MATH 112, SPRING 2019

WITH DENIS AUROUX

CONTENTS

PRELIMINARIES

These notes were taken during the spring semester of 2019 in Harvard's Math 112, *Introductory Real Analysis*. The course was taught by Dr. Denis Auroux and transcribed by Julian Asilis. The notes have not been carefully proofread and are sure to contain errors, for which Julian takes full responsibility. Corrections are welcome at `asilis@college.harvard.edu`.

1. Lecture 1 — January 29, 2019

One of the goals of the course is to rigorously study real functions and things like integration and differentiation, but before we get there we need to be careful about studying sequences, series, and the real numbers themselves.

The real numbers have lots of operations that we use frequently without too much thought: addition, multiplication, subtraction, division, and ordering (inequalities). One of today's goals is to convince you that even before we get there, describing the real numbers rigorously is actually quite difficult.

Definition 1.1. A set is a collection of elements.

Sets can be finite or infinite (there are different kinds of infinities), and they are not ordered. For a set A, $x \in A$ means that x is an element of A. $x \notin A$ means that x is not an element of A. One special set is the empty set , which contains no elements. Other important sets include that of the natural numbers $\mathbb{N} = \{0, 1, 2, 3, \dots\}$, that of the integers $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$, and that of the rationals $\mathbb{Q} = \{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0\}$

If every element of a set A is an element of a set B, we say A is a subset of B, and write $A \subset B$. An example we've already seen is $\mathbb{N} \subset \mathbb{Z}$. For sets, A = B if and only if (iff) A⊂B and B⊂A.

Definition 1.2. A field is a set F equipped with the operations of addition(+) and multiplication($\cdot$), satisfying the field axioms. For addition,

- If $x \in F, y \in F$ then $x + y \in F$
- $x + y = y + x$ (commutativity)
- $(x + y) + z = z + (y + z)$ (associativity)
- F contains an element $0 \in F$ such that $0 + x = x \ \ \forall x \in F$
- $\forall x \in F$, there is $-x \in F$ such that x + (-x) = 0

And for multiplication,

- If $x \in F, y \in F$ then $x \cdot y \in F$
- $x \cdot y = y \cdot x$ (commutativity)
- $(x \cdot y) \cdot z = z \cdot (y \cdot z)$ (associativity)
- F contains an element $0 \neq 1 \in F$ such that $1 \cdot x = x \ \ \forall x \in F$
- $\forall x \in F$, there is $\frac{1}{x} \in F$ such that $x \cdot \frac{1}{x}$ = 1

Finally, multiplication must distribute addition, meaning $x(y + z) = xy + zx \ \ \forall x, y, z \in F$.

The operation of multiplication is usually shortened from $(\cdot)$ to concatenation for convenience's sake, so that $x \cdot y$ be written xy. One example of a field is $\mathbb{Q}$ with the familiar operations of addition and multiplication.

Proposition 1.3. *The axioms for addition imply:*

(1) *If $x + y = x + z$, then $y = z$ (cancellation)*
(2) *If $x + y = x$, then $y = 0$*
(3) *If $x + y = 0$, then $y = -x$*
(4) $-(-x) = x$

Proof. (1). Assume $x + y = x + z$. Then:

$$\begin{aligned} x + y &= x + z \\ (-x) + (x + y) &= (-x) + (x + z) \\ ((-x) + x) + y &= ((-x) + x) + z \\ 0 + y &= 0 + z \\ y &= z \end{aligned}$$

(2) follows from (1) by taking $z = 0$. (3) and (4) take a bit more work, and are good practice to complete on your own. It's worth noting that nearly identical properties (with nearly identical proofs) hold for multiplication. $\square$

Definition 1.4. An <u>ordered set</u> is a set S equipped with a relation $(<)$ satisfying:

- $\forall x, y \in S$, exactly one of $x < y$, $x = y$, or $y < x$ is true.
- If $x < y$ and $y < z$, then $x < z$ (transitivity)

We will write $x \leq y$ to mean $x < y$ or $x = y$ (and because of the above definition, this is an exclusive or).

Definition 1.5. An <u>ordered field</u> $(F, +, \cdot, <)$ is a field with a compatible order relation, meaning:

- $\forall x, y, z \in F$ If $y < z$ then $x + y < x + z$
- If $x > 0$ and $y > 0$ then $xy > 0$

$\mathbb{Q}$ was our example of a field, and fortunately it still works as an example, as $\mathbb{Q}$ is an ordered field under the usual ordering on rationals.

Proposition 1.6. *In an ordered field:*

- *If $x > 0$ then $-x < 0$, and vice versa*
- *If $x > 0$ and $y < z$, then $xy < xz$*
- *If $x < 0$ and $y < z$ then $xy > xz$*
- *If $x \neq 0$, then $x^2 > 0$. Thus $1 > 0$*
- $0 < x < y \implies 0 < \frac{1}{y} < \frac{1}{x}$

Now we'll talk about what's wrong with the rational numbers. As you may expect, we'll begin by considering the square root of 2.

Proposition 1.7. *There does not exist $x \in \mathbb{Q}$ such that $x^2 = 2$*

Proof. Assume otherwise, so $\exists x = \frac{m}{n} \in \mathbb{Q}$ such that $x^2 = 2$. Take x to be a reduced fraction, meaning that m and n share no factors. Then $\frac{m^2}{n^2} = 2$ and $m^2 = 2n^2$ for $m, n \in \mathbb{Z}, n \neq 0$. $2n^2$ is even, so m^2 is even. Since the square of an odd number is odd, m must be even. So $m = 2k$ for some $k \in \mathbb{Z}$. We have $m^2 = (2k)^2 = 4k^2 = 2n^2$. Dividing by 2, we see $2k^2 = n^2$. Using our reasoning from above, we see that n must be even. So m and n are both even, which is a contradiction. $\square$

It seems like we could formally add an element called the square root of 2, and do so for similar algebraic numbers which appear as solutions to polynomials with rational coefficients, but this still wouldn't solve our problem. The problem is that sequences of rational numbers can look to be approaching a number, but not have a limit in $\mathbb{Q}$.

Definition 1.8. Suppose $E \subset S$ is a subset of an ordered set. If there exists $\beta \in S$ such that $x \leq \beta$ for all $x \in E$, then E is bounded above, and β is one of its upper bounds.

The definition for lower bounds is similar. In general, sets may not have upper or lower bounds (think $\mathbb{Z} \subset \mathbb{Q}$).

Definition 1.9. Suppose S is an ordered set and $E \subset S$ is bounded above. If $\exists \alpha \in S$ such that:

(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 $\alpha = \sup E$.

Example 1.10. Consider $\{x \in \mathbb{Q} : x < 0\}$ as a subset of $\mathbb{Q}$. Any rational $y \geq 0$ is an upper bound, and you can see that 0 is the least upper bound.

