 0 \iff x = y$
(2) $d(x, y) = d(y, x)$
(3) (triangle inequality) $d(x, z) \leq d(x, y) + d(y, z)$, for all x, y, z.

The pair (X, d) is called a *metric space*.

It's really hard to overemphasize how important metric spaces are. If you continue studying mathematics, you'll see them everywhere. Here's a the prototypical example of a metric space that you're already familiar with.

Example 5.2. Suppose $X = \mathbb{R}^2$. Endow X with a metric

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

This the 2-dimensional real plane, where you measure distance as you'd expect. Can you verify for yourself that this function satisfies the metric axioms?

If you were a mathematician, and you didn't know what a metric space was, maybe your best clue would have been staring at this example. What does it mean to measure distance? What rules should a distance function satisfy? All the clues are in this example. In practice, many metric spaces may be abstract and the distance functions less intuitive. That's why it's important to have a formalization and axiomatic definition of a metric space.

A remark: if (X, d) is a metric space, and $Y \subseteq X$, then Y inherits the metric from X. In other words, (Y, d) is also a metric space.

Definition 5.3. Let $p \in X$ and fix a real number $r > 0$. An open ball of radius r around p is

$$N_r(p) = \{q \in X | d(p, q) < r\}.$$

When we say something like "a neighborhood of p," we are referring to some ball of radius r around p.

Definition 5.4. Let $E \subseteq X$ and $p \in X$. We say p is an accumulation point (or limit point) if every neighborhood of p contains some $q \neq p \in E$.

Definition 5.5. If $p \in E$ which is not a limit point, we say p is an isolated point

Definition 5.6. A point $p \in E$ is an interior point of E if there is a neighborhood $N_r(p) \subseteq E$.

Definition 5.7. E is open if every point of E is an interior point of E.

Definition 5.8. E is closed if $X \setminus E$ is open.

Definition 5.9. E is perfect if E is closed and every point of E is a limit point of E.

Definition 5.10. E is bounded if there $\exists M > 0$ and $q \in X$ such that $E \subseteq N_M(q)$.

Definition 5.11. E is dense in X if every point of X is a limit point of E.

Our most important example in this class of a dense subset is $\mathbb{Q} \subseteq \mathbb{R}$.

Example 5.12. Let $X = \mathbb{R}$. Then

(1) (a, b) is open
(2) $[a, b]$ is closed
(3) $(a, b]$ is neither open nor closed.

All of these are examples of bounded sets in $\mathbb{R}$.

Proposition 5.13. *Let (X, d) be a metric space, $E \subseteq X$. Then E is closed $\iff$ every limit point of E is a point of E.*

Proof. Forward direction: suppose E is closed. Let $x \in X$ be a limit point of E. Assume for the sake of contradiction that $x \notin E$. Well, if E is a closed set, then $X \setminus E$ is open. We know that $x \in X \setminus E$. But if $X \setminus E$ is open, then $x \in X \setminus E$ is an interior point of $X \setminus E$, and thus has a neighborhood entirely contained in $X \setminus E$. But then we have found a neighborhood of x that does not contain any point of E, contradicting the fact that x is a limit point. Thus, if $x \in X$ is a limit point, we must have $x \in E$.

Reverse direction: now suppose every limit point of E is a point of E. Now assume for the sake of contradiction that E is not closed. Thus, $X \setminus E$ is not open. But if $X \setminus E$ is not open, then there exists some point $p \in X \setminus E$ such that it is not an interior point. Then for every neighborhood $N_\epsilon(p)$ of p, it must intersect E. But this implies that p is a limit point of E. Then we must have $p \in E$. Contradiction, because we said $p \in X \setminus E$. Thus, we must have that E is closed. $\square$

Proposition 5.14. *Let $p \in X$, $r > 0$. Then $B_r(p)$ is open.*

Proof. Let $q \in B$. Then let $h = d(p, q)$. Then $B_{r-h}(q) \subseteq B_r(p)$ (can you see why?), so you're done. $\square$

Theorem 5.15. *If p is a limit point of E, then every neighborhood of p contains infinitely many points of E.*

Proof. Let N be a neighborhood of p. Assume for the sake of contradiction that $N \cap E$ has finitely many points of E, say

$$q_1, \cdots, q_m.$$

But then we have a finite number of distances

$$d(p, q_1), \cdots, d(p, q_m).$$

Then chooose a radius r that is smaller than all of these, and thus we've found a neighborhood $B_r(p)$ which does not intersect E at all. $\square$

Example 5.16. Let $X = \mathbb{R}^2$ and

$$E = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq 1\}.$$

Note you can rewrite the condition of E as $d((x, y), (0, 0) \leq 1$. This set is closed, bounded, and perfect.

If we let $E = \{1, 2, 3\}$, this set is closed and bounded.

If we let $E = \mathbb{Z}$, this is closed but not bounded.

If we let

$$E = \{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots\}$$

this set is bounded but not closed, since 0 is a limit point but not in the set.

Theorem 5.17. *Fix a metric space (X, d).*

1. *If $\{G_\alpha\}_{\alpha\in I}$ is a collection of open sets, then $\bigcup_{\alpha\in I} G_\alpha$ is open.*
2. *If $\{F_\alpha\}_{\alpha\in I}$ is a collection of closed sets, then $\bigcap_{\alpha\in I} F_\alpha$ is closed.*
3. *If you have a finite collection of open sets $G_1, \cdots, G_n$, then $\bigcap_{i=1}^n G_i$ is open.*
4. *If you have a finite collection of closed sets $F_1, \cdots, F_n$, then $\bigcup_{i=1}^n F_i$ is closed.*

In general, an infinite intersection of open sets need not remain open, and an infinite union of closed sets need not remain closed.

Proof.

(1) Pick any $p \in \bigcup_{\alpha\in I} G_\alpha$. Then $p \in G_\alpha$ for some α. Then since G_α is open, there is some neighborhood N of p that is contained in G_α. Then it is contained in $\bigcup_{\alpha\in I} G_\alpha$. Thus, the union is open.

(2) What is $X \setminus \bigcap_{\alpha\in I} F_\alpha$? This is $\bigcup_{\alpha\in I} X \setminus F_\alpha$. But as we just showed, an arbitrary union of opens is open, thus $X \setminus \bigcap_{\alpha\in I} F_\alpha$? is open, and thus $\bigcap_{\alpha\in I} F_\alpha$ is closed.

(3) Pick any $p \in \bigcap_{i=1}^n G_i$. For each i, there is a neighborhood N_i such that $N_i \subseteq G_i$. Then choose the neighborhood of smallest radius, and we have that that neighborhood is contained in $\bigcap_{i=1}^n G_i$.

(4) We have that $X \setminus \bigcup_{i=1}^n F_i = \bigcap_{i=1}^n X \setminus F_i$. We just showed that finite intersections of open sets remains open. Thus, $X \setminus \bigcup_{i=1}^n F_i$ is open, and thus $\bigcup_{i=1}^n F_i$ is closed.

□

Here is an example of an infinite union of closed sets not remaining closed:

Example 5.18. $\bigcup [\frac{1}{n}, 1] = (0, 1]$.

Here is an example of a infinite intersection of open sets not remaining open.

Example 5.19. $\bigcap (-\frac{1}{n}, \frac{1}{n}) = \{0\}$.

Definition 5.20. Let (X, d) be a metric space. Let $E \subseteq X$. Then the closure of E is defined to be

$$\overline{E} := E \cup \{\text{all limits points of E}\}.$$

Theorem 5.21.

1. *$\overline{E}$ is closed*
2. *E is closed $\iff$ $E = \overline{E}$*
3. *If $E \subseteq F$, then $\overline{E} \subseteq \overline{F}$. In particular, if F is closed, then $\overline{E} \subseteq F$.*

Proof. (1) To show $\overline{E}$ is closed, we want to show that $X \setminus \overline{E}$ is open. Pick some $p \in X \setminus \overline{E}$. Then p is not a limit point of E, otherwise it would be in $\overline{E}$. Therefore, there exists a neighborhood of p that does not intersect E at all, thus it is contained in $X \setminus \overline{E}$. Thus, the complemenet is open, so $\overline{E}$ is closed.

(2) Forward direction: suppose E is closed. We showed in proposition 5.13, that E contains all of its limit points. Thus, $E = \overline{E}$. Reverse direction: suppose $E = \overline{E}$. Then again by proposition 5.13, $E = \overline{E}$ means E has all of its limit points, so we have E is closed.

(3) Suppose $E \subseteq F$. We want to show that $\overline{E} \subseteq \overline{F}$. Suppose $p \in \overline{E}$. If $p \in E$, then $p \in F \subseteq \overline{F}$. Now suppose p is a limit point of E. Then p is also a limit point of F, as every neighborhood of p contains some distinct element in E, and thus some distinct element in F. Thus, $p \in \overline{F}$.

□

Theorem 5.22. *If $E \subseteq \mathbb{R}$ is bounded above, and $y = sup(E)$, then $y \in \overline{E}$.*

Proof. If $y = sup(E)$, then consider some open neighborhood of y. If $y \notin \overline{E}$, this would imply that $y \notin E$, and y is not a limit point of E. Then there would exist some open neighborhood $(y-\epsilon, y+\epsilon)$ that does not intersect E. But, this would imply that $y-\epsilon$ is an upper bound of E, and it is smaller than y. This contradicts y being the supremum. Thus, we must have y is a limit point of E and thus $y \in \overline{E}$. □

Remark: let X be a metric space, and $Y \subseteq X$. Then an open set of Y is not necessarily an open set in X. Counterexample: $X = \mathbb{R}$ and $Y = [0,1)$. Then $[0, \frac{1}{2})$ is open in Y, but not in $X = \mathbb{R}$.

6. 9/25/23: Compactness, forward direction of Heine Borel

Last time we talked about metric spaces. These are sets which come equipped with a metric which satisfy the appropriate axioms. We had the notion of a ball of radius r around p, limit points, open sets, and closed sets. Recall that one can take infinite unions of opens, and it will remain open, but taking infinite intersections of opens is not always open. And remember that an infinite union of closed sets is not necessarily closed, but an infinite intersection of closed sets is still closed.

Here's an example of an infinite intersection of open sets not remaining open.

Example 6.1. Take any $p \in \mathbb{R}^2$. Consider increasingly smaller neighborhoods of p of radius $\frac{1}{k}$. Then

$$\bigcap_{i=1}^{\infty} B_{\frac{1}{k}}(p) = \{p\}$$

which is closed.

Recall that the closure of E in a metric space (X, d) is $\overline{E} := E \cup \{\text{ limit points }\}$. Furthermore, $\overline{E}$ is closed and is the smallest such.

Finally, recall that we proved this last time:

Proposition 6.2. *For any subset $E \subseteq \mathbb{R}$ bounded, then $sup(E) \in \overline{E}$.*

Proof. If $y = sup(E)$, then consider some open neighborhood of y. If $y \notin \overline{E}$, this would imply that $y \notin E$, and y is not a limit point of E. Then there would exist some open neighborhood $(y-\epsilon, y+\epsilon)$ that does not intersect E. But, this would imply that $y-\epsilon$ is an upper bound of E, and it is smaller than y. This contradicts y being the supremum. Thus, we must have y is a limit point of E and thus $y \in \overline{E}$. □

Definition 6.3. Let $E \subseteq Y \subseteq X$ where (X, d) is a metric space. We say E is open relative to Y if it is open as a subset of the metric space Y.

Example 6.4. Suppose $X = \mathbb{R}^2$, and Y is the horizontal axis. Then the subset $E = (-1, 1) \times \{0\}$ is an open subset relative to Y. But as you can see, it is not open in X. For any point in E, any open neighborhood will contain points that are not in E.

Theorem 6.5. *Let $Y \subseteq X$ where X is a metric space (X, d). Then for $E \subseteq Y$, then E is open relative to $Y \iff E = Y \cap U$ where U is open in X.*

Proof. Here's the reverse direction. Let $E = Y \cap U$ where U is open. Then for every $p \in E$, there exists $N_r(p) \subseteq U$ for some r. But

$$U \cap Y \supseteq N_r(p) \cap Y$$

and $N_r(p) \cap Y$ is exactly the open neighborhod of p of radius r in the metric space (Y, d).

Here's the forward direction. Let $E \subseteq Y$ and suppose E is open relative to Y. Then for every $p \in E$, there exists neighborhood $N_r^Y(p) \subseteq E$ for some radius r. Then for each of these radius, consider the neighborhood of p in X with that radius and take all of their unions:

$$U = \bigcup_{p \in E} N_{r_i}^X(p).$$

U remains open since it is an infinite union of opens. □

Now we'll discuss the notion of compactness. Consider the following three subsets of $\mathbb{R}$:

$$[0, 1], [0, 1), [0, \infty).$$

Knowing that one is compact and two of them are not, without even knowing the formal definition of compactness, you can probably guess which one is "compact." The answer is that $[0, 1]$ are compact, and the other two are not compact. This is quite intuitive, but getting the right definition is hard – it took humans hundreds of years. So now we'll enjoy the fruits of their labor:

Definition 6.6. Let X be a metric space, and $E \subseteq X$. Then E is compact (relative to X) if for any open cover of E, i.e a collection of open subsets $\{U_\alpha\}_{\alpha \in A}$ in X, such that

$$E \subseteq \bigcup_{\alpha \in A} U_\alpha,$$

there exists a finite subcollection $U_{\alpha_1}, \cdots, U_{\alpha_n}$ such that

$$E \subseteq \bigcup_{i=1}^{n} U_{\alpha_i}.$$

Example 6.7. Consider $[0, \infty) \subset \mathbb{R}$. This is not compact. For example, consider the opens

$$U_n = (-n, n).$$

Then

$$[0, \infty) \subseteq \bigcup U_n$$

but there exist no finite subcollection which still cover $[0, \infty)$. If there were, take the largest such $k \in \mathbb{Z}^+$. Then the union is bounded above by k, and k is not in this union, while $k \in [0, \infty)$.

Compactness is an intrinsic property of a space, as opposed to being a relative notion. The following theorem formalizes this.

Theorem 6.8. *Let $E \subseteq Y \subseteq X$. Then E is compact relative to $Y \iff E$ compact relative to X.*

Proof. Here's the forward direction. Assume E is compact relative to Y. Let $\{U_\alpha\}$ be a collection of opens in X that is an open cover of E. Then note that $\{U_\alpha \cap Y\}$ are opens in Y and which still cover E since $E \subseteq U_\alpha, Y \implies E \subseteq U_\alpha \cap Y$. But E is compact relative to Y, which implies there's a finite subcover $U_{\alpha_1} \cap Y, cdots, U_{\alpha_n} \cap Y$. Thus, there is a finite subcover $U_{\alpha_1}, \cdots, U_{\alpha_n}$ in X of E. This proves that E is compact relative to X.

Here's the reverse direction. Suppose E is compact relative to X. We want to show that E is compact relative to Y. Suppose we have an open cover of E in Y, i.e $\{U_\alpha\}$ are opens in Y such that

$$U \subseteq \bigcup U_\alpha.$$

But note we proved that if U_α is open in Y, then there exist U'_α open in X such that $U'_\alpha \cap Y = U_\alpha$. Thus, we have an open cover of E in X

$$U \subseteq \bigcup U'_\alpha.$$

But since U is compact relative to X, there is a finite subcover $U'_{\alpha_1}, \cdots, U'_{\alpha_n}$. Then

$$U \subseteq \bigcup_{i=1}^{n} U'_{\alpha_i} \cap Y = \bigcup_{i=1}^{n} U_{\alpha_i},$$

thus we have exhibited a finite subcover of the open cover of U in Y. Thus, U is compact relative to Y. □

Here's an extremely important theorem. You might think: why didn't we just define this as compactness? It's much easier. Here are a few reasons why our definition of compactness is much better than taking Heine-Borel as the definition: in mathematics, you'll encounter abstract topological spaces which are not just $\mathbb{R}^n$. In these cases, there might not even be a metric. Furthermore, in practice when proving that something is compact, it is much easier to use our definition than to use Heine Borel.

Theorem 6.9. *$E \subseteq \mathbb{R}^n$ is compact $\iff$ E is closed and bounded.*

Proof. Here's the forward direction. Suppose $E \subseteq \mathbb{R}^n$ is compact. First, let's show that E is closed. Hence, we need to show that $X \setminus E$ is open. For $p \in X \setminus E$, and for any $q \in E$, let $r_q = \frac{1}{2}d(p, q)$ (the number 100 is arbitrary, we can pick anything ≥ 2). Consider the collection

$$\{N_{r_q}(q)\}_{q \in E}.$$

This is an open cover of E. By compactness of E, there is a finite subcover

$$E \subseteq \bigcup_{i=1}^{n} N_{r_{q_i}}(q_i).$$

Take $r = \min\{r_{q_1}, \cdots, r_{q_n}\}$. We have $N_r(p) \cap N_{r_{q_i}}(q_i) = \emptyset$ for every $i = 1, \cdots, n$. Thus, $N_r(p) \cap E = \emptyset$, hence $N_r(p) \subset X \setminus e$. Thus, $X \setminus E$ is open, and hence E is closed.

We now show that E is bounded. But if E were not bounded, we could pick some point $p \in E$, and consider the balls of radius r, $N_r(p)$ for $r \in \mathbb{Z}^+$. Then

$$E \subseteq \bigcup_{r=1}^{\infty} B_r(p).$$

But if E was unbounded, then there could not possibly be a finite subcollection, violating the fact that E is compact. Thus, E must be bounded.

The reverse direction is the hardest part. We'll need to prove a few things before we attempt it, namely theorems 7.2 and 7.3. ☐

7. 9/27/23: Reverse direction of Heine Borel

Let (X, d) be a metric space. recall that $K \subset X$ is compact if for every open cover

$$K \subseteq \bigcup_{\alpha \in I} V_\alpha$$

there exist $J \subseteq I$ where J is finite and

$$K \subseteq \bigcup_{\alpha \in J} V_\alpha.$$

Recall from last time that we proved that if $Z \subseteq Y$ is open relative to Y, then this is equivalent to there being an open $U \subseteq X$ such that $U \cap Y = Z$. Thus, if $Z \subseteq Y$ is closed relative to Y, then this is equivalent to $(Y \setminus Z) \subseteq Y$ being open, which is equivalent to there being an open $U \subseteq X$ such that $U \cap Y = Y \setminus Z$. But this U is equivalent to $X \setminus U$ being closed, and we see that $(X \setminus U) \cap Y = Z$. Thus, $Z \subset Y$ is closed $\iff$ there exist closed $V \subseteq X$ such that $V \cap Y = Z$.

Theorem 7.1. *Let (X, d) be a metric space, and K compact subset of X, and $K' \subseteq K$ closed Then K' is also compact.*

Proof. Since $K' \subseteq K$ closed in K, then there exist $F \subseteq X$ closed of X such that $K' = K \cap F$. Note K' is closed by forward direction of theorem 6.9. Thus, K' is closed.

Let $K' \subseteq \bigcup_{\alpha \in I} V_\alpha$ be an open cover of K', and V_α are open subsets of X. Note

$$K \subseteq \bigcup_{\alpha \in I} V_\alpha \cup (K')^c,$$

and note $(K')^c$ is open. Then by compactness of K, there is a finite subcollection such that

$$K \subseteq \bigcup_{i=1}^{n} V_{\alpha_i} \cup (K')^c,$$

but $(K')^c \cap K' = \emptyset$, and $K' \subseteq X$. Thus, we see that

$$K' \subseteq \bigcup_{i=1}^{n} V_{\alpha_i}.$$

$\square$

Corollary: If $F \subseteq X$ is closed, and $K \subseteq X$ is compact, then $F \cap K$ is compact, since it is closed in K.