Now take $A = \{x \in \mathbb{Q} : x < 0 \text{ or } x^2 < 2\}$ as a subset of $\mathbb{Q}$. The upper bounds of A in $\mathbb{Q}$ are $B = \{x \in \mathbb{Q} : x > 0 \text{ and } x^2 > 2\}$. It turns out that there's no least upper bound here. Though it's a bit opaque, any upper bound y has a lower upper bound $\frac{2y+2}{y+2}$. This suggests that increases sequences of rationals which square to less than 2 have no limit, and likewise for positive, decreasing rationals which square to more than 2.

Theorem 1.11 (Completeness)**.** *There exists an ordered field $\mathbb{R}$ which has the least upper bound property, meaning every non-empty subset bounded above has a least upper bound.*

2. Lecture 2 — January 31, 2019

Last time we talked about least upper bounds and the fact that their existence isn't always guaranteed in $\mathbb{Q}$. Greatest lower bounds are defined analogously, and their existence also isn't guaranteed in $\mathbb{Q}$. As it turns out, this is more than coincidence, since these properties are equivalent.

Theorem 2.1. *If an ordered set S has the least upper bound property, then it also has the greatest lower bound property.*

Proof. We won't prove this rigorously, but here's the idea: given a set $E \subset S$ bounded below, consider its set of lower bounds L. L isn't empty because we assumed E is bounded below, and it's bounded above by all elements of E. So, because S satisfies the least upper bound property, L has a least upper bound. You can show that this is the greatest lower bound of E. $\square$

Last time, we also saw the following important theorem.

Theorem 2.2. *There exists an ordered field $\mathbb{R}$ with the least upper bound property which contains $\mathbb{Q}$ as a subfield.*

Proof. There are two equivalent ways of doing this - one uses things called Cauchy sequences that we'll be encountering later on, and the second uses Dedekind cuts. A cut is a set $\alpha \subset \mathbb{Q}$ such that

(1) $\alpha \neq \emptyset$ and $\alpha \neq \mathbb{Q}$
(2) If $p \in \alpha$ and $q < p$ then $q \in \alpha$
(3) If $p \in \alpha$, $\exists r \in \alpha$ with $p < r$

In practice, $\alpha = (-\infty, a) \cap \mathbb{Q}$, though $(-\infty, a)$ doesn't technically mean anything right now. So we've constructed a set (of subsets) which we claim is $\mathbb{R}$, and now we have to endow it with an order and operations respecting that order in order to get an ordered field. We'll define the order as such: for $\alpha, \beta \in \mathbb{R}$, we write $\alpha < \beta$ if and only if $\alpha \neq \beta$ and $\alpha \subset \beta (\subset \mathbb{Q})$. This is in fact an order.

To see that least upper bounds exist, we claim that the least upper bound of a non-empty, bounded above $E \subset \mathbb{R}$ is the union of its cuts. You have to check that this is a cut and in fact a least upper bound.

We define addition of cuts as $\alpha + \beta = \{p + q : p \in \alpha, q \in \beta\}$. The definition of multiplication is a bit uglier and depends on the 'signs' of cuts. Then you have to check that all the field axioms are satisfied. It's not really worth getting into all of the details here, but people have at some point checked that everything works as we'd like it to. $\square$

Theorem 2.3 (Archimedean property of $\mathbb{R}$)**.** *If $x, y \in \mathbb{R}$, $x > 0$, then there exists a positive integer n such that $nx > y$*

Proof. Suppose not, and consider $A = \{nx : n \text{ a positive integer}\}$. A is non-empty and has upper bound y, so it has a least upper bound, which we'll call α. $\alpha - x < \alpha$ because $x > 0$, so $\alpha - x$ is not an upper bound. Then $\exists nx \in A$ such that $nx > \alpha - x$. But adding x to both sides, we have $nx + x = (n+1)x > \alpha$. But $(n+1)x \in A$, so α was not an upper bound at all. $\square$

Theorem 2.4 (Density of $\mathbb{Q}$ in $\mathbb{R}$)**.** *If $x, y \in \mathbb{R}$ and $x < y$, then $\exists p \in \mathbb{Q}$ such that $x < p < y$.*

Proof. Since $x < y$, we have $y - x > 0$. By the previous theorem, there exists an integer n with $n(y - x) > 1$, meaning $y - x > \frac{1}{n}$. Also by the previous theorem, there exist integers m_1, m_2 with $m_1 > nx$ and $m_2 > -nx$, i.e. $-m_2 < nx < m_1$. Thus there exists an integer m between $-m_2$ and m_1 with $m - 1 \leq nx < m$. Then $nx < m \leq nx + 1 < nx + n(y - x) = ny$. Diving by n, we have $x < \frac{m}{n} < y$, and the $p = \frac{m}{n}$ that we wanted. $\square$

"The rational numbers are everywhere. They're among us." - Dr. Auroux. What we're saying is that between any two reals there's a rational. A problem we encountered last class is that we weren't guaranteed the existence of square roots in $\mathbb{Q}_{\geq 0}$. Fortunately, this has been remedied by constructing $\mathbb{R}$.

Theorem 2.5. *For every real $x > 0$ and every integer $n > 0$, there exists exactly one $y \in R, y > 0$ with $y^n = x$. We write $y = x^{\frac{1}{n}}$.*

Proof sketch. Consider $E = \{t \in R : t > 0, t^n < x\}$. It's non-empty and bounded above, so it has a supremum we'll call α. If $\alpha^n < x$, then α isn't an upper bound of E, and if $\alpha^n > x$, it's not the *least* upper bound of E. $\square$

Definition 2.6. The extended real numbers consist of $\mathbb{R} \cup \{-\infty, \infty\}$ with the order $-\infty < x < \infty$ for all $x \in \mathbb{R}$ and the operations $x \pm \infty = \pm\infty$.

Notice that the extended real numbers don't form a field since, among other reasons, $\pm\infty$ don't have multiplicative inverses.

Definition 2.7. The complex numbers ($\mathbb{C}$) consist of the set $\{(a, b,) : a, b \in \mathbb{R}\}$ equipped with the operations $(a, b) + (c, d) = (a + c, b + d)$ and $(a, b) \cdot (c, d) = (ac - bd, ad + bc)$. These operations make $\mathbb{C}$ a field.

It's convention to write $(a, b) \in \mathbb{C}$ as $a + bi$. The complex conjugate of $z = a + bi$ is $\overline{z} = a - bi$, and the norm of a complex number $z = a + bi$ is $|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$.

Proposition 2.8. *For all $z \in \mathbb{C}$,*

- $|z| \geq 0$ *and* $|z| = 0$ *iff* $z = 0$
- $|zw| = |z||w|$
- $|z + w| \leq |z| + |w|$

Definition 2.9. Euclidean space is $\mathbb{R}^k = \{(x_1, \dots, x_k) : x_i \in \mathbb{R}\}$ equipped with $\overrightarrow{x} + \overrightarrow{y} = (x_1 + y_1, \dots, x_k + y_k)$ and $\alpha \overrightarrow{x} = (\alpha x_1, \dots, \alpha x_k)$ for $\alpha \in \mathbb{R}$.

Theorem 2.10. *Defining $\overrightarrow{x} \cdot \overrightarrow{y} = \sum_{i=1}^{k} x_i y_i$ and $||x||^2 = \overrightarrow{x} \cdot \overrightarrow{x}$, we have:*

- $||x||^2 \geq 0$ *and* $||x||^2 = 0 \iff \overrightarrow{x} = 0$
- $||\overrightarrow{x} \cdot \overrightarrow{y}|| \leq ||\overrightarrow{x}|| \cdot ||\overrightarrow{y}||$
- $||\overrightarrow{x} + \overrightarrow{y}|| \leq ||\overrightarrow{x}|| + ||\overrightarrow{y}||$

Proof. (1) Clear
(2) Some ugly computation

□

3. Lecture 3 — February 5, 2019

Today we'll be talking about sets.

Definition 3.1. For A, B sets, a function $f : A \to B$ is an assignment to each $x \in A$ of an element $f(x) \in B$

A is referred to as the domain of f, and the range of f is the set of values taken by f (in this case, a subset of B). For $E \subset A$, we take $f(E) = \{f(x) : x \in E\}$. In this notation, the range of f is $f(A)$. On the other hand, for $F \subset B$, we define the inverse image, or pre-image, of F to be $f^{-1}(F) = \{x \in A : f(x) \in F\}$. Note that the pre-image of an element in B can consist of one element of A, several elements of A, or be empty. It's always true that $f^{-1}(B) = A$.

Definition 3.2. A function $f : A \to B$ is onto, or surjective, if $f(A) = B$. Equivalently, $\forall y \in B, f^{-1}(y) \neq \emptyset$

Definition 3.3. A function $f : A \to B$ is one-to-one, or injective, if $\forall x, y \in A$, $x \neq y \implies f(x) \neq f(y)$. Equivalently, $f(x) = f(y) \implies x = y$. Also equivalently, $\forall z \in B$, $f^{-1}(z)$ contains at most one element.

Definition 3.4. A function is a one-to-one correspondence, or bijection, if it is one-to-one and onto, i.e. $\forall y \in B, \exists! x \in A$ s.t. $f(x) = y$.

Defining 'size', or cardinality, of finite sets is not too difficult, but extending this notion to infinite sets is fairly difficult. Regardless of what the notion of size for infinite sets should be, it should definitely be preserved by bijections (meaning that if A and B admit a bijection between each other, they should have the same size). So we say that two sets have the same cardinality, or are equivalent, if there exists a bijection between them.

Let $J_n = \{1, \ldots, n\}$ for $n \in \mathbb{N}$ and $J_0 = \emptyset$.

Definition 3.5. A set A is finite if it is in bijection with J_n for some n. Then $n = |A|$. A set A is infinite if it is not finite.

Definition 3.6. A set A is countable if it is in bijection with $\mathbb{N} = \{1, 2, 3, \ldots\}$.

Informally, countability means that a set can be arranged into a sequence.

Definition 3.7. A set A is at most countable if it is finite or countable.

The above definition captures the idea that countability is the smallest infinity.

Definition 3.8. A set A is uncountable if it is infinite and not countable.

When sets are in bijection, we think of them as having the same number of elements. Extremely counter-intuitive pairs of sets which we then think of as having the same number of elements arise.

Example 3.9. $\mathbb{Z}$ is in bijection with $\mathbb{N}$. The map is

$$f(z) \begin{cases} \frac{z-1}{2} & z \text{ is odd} \\ \frac{-z}{2} & z \text{ is even} \end{cases}$$

In the above example, we construct a bijection between $\mathbb{Z}$ and a proper subset of $\mathbb{Z}$, $\mathbb{N}$. This is a property of infinite sets, and in fact can considered the defining property of infinite sets.

Definition 3.10. A sequence in a set A is a function from $\mathbb{N}$ to A.

By convention, $f(n)$ is written x_n, and the sequence itself is written $\{x_n\}_{n\geq 1}$. Despite the brackets, $\{x_n\}_{n\geq 1}$ is *not* a set - it cares about order and allows for repeated elements.

Theorem 3.11. *An infinite subset of a countable set is countable.*

Proof. Let A be countable $E \subset A$ an infinite subset. Then a bijection $\mathbb{N} \to A$ gives a sequence $\{x_n\}_{n\geq 1}$ whose terms lie in A. We construct a sequence of integers $\{n_k\}_{k\geq 1}$ via the procedure n_1 equals the smallest integer n_1 such that $x_{n_1} \in E$. Having chosen $n_1, \ldots, n_{k-1}$, define n_k to be the smallest integer strictly greater than n_{k-1} such that $x_{n_k} \in E$. This procedure never terminates because E is infinite. Now set $f : \mathbb{N} \to E$, $k \mapsto x_{n_k}$. This injects because all the x_i are distinct (because they were defined using an injection $\mathbb{N} \to A$). This surjects because all $e \in E \subset A = \{x_1, x_2, x_3, \ldots\}$ appear at some point in the sequence $\{x_i\}_{i\geq 1}$ and have their indices selected by our procedure. $\square$

Definition 3.12. For sets A, B, the set $A \cup B$ consists exactly of things which are elements of A and/or elements of B. More generally, given a collection of sets E_α indexed by $\alpha \in \Lambda$, define $S = \bigcup_{\alpha \in \Lambda} E_\alpha$ to be the set such that $x \in S$ if and only if there exists $\alpha \in \Lambda$ with $x \in E_\alpha$.

Definition 3.13. For sets A, B, the set $A \cap B$ consists exactly of things which are elements of A and elements of B. Similarly, $S = \bigcap_{\alpha \in A} E_\alpha$ is defined by $x \in S \iff x \in E_\alpha \ \forall \alpha \in \Lambda$.

> **Example 3.14.** Take $A = \{x \in R : 0 < x \leq 1 = (0, 1]$. For $x \in A$, let $E_x = \{y \in R : 0 < y < x\}$. Then $E_x \subset E_{x'}$ if and only if $x \leq x'$. And $\bigcup_{x \in A} E_x = E_1 = (0, 1)$. On the other hand, $\bigcap_{x \in A} E_x = \emptyset$.

Proposition 3.15 (Sets form an algebra)**.** *(1)* $A \cup B = B \cup A$

(2) $(A \cup B) \cup C = A \cup (B \cup C)$

(3) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Theorem 3.16. *Let $\{E_n\}_{n \geq 1}$ be a sequence of countable sets. Then $\bigcup_{i=1}^{\infty} E_n = S$ is countable.*

Proof. Taking $E_1 = \{x_{11}, x_{12}, x_{13}, \ldots\}$, $E_2 = \{x_{21}, x_{22}, x_{23}, \ldots\}$, and so on, we can arrange the elements of S in a sequence like so: $S = \{x_{11}, x_{21}, x_{12}, x_{31}, x_{22}, x_{13}, \ldots\}$. Visually, we're arranging the E_i in a ray and proceeding along diagonal line segments starting on the top left. This certainly isn't rigorous, but it's the essential idea. $\square$

One corollary to this is that if A is at most countable and for each $\alpha \in A$, E_α is at mot countable, then $\bigcup_{\alpha \in A} E_\alpha$ is at most countable.