Let us now prove the reverse direction of Heine-Borel. Recall we need to prove the following facts:

Theorem 7.2. *If K is compact, then any infinite subset of $E \subseteq K$ has a limit point in K.*

Proof. By contradiction, assume that E has no limit point in K. Then for every $q \in K$, there exist $r_q > 0$ such that $N_{r_q}(q) \cap E$ has at most one element. If $q \notin E$, then we can find r_q such that $N_{r_q}(q) \cap E = \emptyset$, and if $q \in E$, we can find r_q such that $N_{r_q}(q) \cap E = \{q\}$. Then

$$E \subseteq K \subseteq \bigcup_{q \in K} N_{r_q}(q).$$

By compactness of K, there exist finite subcollection so that

$$E \subseteq K \subseteq \bigcup_{i=1}^{n} N_{r_i}(q_i).$$

But this is a contradiction, because $N_{r_i}(q_i) \cap E$ is at most one element, so this would imply E is finite, but contradiction because we assumed E is infinite. $\square$

Theorem 7.3. *If $I_1, I_2, \cdots \subseteq \mathbb{R}$ are closed bounded intervals such that*

$$I_1 \supseteq I_2 \supseteq \cdots$$

then

$$\bigcap_{k=1}^{\infty} I_k \neq \emptyset.$$

Proof. Let $I_k = [a_k, b_k]$ such that $I_k \supseteq I_{k+1}$ for every k. Let $x = supE$, where $E = \{a_k\}_{k=1}^{\infty}$. Since for any $x \in I_k$, we have $a_i \leq x \leq b_i$, we see E is bounded above and nonempty, so $sup(E)$ exists. Furthermore, $a_n \leq a_{n+m} \leq b_{n+m} \subseteq b_n$, so for every $m \geq 1$, $a_m \leq b_n$. So the b_n is an upper bound for E, so $x \leq b_n$. Then $a_n \leq x \leq b_n$. So $x \in I_n$ for every n, so $x \in \cap I_n \neq \emptyset$. $\square$

Let's generalize this fact. Some terminology before we move forward. A k-cell in $\mathbb{R}^k$ is

$$I = \prod_{i=1}^{k} [a_i, b_i] = \{(x_1, \cdots, x_k) | a_i \leq x_i \leq b_i\} \subseteq \mathbb{R}^k.$$

For example in $\mathbb{R}^3$, the cube $[0,1] \times [0,1] \times [0,1]$ is a 3-cell.

Theorem 7.4. *Let $\{I_n\}$ be a sequence of k-cells in $\mathbb{R}^k$ such that $I_{n+1} \subseteq I_n$.*

Proof. The proof follows analagously from 7.3. Let $I_n = [a_{1n}, b_{1n}] \times \cdots \times [a_{kn}, b_{kn}]$. Let $z = (\sup\{a_{1i}\}, \cdots, sup\{a_{ni}\})$. Then this $z \in \bigcap I_n$. $\square$

Theorem 7.5. *A k-cell $I \subseteq \mathbb{R}^k$ is compact.*

Proof. Let $I = \prod_{i=1}^k [a_i, b_i] \subseteq \mathbb{R}^k$. Let

$$S = (\sum_{i=1}^{k} (b_i - a_i)^2)^{\frac{1}{2}}.$$

This S is the maximum distance between any two points in this k-cell I. Let's assume for the sake of contradiction that this k-cell I is not compact. Then there exists an open cover

$$I \subseteq \bigcup G_\alpha$$

such that we cannot find a finite subcover that still covers all of I. Let

$$c_i = \frac{a_i - b_i}{2}.$$

Note each $[a_i, b_i] = [a_i, c_i] \cup [c_i, b_i]$. So you can subdivide this k-cell into 2^k smaller k-cells. For example, in $\mathbb{R}^3$, a cube with sidelengths 1 would be subdivided into 8 smaller cubes of sidelengths $\frac{1}{2}$. More formally, these smaller cells are of the form

$$\prod_{i=1}^{k} J_i$$

where $J_i = [a_i, c_i]$ or $[c_i, b_i]$. This shows there are 2^k of these smaller k-cells.

Note if every one of these smaller k-cells could be covered by finitely many G_α, then so could I. But we're assuming that I is not covered by finitely many G_α. Thus, at least one of these smaller k-cells cannot be covered by finitely many V_α. Let's call it I_1. Note that for any $\vec{x}, \vec{y} \in I_1$,

$$d(\vec{x}, \vec{y}) \leq \frac{S}{2}.$$

Take I_1, and subdivide again. Then again, there is some smaller k-cell $I_2 \subset I_1$ that is not covered by finitely many G_α. We can continue this process to obtain a chain

$$\cdots \subset I_{n+1} \subset I_n \subset \cdots \subset I_1 \subset I$$

such that for all $\vec{x}, \vec{y} \in I_{n+1}$, $d(\vec{x}, \vec{y}) \leq \frac{1}{2^{n+1}} S$, and I_{n+1} is not covered by finitely many V_α. But we know from 7.4 that

$$\bigcap_{n \geq 1} I_n \neq \emptyset.$$

Let $x \in \bigcap_{n \geq 1} I_n$. Then there exist α_0 such that for $x \in V_{\alpha_0}$, whch is open, so there exist $r > 0$ $N_r(x) \subset G_{\alpha_0}$. We can choose n such that

$$\frac{S}{2^n} < r.$$

For every $\vec{y} \in I_n$, we have

$$d(\vec{x}, \vec{y}) \leq \frac{S}{2^n} < r.$$

But this implies $I_n \subseteq N_r(\vec{x}) \subseteq V_{\alpha_0}$, contradiction because I_n was supposed to not be covered by finitely many G_α. ∎

Theorem 7.6 (Heine-Borel)**.** *Let* $(\mathbb{R}^k, d)$ *be the metric space with the usual euclidean metric*

$$d(\vec{x}, \vec{y}) = (\sum_{i=1}^{k}(x_i - y_i)^2)^{\frac{1}{2}}.$$

Let $E \subseteq \mathbb{R}^k$*. Then* E *is compact* $\iff$ E *is closed and bounded.*

Proof. Note we proved the forward direction last time. The goal here is to prove the reverse direction.

Let $K \subseteq \mathbb{R}^k$ be closed and bounded. Then $K \subseteq I = [-r, r]^k$ for some large enough $r > 0$ since K is bounded. Since $I = [-r, r]^k$ is compact by 7.5 and K is closed, then K is compact. □

8. 10/2/23: CANTOR SET, CONNECTEDNESS, SEQUENCES

Last time, we finished the proof of the Heine Borel theorem by doing the reverse direction. Recall that Heine-Borel states that $K \subseteq \mathbb{R}^k$ is compact $\iff$ K is closed and bounded.

Theorem 8.1. *Let* $K \subseteq \mathbb{R}^k$*. Then* K *is compact* $\iff$ *every infinite sequence has a limit point in* K*.*

Proof. The forward direction follows from theorem 7.2.

Reverse direction: assume every infinite sequence has a limit point in K. Then we need to show that K is closed and bounded. Let's first show that K is closed. Let x be a limit point of K in $\mathbb{R}^k$. Then for any $n \geq 1$, we have $N_{\frac{1}{n}}(x) \cap E \neq \emptyset$. So there exist $x_n \in E$ such that $d(x, x_n) \leq \frac{1}{n}$. This set $\{x_n\}$ is infinite. Then by assumption, it has a limit point $y \in K$. Now suppose $y \neq x$. Then there exist $\frac{1}{N} > 0$ such that $N_{\frac{1}{N}}(y)$ and $N_{\frac{1}{N}}(x)$ are disjoint. But for all $n \geq N$, we have $x_n \in N_{\frac{1}{N}}(x)$. By the triangle inequality, $d(x, y) \leq d(x, x_n) + d(x_n, y)$, so $d(x, y) - d(x, x_n) < d(x_n, y)$. This implies that for all $n \geq N$, x_n are not within $\frac{1}{N}$ of y, and thus y could not be a limit point. In other words, $N_{\frac{1}{N}}(y)$ does not contain any x_n for $n \geq N$, which contradicts y being a limit point. Thus, we must have $x = y$. Thus, K is closed.

Now we need to show that K is bounded. Assume for the sake of contradiction that K is not bounded. Then we can find a sequence $\{x_n\}$, $x_n \in K$ and $d(0, x_n) > n$. Then $E = \{x_n\}$ is an infinite subset of K. By assumption, E has a limit point $x \in K$. For any N, there exist $n \geq N$ such that $d(x, x_n) \leq 1$. Then

$$N \leq n \leq d(0, x_n) \leq d(0, x) + d(x, x_n) \leq 1 + d(0, x)$$

but this is a contradiction because this is supposed to hold for any N. But $d(0, x)$ is fixed, so by the archimedean principle we can definitely find N greater than $1 + d(0, x)$. Contradiction. Thus, K is bounded. □

Definition 8.2. If (X, d) is a metric space. Then $E \subseteq X$ is perfect if $\forall x \in E$, x is a limit point of E.

Example 8.3. $\mathbb{N}$ is not perfect. $[0, 1] \cup \{2\}$ is not perfect.

Theorem 8.4. *If* $P \subseteq \mathbb{R}^k$ *is perfect, then* P *is uncountable.*

Proof. See Rudin for a proof. □

The following example is a very famous counterexample. Anytime you want to provide a counterexample to some statement, you might want to consider using the cantor set.

Example 8.5 (Cantor set)**.**

First consider the interval $E_0 = [0,1]$. Then remove the middle third, so that we are left with

$$E_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1].$$

Then remove the middle one third of $[0, \frac{1}{3}]$ and remove the middle one third of $[\frac{2}{3}, 1]$. So we are left with

$$E_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1].$$

We define E_i similarly for all i. Then

$$P = \bigcap_{n \geq 1} E_n$$

is the famous **cantor set**. This set is famous because it is a useful counterexample. Here are a few of the properties of the cantor set.

- P is compact and nonempty
- $P = [0,1] \setminus (\bigcup(\frac{3k+1}{3^{n+1}}, \frac{3k+2}{3^{n+1}}))$
- P does not contain any interval of the form $[a,b]$.
- P is a perfect set

We'll introduce one more topological notion: the notion of connected sets. The notion of connectedness is really "how many pieces" are there of your space.

Definition 8.6. Let (X,d) be a metric space. Then X is connected if X cannot be written as $X = U_1 \cup U_2$ where U_1, U_2 are nonempty open subsets and $U_1 \cap U_2 = \emptyset$.

Proposition 8.7. *X is connected $\iff$ X cannot be written as $F_1 \cup F_2$ where F_1, F_2 are nonempty closed subsets and $F_1 \cap F_2 = \emptyset$.*

Proof. Suppose X is connected. Assume for the sake of contradiction we could write $X = F_1 \cup F_2$ where F_1, F_2 are nonempty closed subsets and $F_1 \cap F_2 = \emptyset$. Then define $U_1 = X \setminus F_1$, and $U_2 = X \setminus F_2$. THen $X = U_1 \cup U_2$, where U_1, U_2 are disjoint nonempty opens. But this contradicts X being connected.

Reverse direction: assume for the sake of contradiction that X was not connected. Then we could write $X = U_1 \cup U_2$ as a union of nonempty disjoint opens. But then take $F_1 = X \setminus U_1, F_2 = X \setminus U_2$. THen $X = F_1 \cup F_2$ nonempty disjoint closed, contradiction. So X must be connected. □

Definition 8.8. $A, B \subseteq X$ are said to be separated if $A \cap \overline{B} = \emptyset$, and $B \cap \overline{A} = \emptyset$.

Proposition 8.9. *X is connected $\iff$ X is not union of two nonempty separated subsets.*

Proof. Exercise left to the reader. □

Theorem 8.10. *Let $E \subseteq \mathbb{R}$. Then E is connected $\iff$ for every $x, y \in E$, $[x,y] \subseteq E$.*

Proof. Forward direction: let $x, y \in E$. Let $z \in [x, y]$. Then

$$E = (E \cap (-\infty, z]) \cup (E \cap [z, \infty)).$$

These are closed subsets of E (since $(-\infty, z]$ and $[z, \infty)$ are closed in $\mathbb{R}$ and E inherits the metric from $\mathbb{R}$). Then since E is connected, these pieces cannot be disjoint. Thus, $(E \cap (-\infty, z]) \cap (E \cap [z, \infty)) \neq \emptyset$. But this is a subset of $\{z\}$. Thus the intersection must be $\{z\} \implies z \in E$.

Now suppose if for any $x, y \in E$ we have $[x, y] \subseteq E$. Write $E = F_1 \cup F_2$ where F_1, F_2 are closed nonempty. We want to show that $F_1 \cap F_2 \neq \emptyset$. Suppose $x \in F_1, y \in F_2$. WLOG assume $x < y$. Then $[x, y] \subseteq E$. We have

$$[x, y] = ([x, y] \cap F_1) \cup ([x, y] \cap F_2).$$

Let $E = \{t \geq 0, x + t \in F_1 \cap [x, y]\}$, where $E \neq \emptyset$ and bounded. Note $sup(E)$ exists. Let $t_0 = Sup(E)$. So $x + t_0 \in \overline{F_1 \cap [x, y]} = F_1 \cap [x, y]$. For any $t > t_0$, we have $x + t \notin F_1 \cap [x, y] \implies x + t \in F_2 \cap [x, y]$. note $x + t_0$ is a limit point of $F_2 \cap [x, y]$. But $F_2 \cap [x, y]$ is closed, thus it must contain its limit points $\implies x + t_0 \in F_2 \cap [x, y]$. Thus, F_1, F_2 are not disjoint. □

Up until this point, we've been avoiding phrases such as "sequences" and "series" and "limits of sequences," although we've used ideas reminiscent of these ideas. Let's formalize all of this. Let (X, d) be a metric space.

Definition 8.11. A sequence $\{x_n\}$ in X converges to a point $x \in X$ if for every $\epsilon > 0$, there exist $N \geq 1$ such that for all $n \geq N$,

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

We succinctly write this as $\lim_{n\to\infty} x_n = x$.

If $\{x_n\}$ does not converge, then we say $\{x_n\}$ is divergent.

Theorem 8.12. *Some useful, intuitive facts:*

(1) *If $\{x_n\}$ converges to x and x', then $x = x'$.*
(2) *If $\{x_n\}$ converges, then $\{x_n\}$ is bounded.*
(3) *The converse is not true. A sequence that is bounded is not necessarily convergent.*
(4) *If $E \subseteq X$ and $x \in \overline{E}$ then there exists a sequence $\{x_n\}$ such that $x_n \to x$.*

Proof. (1) Suppose $\{x_n\}$ converges to x, x'. Assume for the sake of contradiction that $x \neq x'$. Choose $\epsilon = \frac{d(x,x')}{2}$. Then $N_\epsilon(x)$ and $N_\epsilon x'$ are disjoint. But there exist N such that for all $n \geq N$ we have $d(x_n, x) < \epsilon$ and there exist N' such that for all $n \geq N'$ we have $d(x_n, x') < \epsilon$. Then take $M = \max\{N, N'\}$. Then for all $n \geq M$, we have

$$d(x_n, x) < \epsilon, d(x_n, x') < \epsilon.$$

But $N_\epsilon(x)$ and $N_\epsilon x'$ are disjoint! So this is an immediate contradiction. Thus, we must have $x = x'$.

□

9. 10/4/23: MORE ON SEQUENCES

Recall what we defined last time. Let (X, d) be a metric space. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X, so a function $\mathbb{N} \to X$ such that $n \mapsto x_n$. Let $y \in X$. Then

$$\lim_{n\to\infty} x_n = y$$

if $\forall \epsilon > 0$, there $\exists N \geq 1$ such that $\forall n \geq N$, $d(x_n, y) < \epsilon$.

Example 9.1. Let $x_n = \frac{1}{n}$. Let's prove that $\lim_{n\to\infty} \frac{1}{n} = 0$. Let $\epsilon > 0$. We have to find $N \geq 1$ such that for every $n \geq N$, $\frac{1}{n} = d(\frac{1}{n}, 0) < \epsilon$. Indeed, it's enough to choose the smallest integer N such that $\frac{1}{\epsilon} < N$, wich we can do by the archimidean principle.

Recall that last class we proved that if (x_n) is convergent, then its limit is unique. Let's explore some further properties.

Proposition 9.2. *(1) If $\lim_{n\to\infty} x_n = y$, then for every $\epsilon > 0$, $N_\epsilon(y)$ contains all but finitely many terms x_n.*
(2) If $(x_n)_{n\in\mathbb{N}}$ converges then, it is bounded.
(3) If $E \subseteq X$ is a subset, and $x \in \overline{E}$ then there exists a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n \in E$ and $\lim_{n\to\infty} x_n = x$.

Proof. First item: follows immediately from the ϵ-definition of a limit. Can you see why?

Second item: Suppose there exists $x \in X$ such that $\lim_{n\to\infty} x_n = x$. Then for $\epsilon = 1$, we have that $N_1(x)$ contains all but finitely many of the x_n. Suppose that $N_1(x)$ contains all but $x_{i_1}, \cdots x_{i_k}$. Then let $R = \max\{1, d(x_{i_t}, x)\}$. Then from how R is defined, $N_R(x)$ contains all of x_n for every n.

Third item: suppose $x \in \overline{E}$. Then for every $n \in \mathbb{N}$, choose some element $x_n \in N_{\frac{1}{n}}(x)$. We know that $N_{\frac{1}{n}}(x)$ for every n will contain some element of E that is not x since x is a limit point. Then this sequence x_n converges to x, as the distance between x_n and x decreases as $\frac{1}{n}$ decreases. More formally, for any $\epsilon > 0$, choose N such that $\frac{1}{\epsilon} < N$. Then for all $n \geq N$, we have that $d(x_n, x) < \frac{1}{n} < \frac{1}{N} < \epsilon$. $\square$

Note that a bounded sequence does not necessarily need to be convergent. The following example demonstrates a bounded sequence that does not converge.

Example 9.3. Consider $x_n = (-1)^n$. To show that this sequence does not converge, we must show that for every $x \in \mathbb{R}$, there exist $\epsilon > 0$ such that for every N, there exist $n \geq N$ such that $d(x_n, x) \geq \epsilon$.

If $x \neq \pm 1$, then take $\epsilon < \min\{d(x, 1), d(x, -1)\}$ and you're done. Can you see why?

If $x = \pm 1$, then take $\epsilon = 1$ and you're done. This is because if $x = 1$, then $d(x_n, x) = 2$ for n odd. If $x = -1$, then $d(x_n, x) = 2$ for n even.

Example 9.4. Note that $\mathbb{Q}$ is countable. So we can write every element of $\mathbb{Q}$ as a sequence $\{x_n\}_{n\in\mathbb{N}}$. Does this converge?

Now that we have the definition of limits, we can ask ourselves: is the sum of limits the limit of sums? Indeed it is.

Proposition 9.5. *Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be sequences in $\mathbb{C}$ such that $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$.*

(1) *Then* $\{x_n + y_n\}_{n\in\mathbb{N}}$ *is also convergent, and*

$$\lim_{n\to\infty} x_n + y_n = x + y.$$

(2) *Then* $\{x_n y_n\}_{n\in\mathbb{N}}$ *is convergent, and*

$$\lim_{n\to\infty} x_n y_n = xy.$$

(3) *If* $x_n \neq 0$ *and* $x \neq 0$*, then* $(\frac{1}{x_n})_{n\in\mathbb{N}}$ *is convergent and*

$$\lim_{n\to\infty} \frac{1}{x_n} = \frac{1}{x}.$$

(4) *If* $\lim_{n\to\infty} x_n = x \neq 0$*. Then there exists* N_0 *such that* $x_n \neq 0$ *for all* $n \geq N_0$*.*

Note our metric here on $\mathbb{C}$ *is* $d(x, y) = ||x - y||$*.*

Proof. (1) Note that $||x + y - (x_n + y_n)|| \leq ||x - x_n|| + ||y - y_n||$. Fix $\epsilon/2$. Then there exist N_x such that for all $n \geq N_x$, we have $||x_n - x|| < \frac{\epsilon}{2}$. And there exist N_y such that for all $n \geq N_y$, we have $||y_n - y|| < \frac{\epsilon}{2}$. Then for all $n \geq N = \max\{N_x, N_y\}$, we have

$$||(x_n + y_n) - (x + y)|| < \epsilon.$$

(2) Note we have that

$$||x_n y_n - xy|| = ||x_n(y_n - y) - y(x - x_n)|| \leq |x_n||y_n - y| + |y||x - x_n|.$$

Recall that since (x_n) are convergent, then the x_n are bounded. So there exist C such that $|x_n| < C$ for all n. Then

$$|x_n y_n - xy| \leq C|y_n - y| + |y||x - x_n|.$$

Fix $\frac{\epsilon}{2C}$. There exist N_y such that for all $n \geq N_y$, $|y_n - y| < \frac{\epsilon}{2C}$. There also exist N_x such that for all $n \geq N_x$, $|x - x_n| < \frac{\epsilon}{2|y|}$. Then for all $n \geq N = \max\{N_x, N_y\}$, we have that

$$|x_n y_n - xy| < C * \frac{\epsilon}{2C} + |y|\frac{\epsilon}{2|y|} = \epsilon.$$

(3) Note that

$$||\frac{1}{x_n} - \frac{1}{x}|| = ||\frac{x - x_n}{x x_n}||.$$

Since $x \neq 0$, there exist $N_1 \geq 1$ such that for all $n \geq N_1$, $||x_n - x||\frac{1}{2}||x||$. Also note that

$$||x|| = ||x_n + x - x_n|| \leq ||x_n|| + ||x - x_n|| < ||x_n|| + \frac{1}{2}||x||.$$

Then there exist $N_2 \geq N_1$ such that for all $n \geq N_2$, we have $||x_n - x|| < \frac{||x||^2}{2}\epsilon$. Then

$$||\frac{1}{x_n} - \frac{1}{x}|| = ||\frac{x - x_n}{x x_n}|| < \frac{2||x - x_n||}{||x||^2} < \frac{2}{||x||^2}\frac{||x||^2}{2}\epsilon = \epsilon.$$

$\square$

Theorem 9.6. *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}^k$, $k \geq 1$, where the metric is the euclidean metric. In other words,*

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

For all n, write

$$x_n = (x_n^{(1)}, \cdots, x_n^{(k)}),$$

and suppose that $(x_n^{(i)})_{n\in\mathbb{N}}$ is a sequence in $\mathbb{R}$ for every i.