Theorem 3.17. *If A is countable, then A^n is countable.*

Proof. We induct on n. When $n = 1$, the claim follows by assumption. If A^{n-1} is countable, then $A^n = \bigcup_{a \in A^{n-1}} A$. Then A^n is a countable union of countable sets, and thus countable. $\square$

A result of this is that $\mathbb{Q}$ is countable, as it can be realized as a subset of $\mathbb{Z}^2$ via the function $\frac{m}{n}$, in reduced form, maps to (m, n).

4. Lecture 4 — February 7, 2019

Last time we saw that the countable union of countable sets is countable. It turns that adding all solutions to polynomials over $\mathbb{Z}$, and forming what are called the algebraic numbers, still leaves you with countably many numbers.

Theorem 4.1. *$\mathbb{R}$ is uncountable. Equivalently, the set A of sequences in $\{0, 1\}$ is uncountable.*

Proof. Suppose A is countable, meaning its elements can be listed sequentially. Then A can be written as the collection

$$\begin{aligned} S_1 &= S_{11},\ S_{12},\ S_{13}, \ldots \\ S_2 &= S_{21},\ S_{22},\ S_{23}, \ldots \\ S_3 &= S_{31},\ S_{32},\ S_{33}, \ldots \\ &\vdots \end{aligned}$$

where each $S_{ij} \in \{0, 1\}$ and every sequence in $\{0, 1\}$ appears exactly once in this sequence of S_i. But consider the sequence

$$M = \begin{cases} 0 & S_{nn} = 1 \\ 1 & S_{nn} = 0 \end{cases}$$

M differs from S_n at the nth term, so the sequence of S_i fails to include all sequences in $\{0, 1\}$. $\square$

A corollary to this is that the set of subsets of $\mathbb{N}$ is uncountable, since there's a correspondence between such subsets and sequences in $\{0, 1\}$ via the rule that a sequence's nth term is 1 if $n \in N$ is in the subset under consideration. This is more than coincidence - the collection of subsets of any set, referred to as the power set of that set, is always strictly larger than that set.

Now we're going to pivot to metric topology. Informally, a metric space is a set equipped with a notion of distance, which is the kind of structure we'll need to discuss limits, continuity, and so on.

Definition 4.2. A metric space consists of a set X equipped with a distance function, or metric, $d : X \times X \to \mathbb{R}$ such that $\forall p, q, r \in X$

(1) $d(p, q) \geq 0$, with equality iff $p = q$
(2) $d(p, q) = d(q, p)$
(3) $d(p, q) \leq d(p, r) + d(r, q)$ [Triangle Inequality]

Our go-to examples for now are $\mathbb{R}$ equipped with the metric $d(x, y) = |x - y|$ and $\mathbb{R}^k$ with the metric $d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_k - y_k)^2}$. From here on out, we'll refer to $\mathbb{R}$ and $\mathbb{R}^k$ as metric space without specifying their metrics, and we'll be using these two metrics. Note that a subset of a metric space is always a metric space, with the metric induced by its parent set.

A natural thing to discuss now is the notion of proximity.

Definition 4.3. Let X be a metric space under the function d:

- A neighborhood of $p \in X$ is a set $N_r(p)$, for some radius $r \in \mathbb{R}_+$, consisting of $q \in X$ such that $d(p, q) < r$.
- p is an interior point of $E \subset X$ if there exists a neighborhood N of p such that $N_r(p) \subset E$ for some $r > 0$.
- $E \subset X$ is open if every point of E is an interior point.

"This stuff is slightly mind-bending and will build on itself and become even more mind-bending by next week." - Dr. Auroux.

Example 4.4. In $\mathbb{R}$, $N_r(p) = (p-r, p+r) = \{x \in R : p-r < x < p+r\}$. Also in $\mathbb{R}$, the interior points of $[a,b]$ are (a,b), meaning $[a,b]$ is not open.

Theorem 4.5. *Every neighborhood is an open set.*

Proof. Let $E = N_r(p)$, and take $x \in E$. Then $d(p,x) < r$, and let $h = r - d(p,x) > 0$. We claim $N_h(x) \subset E$. By the triangle inequality, for any $y \in N_h(x)$, $d(p,y) \leq d(p,x) + d(x,y) < d(p,x) + (1 - d(p,x)) = r$. So $y \in E$, and E contains $N_h(x)$, making x an interior point of E. Since x was selected arbitrarily, all of E's points are interior points and E is open. □

Definition 4.6. Let X be a metric space

- A point $p \in X$ is a limit point of $E \subset X$ if every neighborhood of p contains a point $q \in E$ such that $q \neq p$.
- If $p \in E$ is not a limit point, then it is an isolated point of E.

Notice that isolated points are obligated to members of E while limit points are not.

Example 4.7. Take $E = \{\frac{1}{n} : n = 1, 2, 3, \dots\} \subset \mathbb{R}$. Then 1 is isolated (consider $N_{\frac{1}{4}}(1)$). On the other hand, 0 is a limit point of E, since $\frac{1}{n} < r$ for any $r > 0$ for sufficiently large $n \in \mathbb{N}$, meaning E intersects $N_r(0)$ for any $r > 0$.

In $\mathbb{R}$, the limit points of (a,b) are $[a,b]$. Likewise, the limit points of $[a,b]$ are $[a,b]$.

Definition 4.8. $E \subset X$ is closed if it contains all its limit points.

Proposition 4.9. *In any metric space X, X itself and $\emptyset$ are always both open and closed.*

An important note is that the quality of a set being open or closed is *not* a property of the set itself but of the set in which it lives. Strictly speaking, it doesn't make sense to say E is an open set (though, we'll slightly abuse terminology and start saying that anyway). It only makes sense to say E is an open subset of X.

Theorem 4.10. *If p is a limit point of E in X, then every neighborhood of p contains infinitely many points of E.*

Proof sketch. If there were only finitely many points of E in a neighborhood of p, then one could construct a neighborhood around p whose radius is the minimum of p's distance to these points. This neighborhood doesn't contains any points of E, contradicting the fact that p is a limit point. □

A corollary is that finite sets don't have limit points.

Definition 4.11. A subset E of a metric space X is bounded if there exists $q \in X$ and $M > 0$ such that $E \subset N_M(q)$.

Definition 4.12. The complement of a subset $E \subset X$ is $E^c = \{p \in X : p \notin E\}$.

Theorem 4.13 (De Morgan's Laws)**.** *Let E_α be an arbitrary collection of subsets of X. Then $(\bigcup_\alpha E_\alpha)^c = \bigcap_\alpha E_\alpha^c$.*

Now we reveal an important relationship between open and closed sets, which is not quite one of being 'opposite'.

Theorem 4.14. *$E \subset X$ is open if and only if E^c is closed.*

Proof. "This is a game of negations." - Dr. Auroux. First suppose E^c is closed. Let $x \in E$. Since E^c is closed, x is not a limit point of E^c. Then there exists a neighborhood of x which contains no points in E^c distinct from x. Since x isn't in E^c either, this neighborhood lies entirely in E, meaning x is an interior point of E. We're out of time, but the reverse direction of the proof is very similar. □

5. Lecture 5 — February 12, 2019

Recall our definitions from last class - the interior points of a set are those which admit neighborhoods within the set, limit points of a set are points (not necessarily within the set) whose neighborhoods always contain points of that set, and open sets consist of their interior points while closed sets contain their limit points.