1. *$(x_n)_n$ converges in $\mathbb{R}^k \iff \forall i$, $(x_n^{(i)})$ converges in $\mathbb{R}$. If we have convergence, then*
$$\lim_{n\to\infty} x_n = (\lim_{n\to\infty} x_n^{(1)}, \cdots, \lim_{n\to\infty} x_n^{(n)}).$$
2. *If $(x_n)_n$ and $(y_n)_n$ are two convergent sequences in $\mathbb{R}^k$, then*
$$\lim_{n\to\infty} x_n + y_n = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n$$
(remember that addition here is addition of vectors).

Let us mention subsequences, which we will continue next class. Let (X,d) be a metric space and $(x_n)_{n\in\mathbb{N}}$ a sequence in X. Let $\sigma : \mathbb{N} \to \mathbb{N}$ be a strictly increasing function, i.e $\sigma(n) < \sigma(n+1)$ for every n. Then

$$(x_{\sigma(n)})_{n\in\mathbb{N}}$$

is called a subsequence of $(x_n)_{n\in\mathbb{N}}$.

10. 10/11/23: RELATIONSHIP BETWEEN COMPACTNESS AND CONVERGENCE OF SERIES, CAUCHY SEQUENCES

Assume that a sequence means an infinite sequence. Recall the following fact 7.2, which says that any infinite set in a compact X has a limit point. The following is a very important characterization of compact spaces with a metric.

Theorem 10.1. *Let (A,d) be a metric space. If every infinite sequence has a convergence subsequence then A is compact. Similarly, if A is compact, then every sequence has a convergence subsequence.*

Proof. Let's prove the forward direction. We'll show that if A is compact, every infinite sequence $\{p_n\}$ has a convergent subsequence.

If $\{p_n\}$ is finite, then take a single point. This will be a "convergent subsequence." If $\{p_n\}$ is infinite, there exists a limit point $p \in A$. For each ball of radius $\frac{1}{2^k}$, pick a point $p_{n_k} \in B_{\frac{1}{2^k}}(p)$. This produces a subsequence

$$p_{n_1}, p_{n_2}, p_{n_3}, \cdots.$$

For any ϵ, pick N such that $\frac{1}{2^N} < \epsilon$. Then for all $k \geq N$, we have that

$$d(p_{n_k}, p) < \frac{1}{2^k} \leq \frac{1}{2^N} < \epsilon.$$

Thus, this subsequence converges to p.

The reverse direction is a bit harder, and we won't do. But in general, you can use this fact as given. ∎

But note that we can actually prove this theorem for compact subsets of $\mathbb{R}^k$ using theorem 8.1: $X \subset \mathbb{R}^k$ is compact $\iff$ every infinite sequence has a limit point in X. Here we provide a full proof.

Theorem 10.2. *Let $X \subset \mathbb{R}^k$. Then X is compact $\iff$ every sequence has a convergent subsequence.*

Proof. If X is compact, if the sequence is finite, just take a single point. If the sequence is infinite, then by theorem 8.1, it has a limit point p. Then for each ball of radius $\frac{1}{2^k}$, pick a point $p_{n_k} \in B_{\frac{1}{2^k}}(p)$. This produces a subsequence

$$p_{n_1}, p_{n_2}, p_{n_3}, \cdots .$$

For any ϵ, pick N such that $\frac{1}{2^N} < \epsilon$. Then for all $k \geq N$, we have that

$$d(p_{n_k}, p) < \frac{1}{2^k} \leq \frac{1}{2^N} < \epsilon.$$

Thus, this subsequence converges to p.

If every sequence has a convergent subsequence, why is X compact? Because then every infinite set, we can take an infinite sequence of points from it. Then this will have a convergent subsequence. The point that this subsequence converges to will be a limit point of the infinite set. Thus, X is compact by theorem 8.1. □

Corollary: every bounded infinite sequence in $\mathbb{R}^n$ has a convergent subsequence. This is because any bounded infinite sequence is contained in a closed disk of radius R around the origin, where R is sufficiently large. This is what it means to be bounded. Thus, it is an infinite sequence in a compact subset. Thus it must have a convergent subsequence.

We've discussed the relationship between compactness, sequences, and convergence. Now we'll move on to another extremely important metric space notion: cauchy sequences. Why are we talking about cauchy sequences? It's because we want to eventually show that $\mathbb{R}^n$ is **complete**, which roughly means that it is "not missing any points" with respect to the metric.

Definition 10.3 (cauchy sequence)**.** A sequence $\{p_n\}$ is a metric space X is a cauchy sequence if $\forall \epsilon, \exists N \geq 0$ such that $d(p_n, p_m) < \epsilon$ for all $n, m \geq N$.

Example 10.4. Let $X = \mathbb{R} \setminus \{0\}$. Consider $p_n = \frac{1}{n}$. Then $p_1, p_2, p_3, \cdots$ is a cauchy sequence. Why? Pick any ϵ. Find N such that $\frac{2}{N} < \epsilon$. Then for all $n, m \geq N$, we have

$$|p_n - p_m| = |\frac{1}{n} - \frac{1}{m}| \leq |\frac{1}{n}| + |\frac{1}{m}| \leq \frac{2}{N} < \epsilon.$$

Let $X = \mathbb{Q}$. Consider the decimal expansion of $\sqrt{2}$:

$$1, 1.4, 1.41, 1.414, \cdots$$

This is also a cuachy sequence. But it looks like it converges to $\sqrt{2}$ right? But $\sqrt{2} \notin \mathbb{Q}$. So this sequence is only a cauchy sequence, but doesn't converge.

Here's a rephrasing of what it means for a sequence to be Cauchy.

Definition 10.5. For a subset $E \subset X$, define

$$diam(E) = \sup\{d(p,q) | p, q \in X\}.$$

Let $\{p_n\}$ be a sequence. Let $E_N = \{p_n\}_{n \geq N}$. Then $\{p_n\}$ is Cauchy $\iff$ $\lim_{N \to \infty} diam(E_N) = 0$.

Proposition 10.6. *Let $E \subset X$. Then $diam(E) = diam(\overline{E})$.*

Proof. Note that $E \subset \overline{E}$. Thus, $diam(E) \leq diam(\overline{E})$ immediately. How about $diam(E) \geq diam(\overline{E})$? Let $\bar{p}, \bar{q} \in \overline{E}$. For every $\epsilon > 0$, there exist $p, q \in E$ such that $d(p, \bar{p}) < \epsilon$ and $d(q, \bar{q}) < \epsilon$. Then by triangle inequality,

$$d(\bar{p}, \bar{q}) \leq d(\bar{p}, p) + d(p, q) + d(q, \bar{q}) \leq d(p, q) + 2\epsilon.$$

This ϵ can be taken to be as small as possible. This implies that

$$d(\bar{p}, \bar{q}) \leq d(p, q).$$

Thus, $diam(E) \geq diam(\overline{E})$. □

Proposition 10.7. *If $K_1 \supset K_2 \supset K_3 \supset \cdots$ is a sequence of nested compact subsets of X and $diam(K_n) \to 0$ as $n \to \infty$. Then*

$$\bigcap_{n \geq 0} K_n = \{pt\}.$$

Proof. By theroem 2.36 in Rudin, $\bigcap_{n \geq 0} K_n$ is nonempty. Note that we proved a version of this in $\mathbb{R}^n$ via theorems 7.4 and 7.3. We showed nonempty intersection of compact things are nonempty. Now we just need to show here that the intersection is actually just a single point.

Suppose $p \neq q \in \bigcap K_n = K$. Then $diam K > 0$, because $d(p, q) > 0$. But $K \subset K_n$ for every n, so this would imply $diam(K) \leq diam(K_n)$. This would contradict $diam(K_n) \to 0$ as $n \to \infty$. Thus, $\bigcap K_n$ must be a single point. □

Theorem 10.8. *Let X be a metric space.*

(1) Every convergent sequence is a cauchy sequence.
(2) If X is compact, then a cauchy sequence in X is also convergent.
(3) A cauchy sequence in $\mathbb{R}^n$ is also convergent.

Proof.

(1) Suppose $\{p_n\}$ is a convergent sequence, convering to p. Given $\epsilon > 0$, need to show there exist N such that for all $n, m \geq N$ we have $d(p_n, p_m) < \epsilon$. Note that since the sequence converges to p, you know that for some N, for all $m \geq N$, we have

$$d(p_m, p) < \frac{\epsilon}{2}.$$

Then by triangle inequality, for all $m, n \geq N$, we have

$$d(p_m, p_n) \leq d(p_m, p) + d(p, p_n) < \epsilon.$$

This completes the proof.

(2) Now suppose X is compact. Suppose we have a cauchy sequence in X. Well, we know that infinite sequence in a compact X has a limit point by 7.2. Then it has a subsequence $p_{n_1}, p_{n_2}, \cdots$ which converges to p. We want to show that actually this entire cauchy sequence $\{p_n\}$ converges to p. Fix ϵ. We'd like to show that for all $n \geq N$, we have

$$d(p_n, p) < \epsilon.$$

Note we have the triangle inequality

$$d(p_m, p) \leq d(p_m, p_{n_k}) + d(p_{n_k}, p).$$

For any m and n_k. Then we want to pick N_1 and N_2 such that for $m, n_k \geq N_1$ and $k \geq N_2$, then

$$d(p_m, p_{n_k}), d(p_{n_k}, p) < \frac{\epsilon}{2}.$$

But since the p_{n_i} converge to p, and because the $\{p_i\}$ is a cauchy sequence, then we can do this. Then take $N = \max\{N_1, n_{N_2}\}$. Then for all $m \geq N$, we have that

$$d(p_m, p) \leq d(p_m, p_{n_k}) + d(p_{n_k}, p) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

(3) Now we show that a cauchy sequence $\{p_n\}$ in $\mathbb{R}^n$ is convergent. First we show that this cauchy sequence is bounded. Well, if it is cauchy, then

$$\lim_{N\to\infty} diam\{p_n\}_{n\geq N} = \lim_{N\to\infty} diam E_N = 0.$$

By definition of convergence, there exist N such that for all $n \geq N$, we must have that $diam E_N < 1$. So E_N is bounded, and then consider the points $p_1, \cdots, p_{N-1}$. These all have finite distance from the origin. Thus, $\{p_n\}$ is bounded as a set. Then because this cauchy sequence is forced to be bounded, it must be convergent.

□

We arrive to an extremely important notion of metric spaces:

Definition 10.9. A metric space (X, d) is complete if every cauchy sequence is convergent.

What we've proved is that $\mathbb{R}^n$ is complete.

11. 10/16/23: Series, monotonicity conditions for convergence

Note that last time we showed $(\mathbb{R}^k, d)$ where d is the usual Euclidean metric, is a complete metric space. This means that every Cauchy sequence is convergent.

Proposition 11.1. *If (X, d) is a compact metric space, then (X, d) is complete.*

Proof. Let $\{x_n\}$ be a Cauchy sequence. Since X is compact, there is a convergent subsequence $\{x_{\sigma(n)}\}$ where $\sigma : \mathbb{N} \to \mathbb{N}$ is an increasing function.

Let ℓ be the limit of $\{x_{\sigma(n)}\}$. Fix $\epsilon > 0$. There exist N_0 such that for all $m \geq n \geq N_0$, $d(x_m, x_n) \leq \frac{\epsilon}{2}$. Furthermore, there exist $N_1 \geq N_0$ such that for all $N \geq N_1$, we have $d(x_{\sigma(n)}, \ell) \leq \frac{\epsilon}{2}$. Note $\sigma(n) \geq n$ for every n.

Note for $n \geq N_1$, we have

$$d(x_{\sigma(n)}, x_n) \leq \frac{\epsilon}{2}$$

and

$$d(x_{\sigma(n)}, \ell) \leq \frac{\epsilon}{2}$$

so by the triangle inequality,

$$d(x_n, \ell) \leq \epsilon.$$

This shows that every cauchy sequence is convergent. Hence, (X, d) is complete. □

Example 11.2. The rationals $(\mathbb{Q}, d)$ where $d(x, y) = |x-y|$ is not complete. Recall the example where we took the decimal expansion of $\sqrt{2}$. This will be a cauchy sequence, but since $\sqrt{2} \notin \mathbb{Q}$, the sequence in $\mathbb{Q}$ does not converge.

Definition 11.3. A sequence $(s_n)_{n\geq 1}$ in $\mathbb{R}$ is

- *monotonically increasing* if $s_{n+1} \geq s_n$ for every n.
- *monotonically decreasing* if $s_{n+1} \leq s_n$ for every n.
- *monotonic* if it is either monotically increasing or monotonically decreasing.

Theorem 11.4. *Let $(s_n)_n$ be a sequence in $\mathbb{R}$. If $(s_n)_n$ is increasing and bounded above, then $(s_n)_n$ is convergent. If the sequence is decreasing and bounded below, then $(s_n)_n$ is convergent.*

Proof. First suppose $(s_n)_n$ is increasing and bounded above. We'll show it is convergent. Let

$$E = \{s_n\}.$$

Since E is bounded above, its supremum exists, so set $\ell = sup(E)$. Let $\epsilon > 0$. Note $\ell - \epsilon$ is not an upper bound for E. Since the sequence is increasing, there exist N_0 such that for all $n \geq N_0$, we have

$$\ell - \epsilon \leq s_{N_0} \leq s_n \leq \ell$$

so $-\epsilon \leq s_n - \ell \leq \implies |s_n - \ell| \leq \epsilon$ for all $n \geq N_0$. This shows that the sequence converges to ℓ.

The proof for bounded below is exactly the same. But instead of the supremum, you take the infimum. $\square$

Definition 11.5. Let $(s_n)_n$ be a sequence of real numbers. We say that $(s_n)_n$ diverges to $+\infty$ and we write $\lim_{n\to\infty} s_n = +\infty$, if for every $M > 0$, there exist N such that for all $n \geq N$, we have $s_n > M$.

Similarly, we say that $(s_n)_n$ diverges to $-\infty$ and write $\lim_{n\to\infty} s_n = -\infty$ if for every $-M < 0$, there exist N such that for all $n \geq N$, we have $s_n < -M$.

Note if $(s_n)_n$ is increasing and unbounded, then $\lim_{n\to\infty} s_n = +\infty$. Why? Because to be unbounded, means that for every M, there exist some N_0 such that $s_{n_0} > M$. But since the sequence is increasing, we have $s_n \geq s_{n_0} > M$ for all $n \geq N_0$.

Similarly if $(s_n)_n$ is decreasing and unbounded, then $\lim_{n\to\infty} s_n = -\infty$. Similar proof.

Let $(s_n)_{n\geq 1}$ be a sequence of real numbers. Then define sequences $t_n = \sup\{s_k | k \geq n\}$ and $u_n = \inf\{s_k | k \geq n\}$. Note that t_n is decreasing, and u_n is increasing.

We could have that $\lim_{n\to\infty} t_n = -\infty$ and $\lim_{n\to\infty} u_n = +\infty$. But interesting things happen when these are finite.

Proposition 11.6. *Suppose $(s_n)_n$ is*

- *bounded above. Then t_n is well-defined and decreasing.*
- *bounded below, then u_n is well-defined and increasing.*

If $(s_n)_n$ is bounded. Then we define the upper limit

$$\limsup_{n\to\infty} s_n = \lim_{n\to\infty} t_n$$

(and note that t_n here is decreasing and bounded below). Furthermore,

$$\liminf_{n\to\infty} s_n = \lim_{n\to\infty} u_n$$

and note the $(u_n)_n$ is increasing and bounded above.

Example 11.7. What is $\lim_{n\to\infty} \frac{1}{n^k}$? If $k > 0$, then the limit is zero. If $k = 0$, the limit is 1 since the sequence is a constant 1. And if $k < 0$, then $\lim_{n\to\infty} \frac{1}{n^k} = +\infty$.

Proof. If $k = 0$, then it's clear the limit is 1. For $k > 0$, fix $\epsilon > 0$. We want to find integer N_0 such that for all $n \geq N_0$, we have

$$\frac{1}{n^k} \leq \frac{1}{N_0^k} < \epsilon.$$

So one can take N_0 to be the smallest integer greater than $(\frac{1}{\epsilon})^{\frac{1}{k}}$, and then we're done.

If $k < 0$, then let $M > 0$, we can take N_0 to be the smallest integer greater than $M^{1/k}$. Then for $n \geq n_0$, we have

$$\frac{1}{n^k} \leq \frac{1}{n_0^k} > M.$$

□

Now that we know that this sequence for $k > 0$ goes to 0, the next time you want to show a sequence has limit 0, keep in the mind the following lemma:

Lemma 11.8. *If $(x_n)_n$ and $(s_n)_n$ are two sequences of complex numbers and $|x_n| \leq |s_n| < \epsilon$ and $\lim_{n\to\infty} s_n = 0$, then $\lim_{n\to\infty} x_n = 0$.*

Example 11.9. Consider $p > 0$. Then $\lim_{n\to\infty} p^{\frac{1}{n}} = 1$.

Let $x_n = p^{1/n} - 1$. We want to show that $\lim_{n\to\infty} x_n = 0$. Suppose $p \geq 1$.

Then by the binomial theorem,

$$1 + nx_n \leq (1 + x_n)^n = p,$$

so $0 \leq x_n \leq \frac{p-1}{n}$. We see that this forces $\lim_{n\to\infty} x_n = 0$.

Now suppose $p < 1$. Then $\frac{1}{p} \geq 1$. Then $\lim_{n\to\infty} (\frac{1}{p})^{\frac{1}{n}} = 1$. So

$$\lim_{n\to\infty} \frac{1}{(\frac{1}{p})^{\frac{1}{n}}} = \lim_{n\to\infty} p^{\frac{1}{n}} = 1.$$

Example 11.10. $\lim_{n\to\infty} n^{\frac{1}{n}} = 1$. Let $x_n = n^{\frac{1}{n}} - 1 \geq 0$. Then $1 + nx_n \leq n$ and so $x_n \leq 1 - \frac{1}{n}$.

Then

$$\frac{n(n-1)}{2} x_n^2 \leq (1 + x_n) = n,$$

so $x_n^2 \leq \frac{2}{n-1}$. So $0 \leq x_n \leq \frac{\sqrt{2}}{\sqrt{n-1}}$.

Here is an exercise to the reader: let $p > 0, \alpha \in \mathbb{R}$. Then show

$$\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0.$$

And for $|x| < 1$,

$$\lim_{n\to\infty} x^n = 0.$$

Let's move on to the next topic: series.

Definition 11.11. Let $(a_n)_{n\geq 1}$ be a sequence of complex numbers. We associated a sequence

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

Given a sequence $(a_n)_n$, we want to talk about $\sum_n a_n$. But this is an infinite sum, so there is some limiting process involved here. Thus, we use the partial sums s_n as limiting terms to discuss what the infinite sum might be.

Proposition 11.12. *If $(s_n)_n$ is convergent, then we say the the series $\sum_{n=1}^{\infty} a_n$ is convergent and the limit is defined as*

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

12. 10/18/23: Cauchy criterions for series, examples of series

Recall that we defined a series as $\sum a_n$, where we write the partial sums as

$$S_N = \sum_{k=1}^{N} a_k$$

If $\{S_N\}$ is convergent, then we say that $\sum a_n$ is a convergent series.

Note that $\sum_{n\geq M} a_n$ is a convergent series $\iff$ $\sum_{n\geq 0} a_n$ is a convergent series.

Theorem 12.1 (Cauchy Criterion)**.** *The series $\sum a_n$ of complex numbers is convergent $\iff$ $\forall \epsilon > 0, \exists N \geq 1$ such that $\forall m \geq n \geq N$, we have*

$$|\sum_{k=n}^{m} a_n| < \epsilon.$$

Proof. Let $S_N = \sum_{k=1}^{N} a_k$. If $\sum a_n$ is convergent then S_N is cauchy. This shows the forward direction.

Let's show the reverse direction. The hypothesis that $\forall \epsilon > 0$, there exists $N \geq 1$ s.t. $\forall m \geq n \geq N$, we have

$$|\sum_{k=n}^{m} a_n| < \epsilon$$

is saying that the sequence $\{S_N\}$ is cauchy. But note that as a metric space $\mathbb{C}$ can be thought of as $\mathbb{R}^2$ equipped with the euclidean metric. Thus, $\mathbb{C}$ with its standard metric is a complete metric space. Since $\{S_N\}$ is cauchy, this implies that $\{S_N\}$ converges, which implies that $\sum a_n$ is a convergent series. $\square$

As a corollary, this implies that if $\sum a_n$ is convergent, then $\lim_{n\to\infty} a_n = 0$. Why? Because convergence of the series implies $\{S_N\}$ is Cauchy, so for all ϵ, there exists N such that for all $m \geq n \geq N$, we have $|\sum_{k=n}^{m} a_k| < \epsilon$. Take $m = n$, and we find for all $k \geq N$, $|a_k|\epsilon$. This implies $\lim_{n\to\infty} a_n = 0$.

Here is a useful criterion for proving when a series is *not* convergent.

Proposition 12.2. *If $\lim_{n\to\infty} a_n \neq 0$, or doesn't exist, then $\sum a_n$ is not convergent.*

Example 12.3 (geometric series)**.** Let $0 \leq x < 1$. Then $\sum_{n\geq 0} x^n$ is convergent. This is becuase

$$S_N = \sum_{k=0}^{N} x^k = \frac{1 - x^{N+1}}{1 - x}.$$

Since $0 \leq x < 1$, we see that $\lim_{N\to\infty} x^{N+1} = 0$. This implies that S_N converges to $\frac{1}{1-x}$. So $\sum x^n = \frac{1}{1-x}$.

However, if $x = 1$, then the series $\sum x^n = \sum 1$ is divergent.

Theorem 12.4. *Let $\{x_n\}$ be a sequence such that for all n, $a_n \in [0, \infty)$ i.e non-negative. Then $\sum a_n$ is convergent $\iff$ S_N is bounded.*

Proof. Forward direction: if $\sum a_n$ is convergent, then the S_N are convergent. Then S_N must be bounded.

Reverse direction: suppose the S_N are bounded. We can write $S_{N+1} = a_{n+1} + S_N \geq S_N$ since the a_i are nonnegative. Thus, $\{S_N\}$ is a sequence that is increasing and bounded. But this immediately implies convergence of the sequence $\{S_N\}$, and thus the series $\sum a_n$ converges. □

Corollary: if $0 \leq a_n \leq C_n$ for all $n \geq N_0$ for some N_0, then $\sum C_n$ convergent implies that $\sum a_n$ convergent as well.

Proof. We can just consider the terms starting at N_0 and index that as 1. We have that

$$\sum_{k=1}^{n} a_k \leq \sum_{k=1}^{n} C_k \leq B.$$

We get B from $\sum_{k=1}^{n} C_k$ needing to be bounded because the series converges. Thus, $S_N = \sum_{k=1}^{N} a_k$ is increasing and bounded. Thus, it must converge. □

Another corollary: if $a_n \geq c_n \geq 0$, and $\sum c_n$ diverges, then $\sum a_n$ diverges.

Remark: If $\lim_{n\to\infty} a_n = 0$, the series $\sum a_n$ may or may not converge.

Example 12.5. A famous series is the Riemann-zeta series:

$$\xi(s) = \sum_{n \geq 1} \frac{1}{n^s}.$$

Theorem 12.6 (Cauchy)**.** *Let $a_1 \geq a_2 \geq a_3 \geq \cdots$ be a decreasing sequence of non-negative numbers. Then $\sum_{n\geq 1} a_n$ converges $\iff$ $\sum_{k \geq 0} 2^k a_{2^k}$ converges.*

Proof. It is enough to prove that $\sum a_n$ bounded $\iff$ $\sum 2^k a_{2^k}$ is bounded.

Let $S_n = \sum_{\ell=1}^{n} a_\ell$ and $t_k = \sum_{\ell=1}^{n} 2^\ell a_{2^\ell}$. Let $2^{k-1} \leq n < 2^k$. Note that

$$S_n \leq a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6) + \cdots + (a_{2^{k-1}} + \cdots + a_{2^k - 1}) \leq \sum_{\ell=0}^{k-1} 2^\ell a_{2^\ell} = t_{k-1}.$$

Now if $n > 2^k$, then we have that

$$S_n \geq a_1 + a_2 + (a_3 + a_4) + \cdots + (a_{2^{k-1}} + \cdots + a_{2^k}) \geq \frac{1}{2} a_1 + a_2 + 2a_4 + \cdots + 2^{k-1} a_{2^k} \geq \frac{1}{2} t_k.$$

Thus, we have shown that $\sum a_n$ bounded $\iff$ $\sum 2^k a_{2^k}$ is bounded. □

Theorem 12.7. *We have $\sum_{n \geq 1} \frac{1}{n^p}$ converges $\iff$ $p > 1$.*

Proof. By Cauchy, it is enough to show that $\sum_{k \geq 0} 2^k \frac{1}{2^{kp}} = \sum_{k\geq 0} (\frac{1}{2^{p-1}})^k$ converges. But this is a geometric series, and it converges $\iff$ $\frac{1}{2^{p-1}} < 1 \iff 1 < 2^{p-1} \iff 1 < p$. □

Note that for $0 \leq p < 1$, we have that $\lim_{n\to\infty} \frac{1}{n^p} = 0$, but for $0 \leq p < 1$, the series $\sum \frac{1}{n^p}$ diverges.

Theorem 12.8. *If $p > 1$, then*

$$\sum_{n\geq 2} \frac{1}{n(\log n)^p}$$

is convergent. If $p \leq 1$ then it is divergent.

Proof. Use Cauchy's criterion again. □

Exercise: $\sum_{n\geq 100000} \frac{1}{n \log n (\log \log n)^p}$ converges $\iff$ $p > 1$.

Definition 12.9. A series $\sum_{n\geq 1} a_n$ converges absolutely if $\sum_{n\geq 1} |a_n|$ is convergent.

Theorem 12.10. *If $\sum a_n$ is absolutely convergent, then it is convergent.*

The converse is not always true. Being convergent does not necessarily imply absolute convergence.

Proof. Since $\sum a_n$ is absolutely convergent, we have $\sum_{k=n}^{m} |a_k| < \epsilon$. By triangle inequality, we have

$$|\sum_{k=n}^{m} a_k| \leq \sum_{k=n}^{m} |a_k| < \epsilon.$$

This implies that $\sum_{k=n}^{m} a_k$ is cauchy, thus the series $\sum a_n$ converges.

Here is an example of a series that is convergent but not absolutely convergent:

$$\sum_{n\geq 1} \frac{(-1)^n}{n}.$$

□

13. 10/23/23: ROOT TEST, RATIO TEST, POWER SERIES

Recall that if $\sum a_n$ is a series, then the partial sums $s_N = \sum_{i=1}^{N} a_i$. We say that $\sum a_n$ is a convergent series if the sequence $\{s_N\}$ converges. We say that $\sum |a_n|$ is absolutely convergent if the sequence of $\xi_N = \sum_{i=1}^{N} |a_i|$ is convergent.

Note that absolute convergence of $\sum a_n$ implies convergence of $\sum a_n$. But the converse is not true. For example, $\sum_{n\geq 1} \frac{(-1)^n}{n}$ is convergent but not absolutely convergent.

The root test is one systematic way for testing absolute convergence.

Theorem 13.1 (Root Test)**.** *Let $\sum a_n$ be a series. Let $\alpha = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}}$. Then*

- *If $\alpha < 1$, then $\sum a_n$ converges absolutely.*
- *If $\alpha > 1$, then $\sum a_n$ diverges.*
- *If $\alpha = 1$, anything can happen.*

Proof. Last condition: Here is an example of such a series where anything can happen. Consider $\sum_{n\geq 1} \frac{1}{n^p}$ is convergent for $p > 1$, and divergent for $p = 1$. Suppose

$$\alpha = \lim \sup_{n\to\infty} |a_n|^{\frac{1}{n}} = 1.$$

Then note that $(\frac{1}{n^p})^{\frac{1}{n}} = e^{\frac{1}{n}\log\frac{1}{n^p}} = e^{\frac{-p\log n}{n}}$, and as $n \to \infty$, this converges to $e^0 = 1$, as $\log n/n \to 0$ as $n \to \infty$.

First condition: assume that $\alpha < 1$ and that $\alpha < \beta < 1$. Then there exist N_0 s.t. for all $N \geq N_0$, we have

$$b_N = \sup_{n \geq N} |a_n|^{\frac{1}{n}} \leq \beta.$$

So for all $n \geq N_0$, we have $|a_n| \leq \beta^n$ and $\sum \beta^n$ is convergent since $\beta < 1$. This implies that $\sum |a_n|$ is convergent.

Second condition: if $\alpha > 1$, then there exist $\sigma : \mathbb{N} \to \mathbb{N}$ such that $|\alpha_{\sigma(n)}| \to \alpha$ as $n \to \infty$. Let $1 < \beta < \alpha$. There exist n_0 such that for $n \geq n_0$, we have $|\alpha_{\sigma(n)}|^{\frac{1}{\sigma(n)}} \geq \beta \implies |\alpha_{\sigma(n)}| \geq \beta^{\sigma(n)} \geq 1$. So $\lim_{n\to\infty} a_n$ is nonzero, thus $\sum a_n$ cannot converge. □

Theorem 13.2 (Ratio Test)**.** *Let $\sum a_n$ be a series with $a_n \neq 0$ for every $n \geq 1$.*

(1) $\sum a_n$ converges if $\limsup_{n\to\infty} |\frac{a_{n+1}}{a_n}| < 1$.
(2) $\sum a_n$ diverges if there $\exists n_0$ such that for all $n \geq n_0$, we have $|\frac{a_{n+1}}{a_n}| \geq 1$.

Proof. First we prove the first item. Let $\alpha < \beta < 1$, where $\alpha = \limsup_{n\to\infty} |\frac{a_{n+1}}{a_n}|$. Then there exist N_0 such that for all $n \geq 1$, we have $|\frac{a_{N_0+n}}{a_{N_0}}| < \beta$. Then there exist N_0 such that for all $n \geq 1$, we have

$$|a_{n+N_0}| < \beta^n |a_{N_0}| \implies \sum_{N \geq N_0} |a_N| < (\sum_{N \geq 0} \beta^N)|a_{N_0}|.$$

note that since $\sum \beta^N$ is convergent, this forces $\sum |a_n|$ to be convergent.

For the second item: note $|a_n| \neq 0$ for every $n \geq 1$. The second condition implies that $\lim_{n\to\infty} a_n \neq 0$, so the sum $\sum a_n$ cannot possibly converge. □

Example 13.3. For $x \in \mathbb{R}$, consider $\sum \frac{x^n}{n!}$. This is indeed the Taylor series expansion of e^x centered at $x = 0$. And it converges for any value of x. In fact it is absolutely convergent, since it holds true for all positive $x \in \mathbb{R}$. When x is negative, then $\frac{|x|^n}{n!}$ is still convergent.

Why does this converge everywhere? If $x = 0$, the series is just 1. If $x \neq 0$, then

$$\frac{x^{n+1}/(n+1)!}{x^n/n!} = \frac{x}{n+1}$$

whose limit is 0 as $n \to \infty$ since x is fixed and n grows. Thinking carefully about the lim sup, by the ratio test, this implies that the sum converges.

Note that $e = \sum_{n\geq 0} \frac{1}{n!} \notin \mathbb{Q}$. (See Rudin).

Proposition 13.4. *Let $\{c_n\}_{n\geq 1}$ be a sequence. Then*

(1) $\limsup |c_n|^{\frac{1}{n}} \leq \limsup |\frac{c_{n+1}}{c_n}|$
(2) $\liminf |\frac{c_{n+1}}{c_n}| \leq \liminf |c_n|^{\frac{1}{n}}$.

Proof. This is 3.3.7 in Rudin. This is fun to do at home! (and not so fun in front of people on the blackboard). □

Note this proposition implies that if the ratio test works, i.e that $\limsup |\frac{c_{n+1}}{c_n}| < 1$, then the root test works as well. But the converse is not true.

Now we'll discuss a special class of series: power series.

Definition 13.5. Let $\{c_n\}_n$ be a sequence of complex numbers. The series $\sum_{n\geq 0} c_n z^n$ for $z \in \mathbb{C}$ is called a power series, with $\{c_n\}$ the coefficients of the power series.

Aside: power series play a very important role in complex analysis. In a rough sense, there is a notion of holomorphicity in complex analysis that is the complex-analog of differentiability in real analysis. Here's a miracle: complex functions that are holomorphic everywhere can be written as power series! This is not true for real differentiable functions – the taylor series of a differentiable function in $\mathbb{R}$ is just an approximation.

Example 13.6. Let $z = x + iy$. Then

$$e^z = e^{x+iy} = e^x e^{iy} = e^x(\cos y + i \sin y).$$

We have that $e^z = \sum_{n\geq 0} \frac{z^n}{n!}$ is a power series which absolutely converges for all $z \in \mathbb{C}$.

Lemma 13.7. *If $\sum_{n\geq 0} c_n z_0^n$ converges at nonzero $z_0 \in \mathbb{C}$, then $\sum c_n z^n$ converges absolutely for all $|z| < |z_0|$.*

Proof. We have $\lim_{n\to\infty} c_n z_0^n = 0$. This implies that $\{c_n z_0^n\}$ is bounded. So there exist M such that for all $n \geq 1$, we have $|c_n z_0^n| \leq M$. Let $|z| < |z_0|$. Then

$$|c_n z^n| = |c_n||z_0|^n(\frac{|z|}{|z_0|})^n \leq M\beta^n.$$

where $\beta = |\frac{z}{z_0}| < 1$. ☐

In fact, when we have absolute convergence, we have the radius convergence is not strict.

Lemma 13.8. *If $\sum_{n\geq 0} c_n z_0^n$ converges absolutely at nonzero $z_0 \in \mathbb{C}$, then $\sum c_n z^n$ converges absolutely for all $|z| \leq |z_0|$.*

Remark: if $\sum c_n z^n$ only converges (not absolutely convergent), it need not be convergent for all $|z_0| = 1$. For example, take $\sum \frac{z^n}{n}$. This is convergent at $z = -1$, but divergent at $z = 1$.

Definition 13.9 (Radius of convergence)**.** Let $R = \sup\{r | \forall z \text{ such that } |z| < r, \sum c_n z^n \text{ converges }\}$. Then R is called the radius of convergence of $\sum_{n\geq 0} c_n z^n$.

If $\sum c_n z^n$ converges only at $z = 0$, then we say the radius of convergence is 0. Often, anything can happen on the radius $|z| = R$.

Theorem 13.10. *We have that the radius of convergence is*

$$R = \frac{1}{\limsup_{n\to\infty} |c_n|^{\frac{1}{n}}}.$$

Proof. Use the root test! ☐

Example 13.11. The radius of convergence of $\sum_{n\geq 0} n^n z^n$ is 0 by the root test.

Example 13.12. The radius of convergence of $\sum_{n\geq 1} \frac{z^n}{n^2}$ is 1

Example 13.13. The radius of convergence of $\sum_{n\geq 1} \frac{z^n}{n^p}$ for $p \in \mathbb{R}$ is 1.

Note that when $z = 1$, $p = 1$, the second example still converges, but the third example does not.

14. 10/25/23: ALTERNATING SERIES, SUMS/PRODUCTS OF SERIES, REARRANGEMENTS

We'll discuss a technique today called summation by parts. Let $\sum a_n$ and $\sum b_n$ be two series. Consider the series $\sum a_n b_n$. Assume that b_n satisfies the following properties:

- $b_n \geq b_{n+1}$ for every n
- $\lim_{n\to\infty} b_n = 0$ (note these two imply $b_n > 0$).

Furthermore, the sequence of partial sums $S_N = \sum_{k=1}^{N} a_k$ is bounded. Then

Theorem 14.1. *With the aforementioned assumptions, the series $\sum a_n b_n$ is convergent.*

Proof. Let $\epsilon > 0$. We are looking for N_0 such that for $m \geq n \geq N_0$, we have that

$$|\sum_{k=n}^{m} a_k b_k| < \epsilon.$$

We want to exploit the fact that the partial sums $S_N = \sum_{k=1}^{N} a_k$ are bounded. Note that $a_k = s_k - s_{k-1}$. This implies that

$$\sum_{k=n}^{m} a_k b_k = \sum_{k=n}^{m}(s_k - s_{k-1})b_k = \sum_{k=n}^{m} s_k b_k - \sum_{k=n}^{m} s_{k-1}b_k = \sum_{k=n}^{m} s_k b_k - \sum_{k=n-1}^{m-1} s_k b_{k+1}$$

$$= -s_{n-1}b_n + s_m b_m + \sum_{k=n}^{m-1} s_k(b_k - b_{k+1}).$$

note that $b_k - b_{k+1} \geq 0$ always. Taking the absolute value, we have

$$|\sum_{k=n}^{m} a_k b_k| = |s_{n-1}||b_n| + |s_m||b_m| + \sum_{k=n}^{m-1} |s_k||b_k - b_{k+1}|.$$

Note we can find $A > 0$ such that $|S_p| \leq A$ for all $p \geq 0$. So our inequality becomes

$$|\sum_{k=n}^{m} a_k b_k| \leq A(b_n + b_m + \sum_{k=n}^{m-1} b_k - b_{k+1} \leq 2Ab_n.$$

Thus, we can find N_0 such that for all $n \geq N_0$ we have $b_n \leq \frac{\epsilon}{2A}$. Then for $m \geq n \geq N_0$, we have that

$$|\sum_{k=n}^{m} a_k b_k| \leq \epsilon.$$

By the cauchy criterion, this implies that $\sum a_k b_k$ converges. □

Here are some corollaries. As we'll see, this theorem is quite useful when trying to show that a series is convergent when it is not absolutely convergent.

Example 14.2. The series $\sum_{n\geq 1} \frac{(-1)^n}{n}$ is convergent. Apply the above theorem!

More generally, we have the following alternating series corollary:

Proposition 14.3 (Alternating series)**.** *Let $(c_k)_k$ be a sequence such that*

- $(|c_k|)_k$ *is a decreasing sequence, i.e* $|c_k| \geq |c_{k+1}| \geq |c_{k+2}|$.
- $c_{2k} \geq 0$ *and* $c_{2k-1} \leq 0$ *(alternating signs)*
- $\lim_{k\to\infty} c_k = 0$.

Then $\sum c_k$ *is convergent.*

Proof. Apply the previous theorem. Let $b_n = |c_n|$, and $a_n = (-1)^n$. Then $a_n b_n = c_n$. Check that the sequences b_n and a_n satisfy the necessary properties.

Exercise: check that the limit of the c_k sits between

$$\sum_{k=1}^{\infty} c_{2k+1} \leq \sum_{k=1}^{\infty} c_k \leq \sum_{k=1}^{\infty} c_{2k}.$$

Note that the sequence of $\{c_{2k+1}\}$ is increasing, and that $\{c_{2k}\}$ is decreasing. □

Theorem 14.4. *Let* $\sum_{n\geq 0} c_n z^n$ *be a power series with radius of convergence 1. Assume that the* $c_n \geq 0$*, are decreasing, and* $\lim_{n\to\infty} c_n = 0$*. Then* $\sum_{n\geq 0} c_n z^n$ *is convergent for every* z *with* $|z| = 1$ *except possibly at* 1.