We showed last time that every neighborhood of a limit of a set contains infinitely many points in that set, and that a set is open if and only if its complement is closed.

Theorem 5.1.

(1) If G_α are open in X, $\bigcup_{\alpha \in A} G_\alpha$ is open in X.
(2) If F_α are closed in X, $\bigcap_{\alpha \in A} F_\alpha$ is closed in X.
(3) If $G_1, \ldots, G_n$ are open, $\bigcap_{i=1}^n G_i$ is open.
(4) If $F_1, \ldots, F_n$ are closed, $\bigcup_{i=1}^n F_i$ is closed.

Proof. Because a set is open if and only if its complement is closed, and because of DeMorgan's laws, it suffices to prove only (a) and (c). For (a), assume $x \in \bigcup G_\alpha$. Then $x \in G_\alpha$ for some α, and because G_α is open, $\exists r$ such that $B_r(x) \subset G_\alpha \subset \bigcup_{\alpha \in A} G_\alpha$. For (c), suppose $x \in \bigcap G_i$, and let r_i be the radius such that $B_{r_i}(x) \subset G_i$. Taking $r = \min(r_i)$, we have $B_r(X) \subset G_i\ \forall i$ and thus $B_r(X) \subset \bigcap G_i$. □

It's worth looking at counter-examples to see that we can't do any better than finite intersections or unions for open and closed sets, respectively.

Example 5.2. $\bigcap_{k=1}^\infty (-\frac{1}{k}, \frac{1}{k}) = \{0\}$, so infinite unions of open sets are not in general open. Additionally, $\bigcup_{k=2}^\infty [\frac{1}{k}, 1 - \frac{1}{k}] = (0,1)$, so infinite unions of closed sets are not in general closed.

Definition 5.3. The interior of a set $E \subset X$, written $\mathring{E}$, consists of all interior points of E.

Theorem 5.4.

- *$\mathring{E}$ is open.*
- *If $F \subset E$ and F is open then $F \subset \mathring{E}$ (i.e. $\mathring{E}$ is the largest open subset contained in E).*

Proof.

- Say $x \in \mathring{E}$, so we have r such that $B_r(X) \subset E$. We claim that $B_r(X) \subset \mathring{E}$, meaning x is an interior point of $\mathring{E}$. This follows from openness of open neighborhoods; for any $y \in B_r(X)$, there exists an r_y such that $B_{r_y}(y) \subset B_r(X) \subset E$. So y is an interior point of E and thus x is an interior point of $\mathring{E}$.
- Any $x \in F$ admits a $B_r(X) \subset F$. And $B_r(X) \subset E$, so $x \in \mathring{E}$.

□

Definition 5.5. The closure of E, written $\overline{E}$, is its union with the set of its limit points.

Theorem 5.6.

(1) $\overline{E}$ is closed.
(2) $E = \overline{E} \iff E$ is closed.
(3) If $F \supset E$ and F is closed, then $F \supset \overline{E}$. (i.e. $\overline{E}$ is the smallest closed set containing E).

Proof.

(1) If $p \in X$ and $p \notin \overline{E}$, then p is not in E and it's not a limit point of E. So there exists a $B_r(p)$ which does not intersect E. So p is an interior point of E^c. The interior of E^c is open, by the previous theorem, so $\overline{E}$ is closed.
(2) Clear
(3) Also follows from $(\overline{E})^c = (\mathring{E^c})$

□

Definition 5.7. $E \subset X$ is dense if $\overline{E} = X$

Example 5.8. $\mathbb{Q}$ is dense in $\mathbb{R}$, since any neighborhood around a real number contains rationals.

When $E \subset Y \subset X$, we say E is open relative to Y if E is an open subset of Y. To see why this distinction is important, consider $\{x \in \mathbb{Q} : x^2 < 2\} = (-\sqrt{2}, \sqrt{2}) \cap \mathbb{Q}$. This set is closed in $\mathbb{Q}$, but not in $\mathbb{R}$.

Theorem 5.9. *Let $E \subset Y \subset X$. Then E is open relative to Y if and only if $E = G \cap Y$ for some open $G \subset X$.*

Similarly, $E \subset Y \subset X$ is closed relative to Y if and only if $E = F \cap Y$ for some closed F in X.

6. Lecture 6 — February 14, 2019

"A compact set is the next best friend you can have after a finite set." - Dr. Auroux. You have may have already seen a theorem in calculus which states that continuous functions $f : [a, b] \to \mathbb{R}$ are necessarily bounded and contain their maxima/minima. It turns out to be the case that for a more general continuous function $f : K \to Y$ between metric spaces with K compact, $f(K)$ must be compact as well. This will imply that $f(K)$ is bounded and closed (meaning it contains its maximum/minimum).

Definition 6.1. An open cover of a subset E in a metric space X is a collection of open sets $\{G_\alpha\}$ such that $\bigcup_{\alpha \in A} G_\alpha \supset E$.

Definition 6.2. A subset K of a metric space X is compact if every open cover of K has a finite subcover, meaning $\exists \alpha_1, \ldots, \alpha_n \in A$ such that $K \subset (G_{\alpha_1} \cup \cdots \cup G_{\alpha_n})$.

This definition is pretty opaque right now - let's look at some examples.

Example 6.3. Any finite set is compact. In the worst case, any open cover can be reduced to a subcover containing one open set for each of the set's elements.

It's somewhat miraculous that infinite compact sets exist at all. It would be pretty hard to prove right now that $[a, b]$ is compact given only the definition, but we'll get to a proof next week after developing some tools. As is the case with most definitions containing the word , it's much easier to prove that a set is not compact than to prove that it is.

Example 6.4. $\mathbb{R}$ is not compact. It suffices to provide a single cover which does not admit a finite subcover. Consider the cover $\{(-n, n)\}_{n\in\mathbb{N}}$. This covers, because every element of $\mathbb{R}$ lies in $(-n, n)$ for some n, but any finite collection of subsets amounts to a single interval $(-m, m)$, which fails to cover $\mathbb{R}$.

The problem we have right now is that is that it's very difficult to prove that a set is compact. For now, let's think wishfully and consider the results we could conclude if we knew a set were open. The first remarkable result is that, unlike openness, the compactness of set in a metric space is a function only of the set and its metric, and not of the metric space in which it resides. Simply put, it makes sense to say 'the set K is closed under the metric d', whereas it didn't make sense to say 'the set K is open under the metric d' (in the second case, it matters what set K lives in).