Proof. It suffices to prove that $\sum_{n=0} z^n$ is bounded. We have

$$|\sum_{k=0}^{N} z^k| = |\frac{1 - z^{N+1}}{1-z}| = \frac{|1 - z^{N+1}|}{|1-z|} \leq \frac{2}{|1-z|}.$$

Thus, if $z \neq 1$, the series is convergent. □

Example 14.5. Consider $e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$. This series is convergent for $z = e^{i\theta}$ for all θ except $\theta = 2\pi k$ where $k \in \mathbb{Z}$, at which $e^{i2\pi k} = 1$.

We discussed several classes ago addition and multiplication of sequences. We also have addition and multiplication of series.

Theorem 14.6. *If* $\sum_{n=0}^{\infty} a_n$ *and* $\sum_{n=0}^{\infty} b_n$ *are (absolutely) convergent, then* $\sum_{n=0}^{\infty} a_n + b_n$ *is (absolutely) convergent, with*

$$\sum_{n=0}^{\infty} a_n + b_n = \sum_{n=0}^{\infty} a_n + \sum_{n=0}^{\infty} b_n.$$

To discuss products of series, it will be convenient to use the following idea: note if $c_n = \sum_{k=0}^{n} a_k b_{n-k}$ for $n \geq 0$, then

$$(\sum_{n\geq 0} a_n z^n)(\sum_{n\geq 0} b_n z^n) = \sum c_n z^n.$$

Thus, in general, if we have a product of series

$$(\sum_{k=0}^{\infty} a_k z^k)(\sum_{k=0}^{\infty} b_k z^k) = \sum_{k_1,k_2} a_{k_1} b_{k_2} z^{k_1+k_2},$$

then if we let $c_n = \sum_{k_1+k_2=n} a_{k_1} b_{k_2}$, we can write

$$(\sum_{k=0}^{\infty} a_k z^k)(\sum_{k=0}^{\infty} b_k z^k) = \sum_{k_1,k_2} a_{k_1} b_{k_2} z^{k_1+k_2} = \sum_{n=0}^{\infty} c_n z^n.$$

Theorem 14.7. *Assume that* $\sum_{n\geq 0} a_n$ *is absolutely convergent, and* $\sum_{n\geq 0} b_n$ *is convergent. Then* $\sum c_n$ *is convergent, where* $c_n = \sum_{k_1+k_2=n} a_{k_1} b_{k_2} = \sum_{k=0}^{n} a_k b_{n-k}$. *Furthermore,*

$$\sum c_n = (\sum a_n)(\sum b_n).$$

Let's talk about rearrangements – we've been talking about series, but does the order in which you add terms in your series matter? What if you rearrange some of the terms? More precisely..

Let $\sum a_n$ be a series, and $\sigma : \mathbb{N} \to \mathbb{N}$ is a bijection. Consider the rearrangement $\sum a_{\sigma(n)}$. If $\sum a_n$ is convergent, what can be said about the rearrangement $\sum a_{\sigma(n)}$.

Theorem 14.8. *If $\sum_{n\geq 0} a_n$ is absolutely convergent, then $\sum_{n\geq 0} a_{\sigma(n)}$ is also absolutely convergent (for any bijection $\sigma : \mathbb{N} \to \mathbb{N}$) and $\sum a_{\sigma(n)} = \sum a_n$.*

Proof. We have that $\sum_{k=0}^{N} |a_{\sigma(k)}| \leq \sum_{k=0}^{\infty} |a_k|$. This implies that $\sum_{k\geq 0} a_{\sigma(n)}$ is also absolutely convergent by increasing and bounded. Note the above inequality implies that $\sum_{n=0}^{\infty} |a_{\sigma(n)}| \leq \sum |a_n|$. But this holds for all σ. So we can apply σ^{-1} to obtain

$$\sum_{n=0}^{\infty} |a_{\sigma(n)}| \leq \sum |a_n| \leq \sum_{n=0}^{\infty} |a_{\sigma(n)}|.$$

Finally, we know that $\sum_{n\geq 0} a_\sigma(n)$ is convergent since it is absolutely convergent. But why does it converge to the same value as $\sum a_n$? Fix $\epsilon > 0$, and N such that for all $m \geq n \geq N$, we have

$$|\sum_{k=n}^{m} a_k| \leq \epsilon.$$

Let $p \geq 1$ be an integer such that

$$\{1, \cdots, N\} \subseteq \{\sigma(1), \cdots, \sigma(p)\}.$$

Let $S'_m = \sum_{k=0}^{m} a_{\sigma(k)}$ and $S_m = \sum_{k=0}^{m} a_k$. Then for all $m \geq N_0 = \max\{p, N\}$, we have

$$|S'_m - S_m| < \epsilon.$$

This implies that $\lim_{m\to\infty} S'_m = \lim_{m\to\infty} S_m$. So $\sum a_n$ and $\sum a_{\sigma(n)}$ converge to the same value. $\square$

We won't prove this, but the following proposition from Rudin says that if $\sum a_n$ is convergent but not absolutely convergent, then we have very little control over an arbitrary rearrangement. So convergent sums which are not absolutely convergent depend on the order.

Proposition 14.9. *If $\sum a_n$ is convergent but not absolutely convergent, then for any $\alpha \leq \beta$, there exist $\sigma : \mathbb{N} \to \mathbb{N}$ such that*

$$\liminf \sum a_{\sigma(n)} = \alpha, \limsup \sum a_{\sigma(n)} = \beta.$$

15. 10/30/23: LIMITS AND CONTINUITY

Today we'll discuss limits and continuity. Let (X, d) and (Y, d) be two metric spaces (note that the metrics on X and Y need not be the same, we're just abusing notation by calling them both d here). Let $E \subseteq X$ be a subset. Let $f : E \to Y$. As x gets closer to p in E, what happens to $f(x)$? Does it approach $f(p)$? We would like such a notion. Otherwise, our f would be quite pathological. In our daily life and in other sciences, such a notion would be extremely useful.

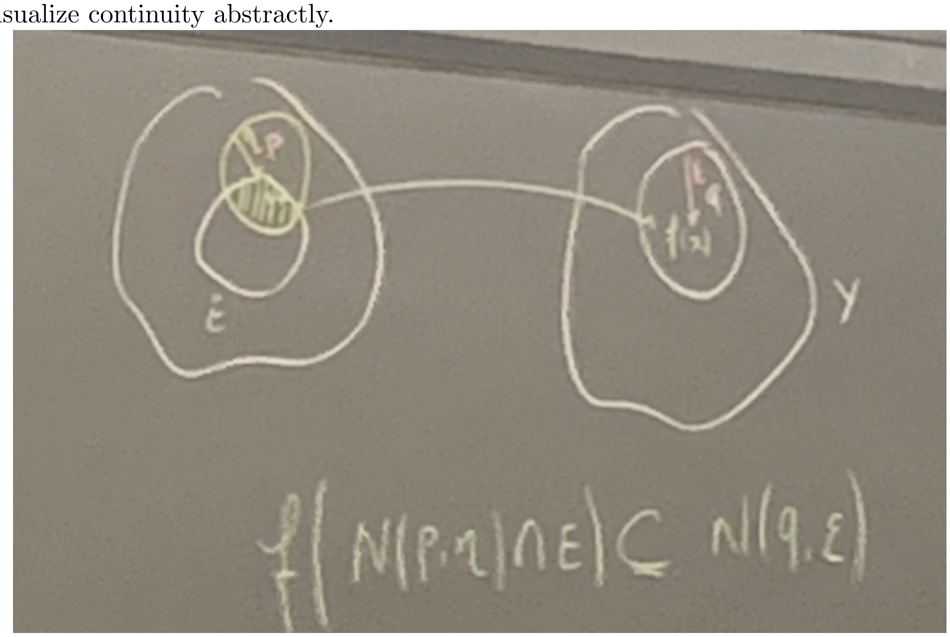
Definition 15.1. Let $f : E \to Y$, p a limit point of E, and $q \in Y$. We say that

$$\lim_{x \to p} f(x) = q,$$

or $f(x) \to q$ as $x \to p$, if for every $\epsilon > 0$, there exists $\eta > 0$ such that for every $x \in E$ not equal to p,

$$0 < d(x, p) < \eta \implies d(f(x), q) < \epsilon.$$

Note from our definition that p is a limit point in E. In other words, p need not be in E. Of course, $q \in X$, the ambient metric space of E. Here's how you should visualize continuity abstractly.



Example 15.2. Here's an example of a function that is not continuous somewhere. Consider some $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = 0$ for $x \neq 2$, and $f(2) = 1$. In this case, we see that this function is not continuous at $x = 2$. Why? Because, say, for $\epsilon = \frac{1}{2}$, there does not exist any η such that

$$x \in (2 - \eta, 2 + \eta) \implies f(x) \in (1 - \frac{1}{2}, 1 + \frac{1}{2}).$$

In fact, for all $x \in (2 - \eta, 2 + \eta)$ such that $x \neq 2$, we have $f(x) = 0$.

Example 15.3. Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = x^2$. Then

$$\lim_{x \to 2} f(x) = 4.$$

Why? Choose any $\epsilon > 0$. Let's find $\eta > 0$ such that

$$0 < |x - 2| < \eta \implies |x^2 - 4| < \epsilon.$$

Note that if we found such an η, then

$$|x^2 - 4| = |x - 2||x + 2| < \eta|x + 2| < \epsilon.$$

Note that we can suppose $\eta < 1$, since once we've found an η, any other smaller quantity still satisfies the implication of inequalities that we want for continuity.

Then $|x+2| \leq |x-2|+4 < 5$. By the triangle inequality, then

$$\eta|x+2| < 5\eta < \epsilon.$$

So we can take $\eta = \min\{1, \frac{\epsilon}{6}\}$. Indeed, we have

$$0 < |x-2| < \frac{\epsilon}{6} \implies |x^2-4| < \epsilon.$$

Example 15.4. We have

$$\lim_{x\to 1} \frac{x^3-1}{x-1} = \lim_{x\to 1} x^2+x+1 = 3.$$

Let $\epsilon > 0$. We want to find η such that

$$0 < |x-1| < \eta \implies |x^2+x+1-3| < \epsilon.$$

Note $x^2+x-2 = (x+2)(x-1)$. So $|x^2+x+1-3| = |x+2||x-1|$. We can take $\eta = \min(1, \frac{\epsilon}{4})$. Then $0 < |x-1| < \eta$ implies $|x-1| < 1, \frac{\epsilon}{4}$, so $|x+2| \leq 4$ and $|x-1| < \frac{\epsilon}{4}$, which implies that

$$|xr+x+1-3| = |x+2||x-1| < \frac{\epsilon}{4}4 = \epsilon.$$

Here's a powerful way to intuitively understand continuity. It relates continuity back to our previous discussions of sequences and convergence of sequences.

Theorem 15.5. *Let $f: E \to Y$ where $E \subseteq X$. Let p be a limit point of E, $q \in Y$. The following statements are equivalent:*

(1) $\lim_{x\to p} f(x) = q$
(2) For all $x_n \to p$ where $x_n \in E, x_n \neq p$, we have $f(x_n) \to q$.

Proof. Let's show (1) $\implies$ (2). Assuming the first statement, this translates to: for every $\epsilon > 0$, there exists $\eta > 0$ such that for all $x \in E$ such that $0 < d(x,p) < \eta$, then $d(f(x), q) < \epsilon$. Now suppose we have a sequence $\{x_n\}$ in E such that $x_n \to p$. We want to show $f(x_n) \to q$ as $n \to \infty$. Let $\epsilon > 0$. It is enough that $0 < d(x_n, p) < \eta \implies d(f(x_n), q) < \epsilon$. Since $\lim_{n\to\infty} x_n = p$, we can find N_0 such that for all $n \geq N_0$ we have $0 < d(x_n, p) < \eta$ for all $n \geq N_0$. Then for all $n \geq N_0$, we have $d(f(x_n), q) < \epsilon$.

Now we show (2) $\implies$ (1). Assume for the sake of contradiction that statement 1 is false, i.e $\lim_{x\to p} f(x) \neq q$ (it may or may not even exist!). This implies that there exist $\epsilon > 0$ such that *for every* $\eta > 0$, there *exists* $x \in E$ where $0 < d(x,p) < \eta$ and $d(f(x), q) \geq \epsilon$. If $\eta = \frac{1}{n}$, then there exists $x_n \in E$ such that

$$0 < d(x_n, p) < \frac{1}{n}.$$

Then we've found a subsequence x_n such that $x_n \to p$ as $n \to \infty$, but we always have that $d(f(x_n), q) \geq \epsilon$, which implies that we could not have $f(x_n) \to q$ as $n \to \infty$, which contradicts our assumption. Thus, we must have that (2) $\implies$ (1). □

Corollary: let $f: E \to Y$ and p limit point of E in X. If the limit $\lim_{x\to p} f(x)$ exists, then it is unique.

Proposition 15.6. *Let $f, g: E \to \mathbb{C}$. Let p be a limit point of E such that*

$$\lim_{x\to p} f(x) = \alpha, \lim_{x\to p} g(x) = \beta.$$

Then

- *(addition)* $\lim_{x\to p}(f+g)(x) = \alpha + \beta$
- *(multiplication)* $\lim_{x\to p}(fg)(x) = \alpha\beta$
- *(nonzero limit, then nonzero on neighborhood) If $\alpha \neq 0$, then there exists $\eta > 0$ such that for every x such that $0 < d(x,p) < \eta$, we have $f(x) \neq 0$.*
- *(nonnegative limit) If $f(x) \geq 0$ for $0 < d(x,p) \leq \eta$ for some $\eta > 0$, then*
$$\alpha = \lim_{x\to p} f(x) \geq 0.$$
- *(squeezing) If $0 \leq f(x) \leq g(x)$ for every x such that $0 < d(x,p) < \eta$, then*
$$0 \leq \lim_{x\to p} f(x) \leq \lim_{x\to p} g(x).$$

Note in the fourth item, if you replace the inequality $f(x) \geq 0$ with a strict inequality $f(x) > 0$, then the inequality $\lim_{x\to p} f(x) \geq 0$ need not be strict. For example, consider $f(x) = x^2$. For every $x \in (-1,1) \setminus \{0\}$ we have $f(x) > 0$. But $\lim_{x\to 0} f(x) = 0$.

16. 11/1/23

We give an equivalent definition of continuity. This definition is must easier to recall, so if you have to choose between definitions to remember, I would recommend this one. Let (X, d_x) and (Y, d_Y) be metric spaces, and $U \subseteq X$.

Definition 16.1. A function $f : U \to Y$ is continuous at a point $x_0 \in U$ if $\forall \epsilon > 0, \exists \delta > 0$ such that
$$d_Y(f(x), f(x_0)) < \epsilon$$
whenever $x \in d_X(x, x_0) < \delta$.

Then f is continuous on U if its continuous at every $x_i \in U$.

Theorem 16.2. *Let $f : U \to Y$ be continuous at $x_0 \in U \iff$ either*

(1) x_0 is an isolated point of U
(2) or $\lim_{x\to x_0} f(x) = f(x_0)$.

Proof. If x_0 is not isolated, this is left to the reader; it is an exercise in unpackaging definitions. If x_0 is isolated, choose $\delta > 0$ such that the open ball of radius δ at x_0 intersecting U contains only x_0, and no other points. Then for any point in this neighborhood of radius δ around x_0, which is just x_0, we see that indeed $d_Y(f(x), f(x_0)) = d_Y(f(x_0), f(x_)) = 0 < \epsilon$ for any ϵ. □

Example 16.3. Here are some examples of continuous functions.

(1) Polynomial functions are conitnuous
(2) Rational functions (fractions of polynomials) are continuous where they are well-defined
(3) $e^x, \sin x, \cos x$
(4) x^a where $a > 0$, is continuous on $x \geq 0$.

Let (X, d_x) and (Y, d_Y) be metric spaces, and $U \subseteq X$. Let (Z, d_Z) be a metric space, and $g : f(U) \to Z$.

Proposition 16.4. *If f is continuous at x_0 and g is continuous at $f(x_0)$, then the composition $h := g \circ f$ is continuous at $x_0 \in U$.*

Proof. Let $\epsilon > 0$. Since g is continuous at $f(x_0)$, there exists $\eta > 0$ such that for all $y \in f(U)$, we have

$$d_Y(y, f(x_0)) < \eta \implies d_Z(g(y), h(x_0)) < \epsilon.$$

Since f is continuous at x_0, there exist δ such that for all $x \in U$,

$$d_X(x, x_0) < \delta \implies d_Y(f(x), x_0) < \eta.$$

But combining these two statements, we have for $x \in U$,

$$d_X(x, x_0) < \delta \implies d_Y(f(x), f(x_0)) < \eta \implies d_Y(h(x), h(x_0)) < \epsilon.$$

Thus, h is continuous at x_0. □

Is the converse true? No. Here is a counterexample. Suppose you define f on $[0,1] \cup [2,\infty)$, where $[0,1]$ $f(x) = x$, and on $[2,\infty)$, $f(x) = x + \frac{1}{x}$. Then consider taking f^{-1}. This inverse f^{-1} will be discontinuous, but their composition is simply identity.

Here's a more topological characterization of continuity.

Theorem 16.5. *$f : U \to Y$ is continuous $\iff$*

$$f^{-1}(V) = \{x \in U | f(x) \in V\}$$

is an open subset of U, for every open subset $V \subseteq Y$.

Proof. Forward direction: suppose f is continuous and let $V \subseteq Y$ be an open subset. To show that $f^{-1}(V)$ is open in U, it suffices to show that for every x_0 in $f^{-1}(V)$, there exist δ such that $B_X(x_0, \delta) \subset f^{-1}(V)$. Note since V is open, there exist some ϵ such that $B_Y(f(x_0), \epsilon) \subseteq V$. But since f is continuous, there exist δ such that

$$d_X(x, x_0) \implies d_Y(f(x), f(x_0)) < \epsilon$$

which means that $f(B_X(x_0, \delta)) \subseteq B_Y(f(x_0), \epsilon) \subseteq V$. Thus, $B_X(x_0, \delta) \subseteq f^{-1}(V)$. Thus, $f^{-1}(V)$ is open.

Reverse direction: suppose for every open subset $V \subseteq Y$, we have $f^{-1}(V)$ is open in X. Then note that for any $x_0 \in X$, we can look at $f(x_0) \in Y$. Then take any ϵ-neighborhood $B_Y(f(x_0), \epsilon)$ of $f(x_0)$. Then by assumption, $f^{-1}(B_Y(f(x_0), \epsilon))$ is open. By definition,

$$f^{-1}(B_Y(f(x_0), \epsilon)) = \{x \in U | d_Y(f(x), f(x_0)) < \epsilon\}.$$

Since this is open, we can find δ such that $B_Y(x_0, \delta) \subset f^{-1}(B_Y(f(x_0), \epsilon))$. But this immediately implies we have found δ such that

$$d_X(x, x_0) < \delta \implies d_Y(f(x), f(x_0)) < \epsilon.$$

Thus, f is continuous at x_0. This is true for every x_0, so f is continuous on U. □

Equivalently:

Proposition 16.6. *$f : U \to V$ is continuous on U $\iff$ $f^{-1}(V)$ is closed in U for every closed $V \subset Y$.*

Proof. This follows from the previous theorem, combined with the fact that $f^{-1}(Y \setminus V) = U \setminus f^{-1}(V)$. □

So far we've been mostly referring to single variables. Take $Y = \mathbb{R}^n$, equipped with the usual euclidean distance metric. We can think of $f : U \to Y = \mathbb{R}^n$ as an n-tuple of coordinate functions $f(x) = (f_1(x), \cdots, f_n(x))$.

Theorem 16.7. *f is continuous at $x_0 \in U \iff$ each f_i continuous at $x_0 \in U$.*

Proof. Here's the main idea: if $|f(x) - f(x_0)| < \epsilon$ (absolute value here to denote the metric in $\mathbb{R}^n$), then $|f_i(x) - f_i(x_0)| < \epsilon$ (absolute value here denotes the actual absolute value on $\mathbb{R}$). If $|f_i(x) - f_i(x_0)| < \frac{\epsilon}{\sqrt{n}}$ for every $i = 1, \cdots, n$. Then

$$|f(x) - f(x_0)| = (\sum |f_i(x) - f_i(x_0)|^2)^{\frac{1}{2}} < \sqrt{n\frac{\epsilon^2}{n}} = \epsilon.$$

□

Let us return back to the situation $n = 1$, so $Y = \mathbb{R}$.

Proposition 16.8. *Let $f, g : U \to \mathbb{R}$ be both continuous at $x_0 \in U$. Then $f + g, fg$ are continous at x_0 as well, and if $g(x_0) \neq 0$, f/g is also continuous at x_0.*

Proof. Let's first check continuity of $f + g$ at x_0. Consider then function

$$\pi : U \to \mathbb{R} \times \mathbb{R}, \pi(x) = (f(x), g(x))$$

and

$$\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \phi(y, y_0) = y + y_0.$$

Note that $f + g = \phi \circ \pi$. But π is continuous by the previous theorem because its coordinate functions f, g are continuous. Furthermore, ϕ is continuous (which we'll let the reader check for themselves). Then since both ϕ and π are continuous, then their composition, which is $f + g$, is also continuous by another previous theorem we proved.

The fact that fg is continuous, and f/g when $g(x_0) \neq 0$ is continuous, has the same proof, except you define ϕ to be multiplication and division, respectively. So the reader at this point must check that the maps $\mathbb{R} \to \mathbb{R} \to \mathbb{R}$ where $(y, y') \mapsto y + y', (y, y') \mapsto yy'$ are continous, and $(y, y') \mapsto y/y'$ is continuous when $y' \neq 0$. □

Recall that we claimed before that polynomials are continuous. For example, $f(x, y, z) = x^2y^3 + z^4 + 5xyz$ is continuous. This is because $f(x, y, z)$ may be thought of as sums of multiplications of functions of the form $\pi_1(x, y, z) = x, \pi_2(x, y, z) = y, \pi_3(x, y, z) = z$. Thus, polynomials are continuous.

Rational functions $\frac{P(X)}{Q(x)}$ are also continuous where it is well-defined. This is because $P(X), Q(X)$ are polynomials, and thus continuous. So the fraction is also continuous, as long as $Q(x) \neq 0$. Of course, if you enlarge the domain of $\frac{P(X)}{Q(x)}$ to be where it is not well-defined, often it is longer continuous at x where $Q(x) = 0, P(x) \neq 0$.

The function $\mathbb{R}^n \to \mathbb{R}$, where

$$x = (x_1, \cdots, x_n) \mapsto |x| = (\sum_{i=1}^{n} x_i^2)^{1/2}$$

is continuous. This is because the square root function $\sqrt{\cdot}\!:\mathbb{R}_{\geq 0} \to \mathbb{R}$ is continuous on the interval of nonnegative numbers. One way to see this via $\epsilon - \delta$ proof is to use the fact that

$$|\sqrt{y} - \sqrt{x}| < (\sqrt{y} - \sqrt{x})(\sqrt{y} + \sqrt{x}) = y - x.$$

Then since the function $x \mapsto \sum_{i=1}^{n} x_i^2$ is continuous and always ≥ 0, the composition with the square root function is continuous.

17. 11/6/23

Last time we discussed the notion of continuity. Today we'll see how it interacts with other topological notions like compactness.

Definition 17.1. A map $f : E \to \mathbb{R}^n$ is bounded if there exists $M \in R_{\geq 0}$ such that for every $x \in E$, $||f(x)|| \leq M$.

Example 17.2. The identity function $\mathbb{R} \to \mathbb{R}$ is not bounded.

Knowing that a function is bounded is quite nice. It's some sort of global control, which you don't always happen.

Theorem 17.3. *Let X be a compact metric space, Y any metric space, and $f : X \to Y$ continuous. Then the image $f(X)$ is compact.*

Proof. We want to show that $f(X)$ is compact, i.e that for any open cover $\{V_\alpha\}$ of $f(X)$, we have a finite subcover. Well, since f is continuous, then $\{f^{-1}(V_\alpha)\}$ is a collection of opens in the metric space X and they must cover X. By compactness of X, there is a finite subcover
$$f^{-1}(V_1) \cup \cdots \cup f^{-1}(V_n).$$
This implies that $V_1 \cup \cdots \cup V_n$ is a finite subcover of $f(X)$. Thus, $f(X)$ is compact. $\square$

Note it is not true in general that the continuous image of an open or closed remains open or closed.

Proposition 17.4. *Let X be a compact metric space, and $Y = \mathbb{R}^k$. Then $f(X)$ is closed and bounded.*

Proof. The previous theorem told us that $f(X)$ is compact. But by Heine Borel, compactness is equivalent to closed and bounded. $\square$

Note that boundedness is not a topological property. To talk about boundedness, we needed to be in a domain like $\mathbb{R}^n$. You can't describe boundedness purely through topological notions. This is why we introduced compactness. Compactness is some intermediary topological notion that allows us to say when something is bounded when the target is something like $\mathbb{R}^n$. Here's another corollary:

Proposition 17.5. *Let X be compact metric space, and $f : X \to \mathbb{R}$ continuous. Let*
$$M = \sup_{p \in X} f(p) \text{ and } m = \inf_{p \in X} f(p).$$
Then f attains its supremum M and infimum m, i.e there exist $p_M, p_m \in X$ such that $f(p_M) = M, f(p_m) = m$.

Proof. We have that $f(X)$ is closed and bounded. Since it is bounded, it's infimum and supremum exists. Since it is closed, it attains its supremum and infimum. $\square$

Question: when are two mathematical objects the same? The answer should be: there is a map $f : X \to Y$ which admits an inverse map. The structure of X, Y and the structure of this map is contingent on the situation at hand. For example, X, Y could just be sets, so we'd be fine with having f be a map of sets. But if X, Y has some kind of structure, then this map f should not only have an inverse, but both f and its inverse should capture the notion that the structures of X and Y are "the same."

More generally, if X, Y are topological spaces (which includes metric spaces), then X and Y are "the same" or homeomorphic, if there exists a homeomorphism $f : X \to Y$. A homeomorphism f is a continuous bijection which admits a continuous inverse.

So what does it mean to say that boundedness is not a topological property? Well, intuitively, we'd like to say that something like $(0,1)$ is bounded. But there is a homoemorphism between $(0,1) \cong (1,\infty)$ via $x \mapsto \frac{1}{1-x}$. So the next best thing to "boundedness" is compactness, which is preserved by homeomorphisms.

WARNING: not every continuous bijection is a homeomorphism. For example, take $f : [0,1) \to S^1$ (the unit circle in $\mathbb{R}^2$) via $x \mapsto (\cos(2\pi x), \sin(2\pi x))$. The inverse to this is not continuous. If one starts at the circle at $(1,0)$, then travels counterclock-wise, then this is the same as starting at 0 on $[0,1)$, and it slowly approaches 1, but actually it will jump back to 0. This jump is a discontinuity.

Proposition 17.6. *Let X be a compact metric space. Let Y be a metric space. Let $f : X \to Y$ be a continuous bijection. Then f^{-1} is continuous and thus f is a homeomorphism.*

Proof. We want to show that f^{-1} is continuous. So start with an open $U \subset X$. Then consider $X \setminus U$, which is closed. A closed subset of a compact metric space is compact. Then $f(X \setminus U) = Y \setminus f(U)$ is compact. But a compact subset of a metric space is closed. Thus, $Y \setminus f(U)$ is closed. But this implies that $f(U)$ is open. Thus, f^{-1} is continuous. $\square$

Definition 17.7. Let $f : X \to Y$ be a map of metric spaces. Then f is uniformly continuous if $\forall \epsilon > 0$, there exists $\delta > 0$ such that for all $p, q \in X$ with $d(p,q) < \delta$, we have $d(f(p), f(q)) < \epsilon$.

Note: continuity is a local property of a function; uniform continuity is stronger because it is a global version of continutity where it doesn't matter what point you are at.

Theorem 17.8. *Let X be compact and $f : X \to Y$ be continuous. Then f is uniformly continuous.*

Proof. Fix $\epsilon > 0$. Since f is continuous, for every $p \in X$, there exist δ_p such that $d(p,x) < \delta_p \implies d(f(p), f(x)) < \frac{\epsilon}{2}$. So for ϵ, we've obtained an open neighborhood $B(p, \delta_p)$ around every point $p \in X$. Let $U_p = \{x | d(p,x) < \delta_p/2\}$. Then by compactness of X, there is a finite subcover

$$U_{p_1} \cup \cdots \cup U_{p_k}$$

of X. Then take the global δ to be $\delta = \frac{\min\{\delta_{p_1}, \cdots, \delta_{p_k}\}}{2}$.

Now we claim that for every $x_1, x_2 \in X$,

$$d(x_1, x_2) < \delta \implies d(f(x_1), f(x_2)) < \epsilon.$$

Why? Well note that $x_1 \in X$ means that it is contained in one of U_{p_i}. Furthermore,

$$d(x_2, p_i) \leq d(x_2, x_1) + d(x_1, p_i) \leq \delta + \delta_{p_i}/2 \leq \delta_{p_i}.$$

Thus, $d(f(x_1), f(x_2)) \leq d(f(x_1), f(p_i)) + d(f(p_i), f(x_2)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. $\square$

We'll briefly mention connectedness.

Proposition 17.9. *Let $f : X \to Y$ be a continuous map of metric spaces. Let $E \subseteq X$ be connected. Then $f(E)$ is connected.*

The intermediate value theorem from calculus is a corollary.

Proposition 17.10. *Let $f : [a, b] \to \mathbb{R}$ be continuous. If $f(a) < f(b)$, and $f(a) < y < f(b)$, then there exist $x \in (a, b)$ such that $f(x) = y$.*

18. 11/8/23

Here's a summary of the most important points of the last few classes. Let $f : X \to Y$ be two metric spaces. Then f is continuous if for every open $U \subseteq Y$, $f^{-1}(U)$ is open in X. Equivalently, if for every closed subset $V \subseteq Y$, then $f^{-1}(V)$ is closed.

We showed before that if f is continuous, then for any compact $U \subseteq X$, we showed $f(U)$ is compact in Y. Furthermore, if U is connected, then $f(U)$ is connected. Why is connectedness preserved by f? This is a more general statement that implies the intermediate value theorem. More precisely, let $f : [a, b] \to \mathbb{R}$ be continuous. Note that $[a, b]$ is connected. Then note that $f(a), f(b) \in f([a, b])$. By connectedness, $[f(a), f(b)] \in f([a, b])$. Thus, for every $c \in [f(a), f(b)]$, there exists $x_0 \in [a, b]$ such that $f(x_0) = c$.

Let $f : X \to Y$. Recall the definition of uniformly continuous. It means that for every $\epsilon > 0$, there exist $\eta > 0$ such that for all x, y,

$$d(x, y) < \eta \implies d(f(x), f(y)) < \epsilon.$$

Here is an example of a continuous function that is not uniformly continuous. Consider $f : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^2$. Assume for the sake of contradiction that it was uniformly continuous. Fix ϵ. Then we get some η such that $|x - y| < \eta < 1$ implies $|x^2 - y^2| < \epsilon$. Note $|x^2 - y^2| = |x - y||x + y|$. Let $A < x$, so that $A < x + y$, so we have $\eta A < \epsilon$. So pick $A \geq \epsilon/\eta$. So for x such that $x \geq \epsilon/\eta$, we see a violation of uniform continuity.

Let $f : (a, b) \to \mathbb{R}$. Let $c \in (a, b)$, we define the left limit

$$\lim_{x \to c_-} := \lim_{x \to c, x < c} f(x) = d = f(c_-)$$

if for every $\epsilon > 0$, there exist η such that for all $c - \eta < x < c$, $|f(x) - d| < \epsilon$. Similarly, we define the right limit to be

$$\lim_{x \to c_+} := \lim_{x \to c, x > c} f(x) = d = f(c_+)$$

if for every $\epsilon > 0$, there exist η such that for all $c < x < c + \eta$, we have $|f(x) - d| < \epsilon$. Note that given some arbitrary function, there's no guarantee that the left or right limits exist, nor is there guarantee that if they simultaneously exist then the left and right limits agree.

If either the limit on the left or the limit on the right do not exist, then we say c is a discontinuity of the second kind. If they both exist, but do not agree, then we say c is a discontinuity of the first kind.

Definition 18.1. Let $f : [a, b] \to \mathbb{R}$. Then f is monotonic if either

- f is increasing, so $x \leq y \implies f(x) \leq f(y)$.
- f is decreasing, so $x \leq y \implies f(x) \geq f(y)$

The following shoes that monotonic functions can't be too crazy. They may not be continuous, but any discontinuities they have must be discontinuities of the first kind, i.e that left and right limits exist but there's no guarantee they'll agree.

Theorem 18.2. *Let I denote any interval ($[a,b],(a,b],(-\infty,a]$, etc). If $f : I \to \mathbb{R}$ is monotonic, then for any $x \in I$, $f(x_+)$ and $f(x_-)$ exist.*

Proof. Suppose that f is increasing. We claim that

$$\sup_{t<x} f(t) = f(x_-).$$

Why? Well, we know that $\sup_{t<x} f(t)$ exists because since f is monotonically increasing and bounded above. Fix some $\epsilon > 0$. Then since $f(x_-)$ is the supremum, we can find η such that $x - \eta < t < x$ implies $0 \le f(x_-) - f(t) < \epsilon$.

Similarly, we claim that

$$\inf_{t>x} f(t) = f(x_+).$$

Why? The infimum exists since f is increasing and bounded below (remember, our limit goes in the negative direction). Because this is an infimum, for any $\epsilon > 0$, we can find η such that $x < t < x + \eta$ implies $0 \le f(t) - f(x_+) < \epsilon$.

If f is decreasing, a similar analysis holds, where

$$\inf_{t<x} f(t) = f(x_-), \sup_{t>x} f(t) = f(x_+).$$

□

Assume that $I = [a, +\infty)$. We say that $\lim_{x\to+\infty} f(x) = b \in \mathbb{R}$ if for every $\epsilon > 0$, there exist $A > 0$ such that for all $x \ge A$, we have $|f(x) - b| < \epsilon$. If $I = (-\infty, b]$, we say $\lim_{x\to-\infty} f(x) = d \in \mathbb{R}$ if for every $\epsilon > 0$, there exists $A > 0$ such that for all $x \le -A$, we have $|f(x) - d| < \epsilon$.

If $b = +\infty$, then for $I \subseteq \mathbb{R} \cup \{\infty, -\infty\}$, then for any $a \in I$ (could be $\pm\infty$), we say

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

if for every $A > 0$, there exist a neighborhood $N(a)$ of a such that for every $x \in N(a)$ not equal to a, we have $f(x)\ geqA$.

To be clear, what is a neighborhood of ∞? It is anything that looks like $[c, +\infty)$. Similarly, a neighborhood of $-\infty$ is $(-\infty, c]$.

Altogether, we can succinctly write all of this as:

Definition 18.3. Let $\overline{R} = \mathbb{R} \cup \{\pm\infty\}$. Let $a, b \in \overline{R}$. Then

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

if for every neighborhood $N(b)$ around b, there exist a neighborhood $N(a)$ of a such that for every $x \in N(a)$ where $x \ne a$, $f(x) \in N(b)$.

Note that f is montonic, and the interval considered includes $+\infty$, then $\lim_{x\to\infty} f$ exists. Remark: for finite limits, all the usual operations are valid (sum, multiplication, etc.) But this is not true for infinite limits.

Example 18.4. $\lim_{x\to 0} \frac{1}{x^2} = +\infty$.

$\lim_{x\to 0} \frac{1}{x}$ does not exist. The right limit is $+\infty$, the left limit is $-\infty$. So the right and left limits do not agree.

$\lim_{x\to+\infty} \sqrt{x+1} - \sqrt{x} = \lim_{x\to+\infty} \frac{1}{\sqrt{x+1}+\sqrt{x}} = 0$.

$\lim_{x\to 1} \frac{x^3}{x-1}$

19. 11/13/23: DIFFERENTIATION

Today we'll talk about differentiation. Let $f : I \to \mathbb{R}$ where $I = [a, b]$ or (a, b).

Definition 19.1. Let x be an interior point of I. If the limit

$$\lim_{t \to x} \frac{f(t) - f(x)}{t - x}$$

exists, then we say that f is differentiable at x, and the derivative $f'(x)$ is the limit. So

$$f'(x) := \lim_{t \to x, t \neq x} \frac{f(t) - f(x)}{t - x}.$$

We say f is differentiable on the interior of I if it is differentiable at every interior point of I. Its derivative is the function

$$f' : int(I) \to \mathbb{R}, \text{ where } x \mapsto f'(x).$$

Given a differentiable function f, what does f' actually mean? Take a look at the expression $\frac{f(t)-f(x)}{t-x}$. Starting from $(x, f(x))$, this fraction measures the ratio of change to $(t, f(t))$. In other words, it is the slope of the line connecting $(x, f(x))$ and $(t, f(t))$. Then as $t \to x$, the limit of these lines is the line tangent to the graph of f at $(x, f(x))$. The derivative $f'(x)$ is thus the slope of this tangent line to the graph of f at $(x, f(x))$. In other words, the derivative $f'(x)$ is a linear approximation of f at x – it measures the infinitesimal rate of change of f at x. In other words, if we perturb x by a very small quantity ϵ, we would expect that $f(x + \epsilon)$ would be approximately $f(x) + \epsilon f'(x)$.

Remark: $I = [a, b]$, there is a notion of a right derivative at $a \in I$, and left derivative at $x = b$.

Example 19.2. Consider the function $| - | : \mathbb{R} \to \mathbb{R}$, where $x \mapsto |x|$. The function is differentiable everywhere except at $x = 0$. Consider

$$\lim_{t \to 0} \frac{|t| - |0|}{t - 0} = \lim_{t \to 0} \frac{|t|}{t}.$$

But this limit does not exist! Because if $t \to 0$ but $t < 0$ always, then the limit is -1. If $t \to 0$ but $t > 0$ always, then the limit is 1. But if the limit existed, it wouldn't matter what sequence of points we chose to approach 0 with. But here, we see that it does matter. Thus, the limit does not exist, and f is not differentiable.

The graph of this function looks like a "V". Differentation is somehow capturing the notion of "smoothness." The graph of the absolute value function is pointed at $x = 0$ and not "smooth" there.

Theorem 19.3. *Let $f : I \to \mathbb{R}$ be a function. Assume that f is differentiable at x an interior point of I. Then f is continuous at x.*

Proof. To show that f is continuous at x, it suffices to show that

$$\lim_{t \to x} f(t) - f(x) = 0.$$

Let $t \in int(I) \setminus \{x\}$. Write

$$f(t) - f(x) = \frac{f(t) - f(x)}{t - x}(t - x).$$

Taking the limit as $t \to x$, we have

$$\lim_{t\to x} f(t)-f(x) = \lim_{t\to x} \frac{f(t)-f(x)}{t-x}(t-x) = \lim_{t\to x} \frac{f(t)-f(x)}{t-x} \lim_{t\to x}(t-x) = f'(x)\lim_{t\to x}(t-x) = 0.$$

Thus, f is continuous at x. □

This shows that differentiability is a stronger notion than continuous. Colloquially: if the graph of f is smooth (dfferentiable), surely you can also draw it without picking up your pen (continuous).

Proposition 19.4. *Let $f, g : I \to \mathbb{R}$, x an interior point of I. If f, g are both differentiable at x, then*

- *(sums) $f+g$ is differentiable at x and*
$$(f+g)'(x) = f'(x) + g'(x).$$
- *(products) fg is differentiable at x and*
$$(fg)'(x) = f(x)g'(x) + f'(x)g(x).$$
- *(quotients) If $g(x) \neq 0$, since g is continuous at x, it is nonzero on an open neighborhood of x. Thus, it makes sense to talk about the derivative of $\frac{f}{g}$ at x.*
  We have $\frac{f}{g}$ is differentiable at x, and
$$(\frac{f}{g})'(x) = \frac{g(x)f'(x) - f(x)g'(x)}{(g(x))^2}.$$

Proof.
- (sums) We have that
$$\lim_{t\to x} \frac{(f+g)(t)-(f+g)(x)}{t-x} = \lim_{t\to x} \frac{f(t)+g(t)-f(x)-g(x)}{t-x}$$
$$= \lim_{t\to x} \frac{f(t)-f(x)}{t-x} + \lim_{t\to x} \frac{g(t)-g(x)}{t-x} = f'(x)+g'(x).$$
- (products) We have that
$$\lim_{t\to x} \frac{f(t)g(t)-f(x)g(x)}{t-x} = \lim_{t\to x} f(t)\frac{g(t)-g(x)}{t-x} + g(x)\frac{f(t)-f(x)}{t-x}$$
$$= \lim_{t\to x} f(t)\frac{g(t)-g(x)}{t-x} + \lim_{t\to x} g(x)\frac{f(t)-f(x)}{t-x}$$
$$= f(x)g'(x) + g(x)f'(x).$$
- (quotients) Observe that $\frac{1}{g}$ is differentiable at x with derivative
$$\lim_{t\to x} \frac{\frac{1}{g(t)} - \frac{1}{g(x)}}{t-x} = \lim_{t\to x} \frac{g(t)-g(x)}{t-x}\frac{1}{g(t)g(x)} = g'(x)\frac{1}{g(x)^2}.$$
  Now apply the product rule for $\frac{f}{g} = f\frac{1}{g}$, and you obtain the quotient rule.