Theorem 6.5. *Suppose $K \subset Y \subset X$ are metric spaces. Then K is compact as a subset of X if and only if K is compact as a subset of Y.*

Proof. Suppose K is compact relative to X. Assume $\{V_\alpha\}$ are open subsets of Y which cover K. For each α, there exists an open $G_\alpha \subset X$ such that $V_\alpha = Y \cap G_\alpha$. The G_α form an open cover of K in X. By compactness of X, this can be reduced to a finite cover $G_{\alpha_1}, \ldots, G_{\alpha_n}$. We then have:

$$\begin{aligned} V_{\alpha_1} \cup \cdots \cup V_{\alpha_n} &= (G_{\alpha_1} \cap Y) \cup \cdots \cup (G_{\alpha_n} \cap Y) \\ &= (G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}) \cap Y \\ &\supset K \cap Y \\ &= Y \end{aligned}$$

So $V_{\alpha_1}, \ldots, V_{\alpha_n}$ form a finite subcover of K in Y, and K is compact in Y. In the other direction, take a cover of K in X, intersect its constituent open sets with Y, and reduce it to a finite subcover of K in Y. Then notice that the corresponding open sets in X form a finite subcover of K. $\square$

Theorem 6.6. *Compact sets are bounded.*

Proof. Consider the open cover $K \subset \bigcup_{p\in K} N_1(p)$. Since K is compact, $K \subset N_1(p_1) \cup \cdots \cup N_1(p_n)$. Then given any two points $q, r \in K$, $q \in N_1(p_i)$ and $r \in N_1(p_j)$ for some i, j. Then, by the triangle inequality, $d(r, q) \leq d(q, p_i) + d(p_i, p_j) + d(p_j, r) \leq 2 + d(p_i, p_j)$. It follows that the distance between any two points in K is at most $\max\{d(p_i, p_j)\} + 2$, so it's bounded. $\square$

Theorem 6.7. *Compact sets are closed.*

Proof. Say $K \subset X$ is compact. Take $p \in X, p \notin K$. The goal is to show that p is not a limit point of K, meaning there's a neighborhood of p that doesn't intersect K. For $q \in K$, we can construct neighborhoods of p and q that don't intersect each other. Take $V_q = N_r(q)$ and $W_q = N_r(p)$ for $r = \frac{d(p,q)}{3}$. Constructing such V_q, W_q for all $q \in K$, we see that the V_q collectively cover K. Since K is compact, they can be reduced to a finite subcover $V_{q_1}, \ldots, V_{q_n}$. Now let $W = W_{q_1} \cap \cdots \cap W_{q_n}$. Since $W \cap V_{q_i} \subset W_{q_i} \cap V_{q_i} = \emptyset$ for each i, W is disjoint from $\bigcup_{i=1}^{n} V_{q_i} \supset K$. So $p \in W$ is not a limit point, and K is closed. $\square$

So no matter how you expand the universe that K lives in, you'll never construct points which are limit points of K.

Theorem 6.8. *Closed subsets of compact sets are compact.*

Proof. Take K compact (in some metric space X, though it doesn't matter), and let $F \subset K$ be closed (in K or, equivalently, in X). Given an open cover of F, consider its union with F^c. This covers K, so reduces to a finite subcover of K. Removing F^c from the finite subcover if necessary, we're left a finite subcover of F, as desired. $\square$

Theorem 6.9 (Nested Interval Property)**.** *Let K be a compact set. Any sequence of non-empty, nested closed subsets $K \supset F_1 \supset F_2 \supset F_3 \supset \ldots$ has non-empty intersection; $\bigcap_{n=1}^{\infty} F_n \neq \emptyset$.*

Proof. Suppose the intersection is empty. Let $G_n = F_n^c$. We have $\bigcup_{n=1}^{\infty} G_n = \bigcup_{n=1}^{\infty} F_n^c = (\bigcap_{n=1}^{\infty} F_n)^c = K$. So the G_n form a cover, and can be reduced to a finite subcover $G_{n_1}, \ldots, G_{n_k}$ for $n_1 < n_2 < \cdots < n_k$. So $F_{n_1} \cap \cdots \cap F_{n_k} = \emptyset$. But this intersection contains F_{n_k}, and we assumed that none of the F_i are empty, so we've arrived at a contradiction. $\square$

Theorem 6.10. *If $E \subset K$ is an infinite subset and K is compact, then E has a limit point.*

Proof. Say E doesn't have a limit point. So every point $p \in K$ admits a neighborhood V_p containing at most 1 point of E (p itself). The V_p cover K, so they can be reduced to a finite subcover of size, say, m. But then there are most m points in E, producing contradiction. $\square$

This property of a set is usually referred to as sequential compactness, because it turns out that it is equivalent to saying that every sequence in K has a convergent subsequence. We don't know what that means yet, but we'll get there in a few weeks.

7. Lecture 7 — February 19,2019

Worksheet! The important takeaway is that compactness and sequential compactness are equivalent in metric spaces.

8. LECTURE 8 — FEBRUARY 21, 2019

The solutions to Tuesday's worksheet have been posted online. Once more, recall that a subset of a metric space is compact if each of its open covers reduce to a finite subcover. We've seen that compactness is an intrinsic property, meaning it doesn't depend on the metric space that a set lives in (only the metric itself), and that compact sets are always closed and bounded.

In general, it's very rare for the converse to be true, meaning that closed and bounded sets are compact. One very important special case, however, is $\mathbb{R}^k$ under the Euclidean metric - sets here are compact if and only if they're closed and bounded.

Theorem 8.1 (Nested Interval Property)**.** *Suppose $I_i = [a_i, b_i]$ are a sequence of non-empty, nested closed intervals in $\mathbb{R}$. Then $\cap_{i=1}^{\infty} I \neq \emptyset$.*

Proof. Take $\alpha = \sup(a_i)$, as all the b_i are upper bounds of the a_i. Since α is an upper bound, it's greater than or equal to all the a_i. And since all the b_i are upper bounds, it's less than or equal to all the b_i. So it's in all the I_i, and it's in their intersection. ☐

Theorem 8.2. *Take $I_i = [a_1, b_1] \times \cdots \times [a_k, b_k]$ to be a sequence of nested k-cells in $\mathbb{R}^k$. Then their intersection isn't empty.*

Sketch. In each coordinate, the setup is as in the previous theorem (a sequence of non-empty, nested closed intervals in $\mathbb{R}$). By the previous result, there's thus a value in each coordinate which lies in the intersection of the sets restricted to that coordinate. Sewing those values together gives a point which is in the intersection of the closed sets. ☐

Theorem 8.3. *Every k-cell in $\mathbb{R}^k$ is compact.*

Sketch. Suppose that a k-cell I is equipped with an open cover which admits no finite subcover. Subdivide I into 2^k cells, (at least) one of which must fail to admit a finite subcover (otherwise I would admit a finite subcover). Subdivide this cell 2^k into 2^k cells, one of which must fail to admit a subcover. Continuing in this fashion, we obtain a sequence of k-cells $I_1, I_2, \ldots$ such that (taking D to be the the distance between the 'corners' of I)t:

(1) $I \supset I_1 \supset I_2 \supset \ldots$
(2) $I_n \subset \bigcup_\alpha G_\alpha$ doesn't have a finite subcover
(3) If $x, y \in I_n$, then $|x - y| \leq D/2^n$.