□

Let's talk about some interesting differentiable functions: $\cos(x)$ and $\sin(x)$. The usual definition you may be used to of these functions are that $(\cos\theta, \sin\theta)$ parameterize the unit circle for $\theta \in [0, frm{-}epi]$. There's another way to define $\cos\theta$ and $\sin\theta$. We have that

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

Then $e^{i(-\theta)} = \cos(-\theta) + i\sin(-\theta) = -\cos\theta + i\sin\theta$. This implies that

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2}, \sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

In fact, in the complex plane, e^z admits a power series which converges uniformly on compact sets

$$e^z = \sum_{n\geq 0} \frac{z^n}{n!}.$$

Then one can express

$$\cos(\theta) = \frac{1}{2}(\sum_{n\geq 0} \frac{|i\theta|^n}{n!} + \sum_{n\geq 0} \frac{(-i\theta)^n}{n!}),$$

and

$$\sin(\theta) = \frac{1}{2i}(\sum_{n\geq 0} \frac{|i\theta|^n}{n!} - \sum_{n\geq 0} \frac{(-i\theta)^n}{n!}).$$

Here's an exercise.

Example 19.5. Using the fact that $\exp : \mathbb{R} \to \mathbb{R}$ is differentiable, whose derivative is itself, with inverse $\log(x)$ on $(0, \infty)$ whose derivative is $\frac{1}{x}$, prove that

$$(\sin)'(x) = \cos x, (\cos)'(x) = -\sin x, \forall x \in \mathbb{R}.$$

Theorem 19.6 (Chain rule)**.** *Let $f : (a, b) \to (c, d)$ and $g : (c, d) \to \mathbb{R}$. Assume that f is differentiable at x, and g is differentiable at $g(x)$.*

Proof. Let's write out f and g in linear approximation form. Since f is differentiable at x, then if we let $\phi(t) = \frac{f(t)-f(x)}{t-x} - f'(x)$, we have that

$$f(t) = f(x) + (t - x)f'(x) + (t - x)\phi(t)$$

where $\lim_{t\to x} \phi(t) = 0$. Since g is differentiable at $f(x)$, we can write

$$g(\mu) = g(f(x)) + (\mu - f(x))g'(f(x)) + (\mu - f(x))\Gamma(t)$$

where $\Gamma(t) = \frac{g(t)-g(f(x))}{t-f(x)}$, where $\lim_{t\to f(x)} \Gamma(t) = 0$.

Note that from the definition of differentiability we obtained these linear approximation equations. But it is reversible: given h such that

$$h(t) = h(x) + (t - x)\beta + (t - x)\epsilon(t)$$

where $\lim_{t\to x} \epsilon(t) = 0$, this implies that

$$\lim_{t\to x} \frac{h(t) - h(x)}{t - x} = \beta$$

which implies that h is differentiable at x with derivative $h'(x) = \beta$. We'll use this to finally prove the chain rule.

Substituting $\mu = f(t)$ in the linear approximation of g, we obtain

$$g \circ f(t) = g \circ f(x) + (f(t) - f(x))g'(f(x)) + (f(t) - f(x))\Gamma(f(t))$$
$$= g \circ f(x) + (t - x)\frac{f(t) - f(x)}{t - x}g'(f(x)) + (t - x)\frac{f(t) - f(x)}{t - x}\Gamma(f(t)).$$

Then

$$(g \circ f)'(x) = \lim_{t\to x} \frac{g \circ f(t) - g \circ f(x)}{t - x}$$

$$= \lim_{t\to x} \frac{f(t)-f(x)}{t-x} g'(f(x)) + \lim_{t\to x} \frac{f(t)-f(x)}{t-x} \Gamma(f(t))$$
$$= f'(x)g'(f(x)) + f'(x) * 0 = f'(x)g'(f(x)).$$

□

Definition 19.7. Let $f : (X, d) \to \mathbb{R}$ be continuous. Let $x \in X$. Then x is said to be

(1) a local minimum if there exists a neighborhood V of x such that for every $y \in V$, $f(y) \geq f(x)$
(2) a local maximum if there exists a neighborhood V of x such that for every $y \in V$, $f(y) \leq f(x)$
(3) x is a local extremal value if either x is a local minima or local maxima.

Theorem 19.8. *Let $f : I \to \mathbb{R}$, $x \in I$ an interior point. If x is a local extremal value and f is differentiable at x, then*

$$f'(x) = 0.$$

Proof. Assume that x is a local maximum. Then for $t > x$ in some interval $(x, x+\epsilon)$, we have $f(t) \leq f(x)$. Thus

$$\lim_{t\to x} \frac{f(t)-f(x)}{t-x} \leq 0 \implies f'(x) \leq 0.$$

Furthermore, for $t < x$ in some interval $(x-\epsilon, x)$ we have $f(t) \leq f(x)$ so

$$\lim_{t\to x} \frac{f(t)-f(x)}{t-x} \geq 0 \implies f'(x) \geq 0.$$

Thus, $f'(x) = 0$. A similar proof holds for when x is a local minimum. □

20. 11/15/23: MEAN VALUE THEOREM, TAYLOR SERIES

To clarify: when we say $f : [a, b] \to \mathbb{R}$ is continuous, we mean it is continuous on (a, b), and continuous at a on the right, and on the left on b.

Theorem 20.1 (Mean value theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ differentiable on (a, b) such that $f(a) = f(b)$. Then there exist $x \in (a, b)$ such that $f'(x) = 0$.*

Proof. If f is constant, then $f(x) = f(a) = f(b), \forall x \in (a, b)$. Then $f'(x) = 0 \forall x \in [a, b]$.

If not constant, there exist $x_0 \in (a, b)$ such that $f(x_0) \neq f(a)$. WLOG we can assume that $f(x_0) > f(a)$ (the case where $f(x_0) < f(a)$ follows similarly). Since $f : [a, b] \to \mathbb{R}$ is continuous, and $[a, b]$ is compact, then $f([a, b])$ is compact and thus closed and bounded by Heine-Borel, so

$$\sup_{x\in[a,b]} f(x) = \max_{x\in[a,b]} f(x),$$

because $\sup_{x\in[a,b]} f(x)$ is a limit point of $f([a, b])$, which is closed, so it must be achieved. So let the supremum/maximum be $f(x_1)$ for some $x_1 \in [a, b]$. Thus, we have

$$f(x_1) \geq f(x_0) > f(a) = f(b).$$

Hence $x_1 \in (a, b)$. Our discussion implies x_1 is an extremal value for f, thus $f'(x_1) = 0$. □

This is equivalent to the 2nd version of the mean value theorem, which at first glance, seems more general. But really these are the same theorems, once you "rotate" (not true rotation, you just subtract a linear function) the graph in the 2nd version.

Theorem 20.2 (Mean value theorem 2nd version)**.** *Let $f : [a, b] \to \mathbb{R}$ continuous, and differentiable on (a, b). Then there exists $x \in (a, b)$ such that*

$$f'(x) = \frac{f(b) - f(a)}{b - a}.$$

Proof. Let $g(x) = (x - a)\frac{f(b)-f(a)}{b-a} + f(a)$. So $g(a) = f(a)$, and $g(b) = f(b)$. In other words, this $g(x)$ is the line through $(a, f(a))$ and $(b, f(b))$.

Then let $h = f - g$. Then $h(a) = h(b) = 0$. Then by the first version of MVT 20.1, there exist $x_0 \in (a, b)$ such that $h'(x_0) = 0$. But this means $f'(x_0) - g'(x_0) = 0$, but $g'(x_0) = \frac{f(b)-f(a)}{b-a}$. So

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$

□

So these two statements of the mean value theorem are equivalent.

Theorem 20.3. *Let $f : [a, b] \to \mathbb{R}$ continuous, and differentiable on (a, b). Then*

(1) If $f'(x) \geq (>) 0$, $\forall x \in [a, b]$, then f is nondecreasing (increasing).
(2) If $f'(x) \leq (<) 0$, $\forall x \in [a, b]$, then f is non-increasing (decreasing).

Proof. For any $x, y \in [a, b]$ where $x < y$, there exist $z \in (a, b)$ such that

$$f'(z) = \frac{f(y) - f(x)}{y - x}.$$

This implies all four of the statements, by examining what the implications are if the derivative is always positive, always nonnegative, always non-positive, or always negative. □

We'll briefly mention L'Hospital rule. Suppose $f, g : [a, b] \to \mathbb{R}$, $t \in (a, b)$. Consider

$$\lim_{x \to t} \frac{f(x)}{g(x)}.$$

If $\lim_{x\to t} g(x) = g(t) \neq 0$, then we're all good! But if $\lim_{x\to t} f(x) = \lim_{x\to t} g(x) = 0$, then we can't conclude anything!

Theorem 20.4 (L'Hospital rule)**.** *Let $f, g : [a, b] \to \mathbb{R}$ differentiable, $t \in (a, b)$, such that $g'(t) \neq 0$, and*

$$\lim_{x\to t} f(x) = 0, \lim_{x\to t} g(x) = 0, \lim_{x\to t} \frac{f'(x)}{g'(x)} = A.$$

Then

$$\lim_{x\to t} \frac{f(x)}{g(x)} = A.$$

Proof. Proof in 5.13 Rudin. □

Example 20.5. $\lim_{x\to 0} \frac{1-\cos x}{x^2} = \lim_{x\to 0} \frac{\sin x}{2x} = \lim_{x\to 0} \frac{\cos x}{2} = \frac{1}{2}$ by L'Hospital rule.

Definition 20.6 (Higher order derivatives)**.** Let $f : I \to \mathbb{R}$, I open interval. We write $f^{(k)} : I \to \mathbb{R}$ for the derivative of order k, where by convention, we let $f^{(0)} := f$ itself, and

$$(f^{(n)}) = (f^{(n-1)})'.$$

So for $f^{(n)}$ to exist, note that $f, f', f^{(2)}, \cdots, f^{(n-1)}$ must all be differentiable.

For $n \geq 0$, we denote

$$\mathcal{C}^n(I, \mathbb{R}) = \{f : I \to \mathbb{R} | f^{(n)} \text{ exists and is continuous } \}.$$

We define

$$C^\infty(I, \mathbb{R}) = \bigcap \mathcal{C}^n(I, \mathbb{R}) = \{f : I \to \mathbb{R} | \text{ derivatives of any order exist}\}.$$

In other words, $C^\infty(I, \mathbb{R})$ is infinitely differentiable functions, the nicest "smoothest" functions.

We said before that a function differentiable at a point can be approximated at that point by a linear function. Taylor's theorem gneneralizes this: if the function is infinitely differentiable, it can be approximated by a polynomial-like function.

Theorem 20.7 (Taylor's theorem)**.** *Let $f \in C^n([a, b], \mathbb{R})$. Let $\alpha \in (a, b)$ and let*

$$P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(\alpha)}{k!}(t - \alpha)^k.$$

$$= f(\alpha) + (t - \alpha)f'(\alpha) + \frac{(t - \alpha)^2}{2} f^{(2)}(\alpha) + \cdots + \frac{f^{(n-1)}(\alpha)}{(n - 1)!}(t - \alpha)^{n-1}.$$

$P(t)$ is what our f is approximated by, and here is what the approximation says precisely: let $(\alpha, \beta) \subseteq [a, b]$. Then there exists $x \in (\alpha, \beta)$ such that

$$f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta - \alpha)^n.$$

Proof. Note that by construction, $P(\alpha) = f(\alpha), P'(\alpha) = f'(\alpha), \cdots, P^{(n-1)}(\alpha) = f^{(n-1)}(\alpha)$.

Let $M = n! \frac{f(\beta) - P(\beta)}{(\beta - \alpha)^n}$. Let

$$g(t) = f(t) - P(t) - M(t - \alpha)^n.$$

Then because P, f agree up to order $n - 1$, we have that

$$g(\alpha) = 0, g'(\alpha) = 0, \cdots, g^{(n-1)}(\alpha) = 0.$$

Note that $g(\beta) = 0$. Then applying the MVT 20.1, we have there exists $\beta_1 \in (\alpha, \beta)$ such that $g'(\beta_1) = 0$. Then since $g'(\alpha) = 0$, by MVT 20.1, can find $\beta_2 \in (\alpha, \beta_1)$ such that $g^{(2)}(\beta_2) = 0$. Continuing, we find there exist $\beta_n \in (\alpha, \beta_{n-1}$ such that

$$g^{(n)}(\beta_n) = 0 \implies f^{(n)}(\beta_n) - n!M = 0 \implies M = \frac{f^{(n)}(\beta_n)}{n!}.$$

□

In general, if you have a vector-valued function

$$f : I \to \mathbb{R}^k,$$

we can write $f(x) = (f_1(x), \cdots, f_k(x))$. Then f is differentiable at x if

$$\lim_{t \to x} \frac{f(t) - f(x)}{t - x} \text{ exists in } \mathbb{R}^k.$$

Proposition 20.8. *$f : I \to \mathbb{R}^k$ is differentiable $\iff$ each f_i is differentiable.*

21. 11/20/23: Riemann integrals

Today we'll define the Riemann integral, which is meant to capture the notion of "area under the curve." Assume $f : [a, b] \to \mathbb{R}$ be bounded.

Definition 21.1. A partition of $[a, b]$ is a set $\{x_0, \cdots, x_n\} \subseteq [a, b]$ such that

$$a = x_0 \le x_1 \le \cdots \le x_n = b.$$

Define $\Delta x_i = x_i - x_{i-1}$ for $i = 1, \cdots, n$.

Let

$$M_i := \sup\{f(x)|x_{i-1} \le x \le x_i\}, m_i := \inf\{f(x)|x_{i-1} \le x \le x_i\},$$

and define

$$U(P, f) = \sum_{i=1}^{n} M_i \Delta x_i, L(P, f) = \sum_{i=1}^{n} m_i \Delta x_i.$$

The quantities $U(P, f)$ and $L(P, f)$ are upper and lower bounds, respectively, for the area under the curve. As we take finer partitions, these quantities will squeeze towards each other to give us the value of the integral.

Definition 21.2. Let $f : [a, b] \to \mathbb{R}$ be bounded. Then define

$$\overline{\int_a^b} f(x)dx = \inf\{U(P, f)| \text{ for all partitions } P \text{ of } [a, b]\}$$

and define

$$\underline{\int_a^b} f(x)dx = \sup\{L(P, f)| \text{ for all partitions } P \text{ of } [a, b]\}.$$

A bounded function $f \in \mathbb{R}^{[a,b]}$ is **Riemann integrable** if

$$\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx.$$

Why is $\int_a^b f(x)dx$ finite? Because f is bounded, so there is a global lower bound and an upper bound m, M for f over $[a, b]$. Thus, for all partitions, $U(P, f) \le M(b - a)$ and $L(P, f) \ge m(b - a)$.

Now do we need the function to be bounded? No, you can still do integrals with unbounded functions, but you have to be careful. For example, we have $\int_0^1 \frac{1}{2\sqrt{x}}dx = \sqrt{x}|_0^1$. Although, really the integral should be written as something like $\lim_{\epsilon\to 0} \int_\epsilon^1 \frac{1}{2\sqrt{x}}dx = \lim_{\epsilon\to 0} 1 - \sqrt{\epsilon} = 1$, since from the definition of Riemann integrable, $\sqrt{x}$ is bounded on $[\epsilon, 1]$, but it is not on $[0, 1]$.

This shows that our assumption for boundedness for Riemann integration is quite limited. There are ways to extend the Riemann integral as we have defined it to integrate unbounded functions. There are also even more advanced generalizations of the Riemann integral: the Lebesgue integral. See Math 114. But often in daily life, the functions you will encounter will be bounded on a given interval, so we'll still gain something substantive with our definition of Riemann integrable.

We will now prove the following theorem.

Theorem 21.3. *If (bounded) f is continuous on $[a,b]$ except possible at finitely many points, then f is integrable.*

First let's develop some technology.

Lemma 21.4. *Suppose P' is a refinement of P, i.e $P \subseteq P'$. Then*
$$U(P,f) \geq U(P',f) \text{ and } L(P,f) \leq L(P',f).$$

Proof. It suffices to prove the statement when $P' = P \cup \{x^*\}$ where, say, $x_{i-1} < x^* < x_i$. Let
$$\omega_1 = \inf\{f(x)|x_{i-1} \leq x \leq x^*\} \geq m_i$$
$$\omega_2 = \inf\{f(x)|x^* \leq x \leq x_i\} \geq m_i.$$
Then
$$L(P',f) = m_1\Delta x_1 + \cdots + \omega_1(x^* - x_{i-1}) + \omega_2(x_i - x^*) + m_{i+1}\Delta x_{i+1} + \cdots + m_n\Delta x_n.$$
$$L(P,f) = m_1\Delta x_1 + \cdots + m_i\Delta x_i + \cdots + m_n\Delta x_n.$$
Note that the terms in $L(P',f)$ and $L(P,f)$ all agree except in $L(P',f)$, the quantity $m_i\Delta x_i$ in $L(P,f)$ is replaced by $\omega_1(x^* - x_{i-1}) + \omega_2(x_i - x^*)$. But $\omega_1, \omega_2 \geq m_i$, so
$$\omega_1(x^* - x_{i-1}) + \omega_2(x_i - x^*) \geq m_i(x^* - x_{i-1} + x_i - x^*) = m_i\Delta x_i.$$
A similar proof proves the claim that $U(P,f) \geq U(P',f)$. Exercise left to the reader. □

Corollary:

Proposition 21.5. $\underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx.$

Proof. Note that for *any* partitions P_1, P_2, we have
$$L(P_1,f) \leq U(P_2,f).$$
If we prove this, then if we let $P = P_1 \cup P_2$, we have
$$L(P_1,f) \leq L(P_1 \cup P_2, f) \leq U(P_1 \cup P_2, f) \leq U(P_2,f).$$
This follows from lemma 21.4. This proves the claim that $\underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx$ by definition (unpackage the definition in terms of sup and inf of $L(P,f), U(P',f)$ over all partitions). □

Lemma 21.6. *A bounded f on $[a,b]$ is integrable $\iff$ $\forall \epsilon > 0$, there $\exists P$ partition of $[a,b]$ such that*
$$U(P,f) - L(P,f) < \epsilon.$$

Proof. Suppose f is integrable. Then
$$I = \inf_P U(P,f) = \sup_P L(P,f).$$
So there exists P_1 and P_2 such that
$$I - L(P_1,f) < \frac{\epsilon}{2}, U(P_2,f) - I < \frac{\epsilon}{2}.$$
So $U(P_2,f) - L(P_1,f) < \epsilon \implies U(P_1 \cup P_2, f) - L(P_1 \cup P_2, f) < \epsilon$.

Now suppose that for every $\epsilon > 0$, there exists a partition P of $[a,b]$ such that
$$U(P,f) - L(P,f) < \epsilon.$$

Fix such an ϵ. Since there exists such a partition, by the lemma 21.4, this implies

$$\inf_P U(P,f) - \sup_P L(P,f) \le U(P,f) - L(P,f) < \epsilon.$$

So they must be equal. ☐

Proposition 21.7. *If f is continuous on $[a,b]$, then f integrable.*

Proof. Note $[a,b]$ is compact. This implies f is uniformly continuous on $[a,b]$. We will show that for any $\epsilon > 0$, there exists partition P such that

$$U(P,f) - L(P,f) < \epsilon.$$

Note that f is uniformly continuous. Then consider $\frac{\epsilon}{b-a}$. Then uniform continuity implies there exists δ such that $|x-y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{b-a}$. Then pick partition P such that $\Delta x_i < \delta$. Then with respect to this partition P, $M_i - m_i < \frac{\epsilon}{b-a}$ for each index. Then

$$U(P,f) - L(P,f) = \sum (M_i - m_i)\Delta x_i < \sum \frac{\epsilon}{b-a}\Delta x_i = \frac{\epsilon}{b-a}(b-a) = \epsilon.$$

Thus, by lemma 21.6, f is integrable. ☐

22. 11/27/23

Theorem 22.1. *(1) If f is continuous on $[a,b]$, then f is Riemann integrable.*
(2) If $f : [a,b] \to \mathbb{R}$ is montonic and bounded, then f is integrable.
(3) If f is bounded and piecewise continuous on $[a,b]$, then f is integrable. In other words, there $\exists a_1, \cdots, a_r \in [a,b]$ such that f is continuous on $[a,b] \setminus \{a_1, \cdots, a_r\}$.