By the previous theorem, the intersection of the I_n is non-empty. Select some x in the intersection - it lies in G_{α_0} for some G_{α_0}. Since G_{α_0} is open, there exists an $r > 0$ such that $N_r(X) \subset G_{\alpha_0}$. Pick n such that $D/2^n < r$. Then $\forall y \in I_n$, $d(x, y) \leq D/2^n < r$, so $I_n \subset N_r(x) \subset G_{\alpha_0}$. But this contradicts (b), as we've found a finite subcover for I_n, namely just G_{α_0}. ☐

Theorem 8.4 (Heine-Borel)**.** *Subsets of $\mathbb{R}^k$, under the Euclidean metric, are compact if and only if they're closed and bounded.*

Proof. We've already seen that compact sets are closed and bounded (in fact, this is always true). In the other direction, any closed, bounded set can be witnessed as a subset of a sufficiently large k-cell in $\mathbb{R}^k$. So it's a closed subset of a compact set, which means it's compact. ☐

Theorem 8.5 (Weierstrauss)**.** *Every infinite bounded subset of* $\mathbb{R}^k$ *has a limit point.*

Proof. Since the set is bounded, it lives in a compact k-cell. Infinite subsets of compact sets have limit points, so the set has a limit point. □

Definition 8.6. Subsets $A, B \subset X$ are separated if $A \cap \overline{B} = \emptyset$ and $\overline{A} \cap B = \emptyset$.

Example 8.7 (0,1)**.** and (1,2) are disjoint but not separated. (0,1) and (1,2) are both disjoint and separated.

Definition 8.8. $E \subset X$ is connected if it cannot be decomposed into the union of non-empty separated sets.

As with compactness, this is an intrinsic property of E, irrespective of the larger metric space in which it lives. More explicitly, E is connected in X if and only if E is connected in E.

Notice that if $X = A \cup B$ with A, B separated then $B = A^c$. And $\overline{A} \cap B = \emptyset$, so $\overline{A} = A$, meaning A is closed. Similarly, B is closed. So A and B are both closed and (because their complements are closed) open. So X is connected if and only if the only 'clopen' sets are $\emptyset$ and X.

Theorem 8.9. $E \subset \mathbb{R}$ *is connected if and only if* $x, y \in E$ *and* $x < z < y \in E$ *implies* $z \in E$.

Proof. First suppose $x < z < y$ with $x, y \in E, z \notin E$. Then $E = (-\infty, z) \cup (z, \infty)$, so it's not connected. Now suppose E is not connected, meaning $E = A \cup B$ for A, B separated. Take $x \in A$ and $y \in B$. Assume without loss of generality that $x < y$. Let $z = \sup(A \cap [x, y])$. If $z \notin A$, then we're done. If $z \in A, z \notin B$, and you can find a nearby z' that produces contradiction. □

9. Lecture 9 — February 26, 2019

Definition 9.1. A sequence $\{p_n\}$ in a metric space X converges if $\exists p \in X$, the limit of the sequence, such that $\forall \epsilon > 0, \exists N$ such that $\forall n \geq N, d(p_n, p) < \epsilon$. Then we write $p_n \to p$, or $\lim_{n\to\infty} p_n = p$. If there exists no such p, we say that $\{p_n\}$ diverges.

This definition is a bit intimidating, but all it's saying is that for any open ball around the limit, the elements of a sequence eventually stay in the ball.

Definition 9.2. The range of a sequence $\{p_n\} \subseteq X$ is the set consisting of the sequence's elements. A sequence is bounded if its range is a bounded subsets of X.

Because sequences allow repetition, the range of a sequence can be finite - for instance, consider the range of $p_n = (-1)^n \in \mathbb{R}$.

Example 9.3. $p_n = (-1)^n \in \mathbb{R}$ diverges. On the other hand, $\lim_{n\to\infty} \frac{1}{n} = 0$.

Proposition 9.4. $p_n \to p \iff d(p_n, p) \to 0$

Proof. Note that the right hand side is a sequence in $\mathbb{R}$. That it converges to 0 means that $\forall \epsilon > 0 \exists N$ s.t. $\forall n \geq N$, $|d(p, p_n) - 0| < \epsilon$. But $|d(p, p_n) - 0|$ is just $d(p, p_n)$, so this in fact corresponds to the statement that $p_n \to p$. The other direction follows fairly directly from definition. □

Theorem 9.5. *$p_n \to p$ if and only if every neighborhood of p contains p_n for all but finitely many n.*

Proof.

$$\begin{aligned} p_n \to p &\iff \forall \epsilon > 0\ \exists N \text{ s.t } \forall n \geq N, p_n \in N_\epsilon(p) \\ &\iff \forall \epsilon > 0, \text{ for all but finitely many } n, p_n \in N_\epsilon(p) \end{aligned}$$

□

The second line used the fact that for a set of integers, 'all but finitely many' is the same as 'all the sufficiently large'.

Theorem 9.6. *Limits are unique.*

Proof. Suppose $p_n \to p$ and $p_n \to p'$. If $p \neq p'$, then take $\epsilon = \frac{1}{3}d(p, p')$. Note that $N_\epsilon(p)$ and $N_\epsilon(p')$ are disjoint. That $p_n \to p$ implies that all but finitely many of the p_n are in $N_\epsilon(p)$, and likewise for p' and $N_\epsilon(p')$. Since they're disjoint, this is a contradiction. □

Proposition 9.7. *Convergent sequences are bounded.*

Sketch. Say $p_n \to p$. Then only finitely many of the p_n aren't in $N_1(p)$. Those in $N_1(p)$ are certainly bounded, and the finitely many terms which aren't in $N_1(p)$ are bounded (finite collections of numbers are always bounded). The union of bounded things is bounded, so this is bounded. □

Proposition 9.8. *If $E \subset X$ and p is a limit point of E, then there exists a sequence $\{p_n\}$ with terms in E such that $p_n \to p$ in X.*

Proof. Since p is a limit point of E, then within any neighborhood of size $\frac{1}{n}$ lies a point of E. Form a sequence in this way, so that p_n lies in $N_{\frac{1}{n}}(p)$. Then $d(p, p_n) \to 0$ and thus $p_n \to p$. □

Theorem 9.9. *Suppose $\{s_n\}, \{t_n\}$ are sequence in $\mathbb{R}$ or $\mathbb{C}$ with limits s and t, respectively. Then*

- *$s_n + t_n \to s + t$*
- *$cs_n \to cs$ and $s_n + c \to s + c$*
- *$s_n t_n \to st$*
- *If $s_n \neq 0$ and $s \neq 0$, then $\frac{1}{s_n} \to \frac{1}{s}$*

Proof.