Proof. (1) Proved last time.
(2) Recall that f is integrable on $[a,b] \iff \forall \epsilon > 0$ there exists a partition P such that $U(P,f) - L(P,f) < \epsilon$. Since f is monotonic, WLOG we can assume that f is increasing. Let $x_k := a + k\frac{b-a}{n}$ for $0 <\le k \le n$. This gives a uniform partition $P = \{x_0, \cdots, x_n\}$. Then

$$U(P,f) = \sum_{k=0}^{n-1} \max_{x_k \le t \le x_{k+1}} f(t)(x_{k+1} - x_k) = \sum_{k=0}^{n-1} f(x_{k+1})\frac{b-a}{n}$$

$$L(P,f) = \sum_{k=0}^{n-1} f(x_k)\frac{b-a}{n}.$$

Thus,

$$U(P,f) - L(P,f) = \frac{b-a}{n}\sum_{k=0}^{n-1} f(x_{k+1}) - f(x_k) = \frac{b-a}{n}(f(b) - f(a)).$$

Then for every ϵ, we simply need to choose n such that

$$\frac{b-a}{n}(f(b) - f(a)) < \epsilon.$$

So given ϵ, we can find n by the archimedean property such that

$$n > \frac{(b-a)(f(b) - f(a))}{\epsilon}.$$

This gives a uniform partition P given by $x_0, \cdots, x_n$ such that $U(P,f) - L(P,f) < \epsilon$. Thus, f is integrable.

(3) Since f is bounded note $|f| \leq M$ for some M. Let $\epsilon > 0$. We will find partition P such that $U(P,f) - L(P,f) < \epsilon$. Find intervals $[\alpha_0, \beta_0], \cdots, [\alpha_k, \beta_k]$ so that each segment $[\alpha_i, \beta_i]$ contains a_i, where we're assuming that the indices are ordered so that $a_0 < \cdots < a_k$. Make the $[\alpha_i, \beta_i]$ small enough so that $\sum \beta_i - \alpha_i < \frac{\epsilon}{4M}$. Then consider $J = [a,b] \setminus \bigcup(\alpha_i, \beta_i)$. Then J is a finite union of closed intervals and f is continuous on J. So we can find a partition P of J so that

$$U_J(P,f) - L_J(P,f) < \frac{\epsilon}{2}.$$

This partition P naturally gives a partition of the entire $[a,b]$, and we see that

$$U(P,f) - L(P,f) < U_J(P,f) - L_J(P,f) + \sum_i (\max_{[\alpha_i,\beta_i]} f - \min_{[\alpha_i,\beta_i]} f)(\beta_i - \alpha_i) < \frac{\epsilon}{2} + \frac{\epsilon}{4M} 2M = \epsilon.$$

where note that $(\max_{[\alpha_i,\beta_i]} f - \min_{[\alpha_i,\beta_i]} f) < 2M$.

□

Theorem 22.2. *Let $f : [a,b] \to [m,M]$ be an integrable function, and $\Phi : [m,M] \to \mathbb{R}$ continuous. Then $\Phi \circ f$ is integrable.*

Proof. See 6.11 in Rudin. □

Applications: suppose $f : [a,b] \to \mathbb{R}$ is bounded and integrable, then f^2 is also integrable. This is because f^2 is the composition with x^2 which is continuous. Similarly, this means e^f, $\sin(f)$ are all integrable.

Theorem 22.3. *Let $f, g : [a,b] \to \mathbb{R}$ be integrable functions.*

(1) *Then $f + g$ is integrable: $\int_a^b f + g dx = \int_a^b f dx + \int_a^b g dx$.*
(2) *Then for any $c \in \mathbb{R}$, cf is integrable: $\int_a^b cf = c \int_a^b f$.*
(3) *If $f(x) \leq g(x)$ for every $x \in [a,b]$, then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$. Note this means if $g \geq 0$ everywhere, then $\int_a^b g(x)dx \geq 0$.*
(4) *For any $c \in [a,b]$, $\int_a^c f(x)dx + \int_c^b f(x)dx = \int_a^b f(x)dx$.*
(5) *f is integrable on any interval $[c,d] \subseteq [a,b]$. (This amounts to showing that for every $\epsilon > 0$, there exist P such that $U_{[c,d]}(P,f) - L_{[c,d]}(P,f) < U(P,f) - L(P,f) < \epsilon$).*
(6) *If $|f(x)| \leq M$ for every $x \in [a,b]$, then*

$$\left| \int_a^b f(x)dx \right| \leq M(b-a).$$

(7) *The product fg is also integrable. This comes from noting that $fg = \frac{(f+g)^2 - f^2 - g^2}{2}$.*

Proof. (1) Let $\epsilon > 0$. There exist partitions P_1, P_2 such that $U(P_1,f) - L(P_1,f) < \frac{\epsilon}{2}$ and $U(P_2,f) - L(P_2,f) < \frac{\epsilon}{2}$. Let P be a refinement of P_1 and P_2. Then $L(P,f) + L(P,g) \leq L(P, f+g)$. This comes from the fact that generally $\min_I f + \min_I g \leq \min_I f + g$. Note also that $U(P, f+g) \leq U(P,f) + U(P,g)$, since generally over an interval we have $\max_I f + g \leq \max_I f + \max_I g$. Then

$$L(P,f) + L(P,g) \leq L(P,f+g) \leq U(P,f+g) \leq U(P,f) + U(P,g).$$

We can assume that $U(P,f) \leq \epsilon + \int_a^b f$ and $U(P,g) \leq \epsilon + \int_a^b f$. Then $\int_a^b f + g \leq U(P, f+g) \leq 2\epsilon + \int f + \int g$. Since this is true for every ϵ, this implies that $\int_a^b f + g \leq \int_a^b f + \int_a^b g$. The analagous argument for $L(P,f)$ shows the reverse inequality that $\int_a^b f + \int_a^b g \leq \int_a^b f + g$.

(2)

□

23. 11/29/23

Theorem 23.1 (Change of variables)**.** *Let $\phi : [A,B] \to [a,b]$ be differentiable, and ϕ' continuous. Furthermore, suppose ϕ is bijective and strictly increasing. Let $f : [a,b] \to \mathbb{R}$ be integrable. Then*

$$\int_a^b f(x)dx = \int_A^B f \circ \phi(t)\phi'(t)dt.$$

Due to time we won't prove this. But we note that a very important part of this statement is that ϕ is bijective and strictly increasing. This necessarily means $\phi(A) = a, \phi(B) = b$. In real life, you may encounter some $\phi : [A,B] \to \mathbb{R}$ and if $\phi'(t) \neq 0$ for every $t \in [A,B]$, then we must have either $\phi'(t) > 0$ for all $t \in [A,B]$ or $\phi'(t) < 0$ for all $t \in [A,B]$ by the intermediate value theorem. But if, say, $\phi'(t) > 0$ for every t, then $\phi : [A,B] \to [\phi(A), \phi(B)]$ is a bijection and strictly increasing.

Now what if you did change of variables where for every t, $\phi'(t) < 0$? Consider $[A,B] = [0,1]$. Then consider $\phi(1-t)$. Then $\phi(1-t)$ has derivative strictly positive everywhere. So

$$\int_a^b f(x)dx = \int_A^B f \circ \phi(1-t)(-\phi'(1-t))dt = \int_B^A f \circ \phi(1-t)\phi'(1-t)dt = -\int_A^B f \circ \phi(t)\phi'(t)dt.$$

In general, if $\phi'(t) \neq 0$ for every t, then

$$\int_a^b f(x)dx = \int_A^B f \circ \phi(t)|\phi'(t)|dt.$$

Let's discuss the relationship between integration and differentiation.

Theorem 23.2. *Let $f : [a,b] \to \mathbb{R}$ be integrable. Let*

$$F(x) = \int_a^x f(t)dt.$$

Then F is continuous. If f is continuous at x_0, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$.

Proof. Let $M > 0$ such that $|f(x)| \leq M$ for all $x \in [a,b]$. For any $x, y \in [a,b]$, we have

$$|F(x) - F(y)| = |\int_y^x f(t)dt| \leq M|x-y|.$$

This shows that F is continuous. Why? Becuase if we fix an $\epsilon > 0$, then for $\eta = \frac{\epsilon}{M}$, we have that $|x-y| < \eta \implies |F(x) - F(y)| < \epsilon$, thus F is actually uniformly continuous.

Now we show the second part. Suppose that f is continuous at x_0. Let $\epsilon > 0, \eta > 0$ such that $|x - x_0| < \eta \implies |f(x) - f(x_0)| < \epsilon$. We have for $x \neq x_0$

$$\frac{F(x) - F(x_0)}{x - x_0} - f(x_0) = \frac{1}{x - x_0}(\int_{x_0}^x f(t) - f(x_0)dt).$$

Our goal is to show that $\lim_{x\to x_0} \frac{F(x)-F(x_0)}{x-x_0} - f(x_0) = 0$. Because this would imply $F'(x_0) = f(x_0)$.

If $|x - x_0| < \eta$, then for any $t \in [x_0, x]$, we have $|t - x_0| \le |x - x_0| < \eta$. Hence

$$|\int_{x_0}^{x} f(t) - f(x_0)dt| \le \int_{x_0}^{x} |f(t) - f(x_0)|dt < \epsilon|x - x_0|.$$

Thus,

$$|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)| = |\frac{1}{x - x_0}\int_{x_0}^{x} f(t) - f(x_0)dt| \le \frac{1}{x - x_0}\int_{x_0}^{x} |f(t) - f(x_0)| < \epsilon.$$

Hence $\lim_{x\to x_0} \frac{F(x)-F(x_0)}{x-x_0} = f(x_0)$. □

Remark: Here is an example of a non Lipschitz continuous function. Take $\sqrt{}:[0,1] \to \mathbb{R}$. To see that it is not lipschitz, here's a hint: $x - y = (\sqrt{x} - \sqrt{y})(\sqrt{x} + \sqrt{y})$.

Theorem 23.3 (Fundamental Theorem of Calculus)**.** *If $f : [a, b] \to \mathbb{R}$ integrable, and there exists $F : [a, b] \to \mathbb{R}$ differentiable such that $F' = f$, then*

$$\int_a^b f(x)dx = F(b) - F(a).$$

F is called the primitive of f.

Proof. Let $\epsilon > 0$. There exists $P = \{x_0, \cdots, x_n\}$, where $x_0 = a, x_n = b$ such that $U(P, f) - L(P, f) < \epsilon$. By the mean value theorem,

$$F(x_i) - F(x_{i-1}) = f(t_i)(x_i - x_{i-1})$$

for $t_i \in [x_{i-1}, x_i]$. Then

$$\sum_{i=1}^{n} f(t_i)(x_i - x_{i-1}) = F(b) - F(a),$$

and note that $L(P, f) \le \sum_{i=1}^{n} f(t_i)\Delta x_i \le U(P, f)$ implies both that

$$L(P, f) \le F(b) - F(a) \le U(P, f) \text{ and } L(P, f) \le \int_a^b f(t)dt \le U(P, f).$$

This implies that

$$|F(b) - F(a) - \int_a^b f(t)dt| < \epsilon, \forall\epsilon > 0$$

which means $F(b) - F(a) = \int_a^b f(t)dt$. □

Theorem 23.4 (Integration by parts)**.** *Let $F, G : [a, b] \to \mathbb{R}$ such that $f = F', g = G'$ are integrable. Then*

$$\int_a^b F(t)g(t)dt = F(b)G(b) - F(a)G(a) - \int_a^b f(t)G(t)dt.$$

Proof. By the fundamamental theorem of calculus, we have $[FG]' = Fg + fG$. Then $FG(b) - FG(a) = F(b)G(b) - F(a)G(a) = \int_a^b Fg + fG$. □

Example 23.5. If $f(x) = x^n$, then $F(x) = \frac{1}{n+1}x^{n+1} + C$.

If $f(x) = \cos(x)$, then $F(x) = \sin(x) + C$.

If $f(x) = \sin(x)$, then $F(x) = -\cos(x) + C$.

If $f(x)e^x$, then $F(x) = e^x + C$.

If $f(x) = \frac{1}{x}$, then $F(x) = \log(x) + C$.

If $f(x) = \frac{1}{1-x^2}$ on some interval $x \neq 1, -1$, then $F(x) = \frac{1}{2}\log(1+x) - \frac{1}{2}\log(1-x) + C$.

If $f(x) = \frac{1}{1+x^2}$, then $F(x) = \arctan(x) + C$. One computes this knowing $\tan(x) : (-\pi/2, \pi/2) \to \mathbb{R}$, $\arctan(x) = \tan^{-1}(x) : \mathbb{R} \to (-\pi/2, \pi/2)$, and using the inverse function theorem that $(f^{-1}(x))' = \frac{1}{f' f^{-1}(x)}$.

How about $f(x) = \frac{1}{\sqrt{1-x^2}}$ for $x \in [-1, 1]$? Note $\sin : [-\pi/2, \pi/2] \to [-1, 1]$. The inverse is $\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$. Note that if we consider

$$\int_{-1}^{x} \frac{1}{\sqrt{1-t^2}} dt = \int_{-\pi/2}^{\arcsin x} \frac{1}{\cos u} \cos u du = \arcsin(x) + \frac{\pi}{2},$$

where we performed a change of variables with $t = \sin u$ so $dt = \cos(u)dt$.

24. 12/4/23: SEQUENCES AND SERIES OF FUNCTIONS

Today is the last class. We will sketch some ideas that we would have encountered if we had another semester. In a way, it is a culmination of everything that we have seen this semester. Let's talk about sequences and series of functions.

Let (X, d) be a metric space, and $F(X, \mathbb{R})$ the space of real-valued functions on X. We denote $\mathcal{J}(X, \mathbb{R}) \subset C^0(X, \mathbb{R})$ be a space of "nice" special functions. This not a precise definition. The exact definition of this subspace depends on the context. But in general we will, loosely speaking, want to require that every $f \in F(X, \mathbb{R})$ to be approachable by functions in $\mathcal{J}(X, \mathbb{R})$.

Let's consider an example. Suppose we had some continuous function on $[0, 1]$. Then we could try to approximate this function by increasingly higher degrees of polynomials.

Another example is if we consider $X = \mathbb{R}$ and 2π-periodic functions. Then the goal would be to be able to approximate this 2π-periodic function by $\sin(x)$ and $\cos(x)$. Let's make this notion of approximation more precise.

Let (X, d) be a metric space. Let $\{(f_n)\}_n$ be a sequence of functions. Let $f : X \to \mathbb{R}$ be a function. Note that for every $x \in X$, $(f_n(x))_n$ is a sequence of real numbers.

If $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in X$, then we say that $(f_n)_n$ *converges pointwise* to f.

Note that $\{\sum_{k=1}^{n} f_k\}_n$ is called a series of functions and the series is denoted as $\sum_{n\geq 0} f_n$. Then we say $\sum_{n\geq 0} f_n$ pointwise converges if for every $x \in X$, the series $\sum_{n\geq 0} f_n(x)$ is convergent.

Question: what properties are preserved by pointwise convergence? If f_n is continuous for every n (written as $f_n \in C^0$) and f_n converges pointwise to f, is $f \in C^0$?

The answer is no. For example, sin and cos are continuous functions, but there are linear combinations of powers of sin and cos which pointwise converge to functions with discontinuities. Here's another example. If we let $X = [0, 1]$ we can consider $f_n(x) = x^n$ on this interval. Then f_n pointwise converges to f such that $f(x) = 0$ for $0 \leq x < 1$ and $f(1) = 1$.

So pointwise convergence doesn't even preserve continuity. If we want to preserve something like continuity, we need a stronger notion of convergence of functions.

Definition 24.1 (uniform convergence)**.** We say that $(f_n)_n$ converges uniformly to f if, not only does $f_n \to f$ pointwise, but for every $\epsilon > 0$, there exists N_ϵ such that for every $n \geq N_\epsilon$ and every $x \in X$, we have

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

Theorem 24.2 (Cauchy criterion)**.** *Let $\sum_{n\geq 0} f_n$ be a series of functions. Then it converges unformly $\iff$ for every $\epsilon > 0$, there exists $N \geq 1$ such that for every $n, m \geq N$, for every $x \in X$ we have*

$$|\sum_{k=n}^{m} f_k(x)| < \epsilon.$$

Proof. Proof sketch: For every $x \in X$, we have $\sum_{n\geq 0} f_n(x)$ is Cauchy so it converges so that $f(x) = \sum_{n\geq 0} f_n(x)$. For every $\epsilon > 0$, there exists N such that for every $n \geq N$, and letting $m \to \infty$, we have for every $x \in X$

$$|\sum_{k=0}^{n-1} f_k(x) - f(x)| = |\sum_{n}^{\infty} f_k(x)| \leq \epsilon.$$

□

Definition 24.3. $\sum_{n\geq 0} f_n$ converges pointwise absolutely if for every $x \in X$, $\sum_{n\geq 0} |f_n(x)|$ is a convergent series.

Definition 24.4 (Normal convergence/Weierstrass notion of convergence)**.** The series $\sum_{n\geq 0} f_n$ converges normally if $\sum_{n\geq 0} \sup_{x\in X} |f_n(x)|$ is convergent.

Note that in the notion of normal convergence, we are assuming that for every n, f_n is bounded, so that $|f_n(x)| \leq M_n$ for every $x \in X$. Denote $||f_n||_\infty = \sup_{x\in X} |f_n(x)| \in \mathbb{R}$. note that from the definition of normal convergence, this impies $\sum ||f_n||_\infty$ is convergent.

Example 24.5. If there exists C_n such that for every $x \in X$, $|f_n(x)| \leq C_n$ and $\sum C_n$ converges, then $\sum_n f_n$ converges normally.

Theorem 24.6. *If $\sum_n f_n$ converges normally. Then*

(1) $\sum f_n$ converges uniformly
(2) $\sum f_n$ converges pointwise absolutely

Proof. For (2): for every $x \in X$, we have

$$\sum_{n\geq 0} |f_n(x)| \leq \sum_{n\geq 0} ||f_n||_\infty$$

and since the right hand side is convergent by assumption, then the left hand side is also convergent.

For (1): we have

$$|\sum_{k=n}^{m} f_k(x)| \leq \sum_{k=n}^{m} |f_k(x)| \leq \sum_{k=n}^{m} ||f_k||_\infty < \epsilon$$

so we conclude by Cauchy. □

Summary: let $(f_n)_n$ be a sequence of functions. If $f_n \to f$ uniformly, then $f_n \to f$ pointwise.

If $\sum_{n\geq 0} f_n$ is a series of functions, then we have a notion of: pointwise convergence, pointwise absolute convergence, uniform convergence, and normal convergence. Note normal convergence implies both uniform convergence and pointwise absolute convergence. Both uniform convergence and pointwise absolute convergence implies pointwise convergence.

Theorem 24.7. *Assume that $f_n \to f$ converges uniformly on $E \subseteq X$. Let $x \in E$ a limit point. Let $A_n = \lim_{t\to x} f_n(t)$. Then $(A_n)_n$ converges and f has a limit at x and $\lim_{n\to\infty} A_n = \lim_{t\to x} f(t)$. Then*

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

Corollary: if $f_n \to f$ uniformly, and each f_n is continuous, then f is continuous.

Theorem 24.8. *If $f_n \to f$ uniformly. If f_n is integrable, then f is integrable, and*

$$\lim_{n\to\infty} \int_a^b f_n(t)dt = \int_a^b \lim_{n\to\infty} f_n(t)dt = \int_a^b f(t)dt.$$

Theorem 24.9. *If f_n differentiable on $[a, b]$, and f_n converges pointwise to f and f_n' converges uniformly to some g, then f_n converges uniformly to f such that f is differentiable with $f' = g$.*