- Given $\epsilon > 0$, $\exists N_1$ s.t. $\forall n \geq N_1$, $|s_n - s| < \epsilon$. And $\exists N_2$ s.t. $\forall n \geq N_2$, $|t_n - t| < \epsilon$. Then for $n \geq \max(N_1, N_2)$, $|(s_n + t_n) - (s + t)| = |(s_n - s) + (t_n - t)| \leq |s_n - s| + |t_n - t| \leq \epsilon + \epsilon$. We've slightly exceeded the distance ϵ that we're allowed to move. If we had just selected the N_i to restrict s_n and t_n within $\epsilon/2$ of their limits, this would have worked. Many proofs of convergence will be of this general form.
- Exercise

- We have $s_n t_n - st = (s_n - s)(t_n - t) + s(t_n - t) + t(s_n - s)$. Fix $\epsilon > 0$. $\exists N_1$ s.t $\forall n \geq N$, $|s_n - s| < \sqrt{\epsilon}$, and $\exists N_2$ s.t. $\forall n \geq N_2$, $|t_n - t| < \sqrt{\epsilon}$. For $n \geq \max(N_1, N_2)$, $|(s_n - s)(t_n - t)| < \sqrt{\epsilon}\sqrt{\epsilon} < \epsilon$. Hence $(s_n - s)(t_n - t) \to 0$. It's easier to see that $s(t_n - t) + t(s_n - s)$ converges to 0 (they're just scaled sequences which converge to 0). So our original term is the sum of two sequences which converge to 0, and thus it converges to 0.
- Exercise

□

Theorem 9.10. *(1) $\{x_n\} \in \mathbb{R}^k$ converges to $x = (\alpha_1, \ldots, \alpha_k)$ if and only if each coordinate of the x_n correspond to the appropriate α_i.*
(2) If $x_n \to x$, $y_n \to y$ in $\mathbb{R}^k$ and $\beta_n \to \beta$ in $\mathbb{R}$, then $x_n + y_n \to x + y$, $\beta_n x_n \to \beta x$, and $x_n \cdot y_n \to x \cdot y$.

Is there a way to consider whether a sequence 'wants to converge' or 'should converge' without considering its limit? A mathematician called Cauchy answered this question in the 19th century.

Definition 9.11. A sequence $\{p_n\}$ in a metric space X is a Cauchy sequence if $\forall \epsilon > 0$, $\exists N$ such that $\forall m, n \geq N$, $d(p_m, p_n) < \epsilon$.

Definition 9.12. The diameter of a non-empty, bounded subset $E \subset X$ is $\text{diam} E = \sup\{d(p, q) : p, q \in E\}$.

Now given a sequence p_n, take $E_n = \{p_n, p_{n+1}, \ldots\}$. Then the definition of a sequence being Cauchy is equivalent to the condition that the diameters of E_n converge to 0.

Theorem 9.13. *(1) In any metric space, every convergent sequence is Cauchy.*
(2) If X is a compact metric space, and $\{p_n\}$ is a Cauchy sequence in X, then $\{p_n\}$ converges in X.
(3) In $\mathbb{R}^k$, every Cauchy sequence converges.

Proof. (1) If $p_n \to p$ and $\epsilon > 0$, then $\exists N$ such that $\forall n \geq N$, $d(p_n, p) < \frac{\epsilon}{2}$. Then, by the triangle equality, for $m, n \geq N$, $d(p_m, p_n) \leq \epsilon$.
(2) We'll need two results to prove this: first, for bounded $E \subset X$, diam($\overline{E}$) = diam(E). Secondly, if K_n are a sequence of nested, non-empty compact sets and diam$(K_n) \to 0$, then $\cap_{n=1}^{\infty} K_n$ contains exactly one point. To see the first claim, note that given $p, q \in \overline{E}$ and $\epsilon > 0$, $\exists p', q' \in E$ such that $d(p, p') < \epsilon$ and $d(q, q') < \epsilon$. Then $d(p, q) \leq \epsilon + \text{diam} E + \epsilon$. Since ϵ can be made arbitrarily small, it must be that diam($\overline{E}$) = diam(E). To see that the second claim holds, recall that we've already shown that the intersection of the K_n is not empty. But it has arbitrarily small diameter (as its contained in each of the K_n), so it must contain exactly one point.

□

The third result is sometimes called the Cauchy criterion of convergence. We'll pick up this proof next time.

10. LECTURE 10 — FEBRUARY 28, 2019

Recall that a sequence p_n converges to p if eventually the points of p_n stay as close to p as you'd like them to. We also saw the following big theorem last time:

Theorem 10.1. *(1) Every convergent sequence is Cauchy.*
(2) In a compact space, every Cauchy sequence converges.
(3) In $\mathbb{R}^k$, Cauchy sequences converge.

Proof. We proved (1) last time. For (2), let p_n be a Cauchy sequence in a compact space K. Let $E_n = \{p_n, p_{n+1}, \dots\}$, and consider $K \supset \overline{E_1} \supset \overline{E_2} \supset \dots$. This is a decreasing sequence of non-empty compact subsets, so its intersection is non-empty. And since the diameters of the $\overline{E}_n$ approach zero, this intersection contains exactly one point, say p. To see that $p_n \to p$, fix ϵ. We know that $\exists N$ such that diam($\overline{E}_N$) = diam(E_n) $< \epsilon$. Then $\forall n \geq N$, $p_n, p \in \overline{E}_n$. So $d(p, p_n) \leq$diam($\overline{E}_n$) $< \epsilon$.

For (3), first note that Cauchy sequences are bounded (only finitely many terms are not within distance ϵ of an appropriately chosen $p_{n'}$). So the Cauchy sequence lies in a k-cell, which is compact, and we can apply (2). □

Definition 10.2. A metric space X is <u>complete</u> if its Cauchy sequences converge.

Thus, we've shown that compact spaces are complete and that $\mathbb{R}^k$ is complete. On the other hand, $\mathbb{Q}$ is not complete because 1, 1.4, 1.41, 1.414, ... is Cauchy but does not converge (as $\sqrt{2} \notin \mathbb{Q}$).

For a metric space X which fails to be complete, it's possible to build a larger metric space $X^* \supset X$, the <u>completion</u> of X, which is complete. In fact, one can define $\mathbb{R}$ to be the completion of $\mathbb{Q}$.

Definition 10.3. Given a sequence $\{p_n\} \in X$ and a strictly increasing sequence of positive integers $\{n_k\}$, the sequence $\{p_{n_k}\} = p_{n_1}, p_{n_2}, p_{n_3}, \dots$ is called a <u>subsequence</u> of $\{p_n\}$.

If $\{p_{n_k}\} \to p$, we say that p is a subsequential limit of $\{p_n\}$.

Example 10.4. Consider $p_n = (-1)^n$. The subsequence p_{2k} converges to 1, while the subsequence p_{2k+1} converges to -1.

Proposition 10.5. *$p_n \to p \iff$ every subsequence of $\{p_n\}$ converges to p*

Proof. If every subsequence of $\{p_n\}$ converges to p, then the subsequence consisting of all terms converges to p, so p_n converges to p. In the opposite direction, suppose p_n converges to p. Then for any choice of ϵ, there's an N such that $n \geq N$ implies $d(p_n, p) < \epsilon$. Then this N works for any subsequence of p_n, because the Nth term in any subsequence is either the Nth term in the original subsequence or a term of strictly higher index. □

Theorem 10.6. *(1) Every sequence in a compact metric space has a convergent subsequence.*
(2) Every bounded sequence in $\mathbb{R}^k$ has a convergent subsequence

Proof. (1) If $\{p_n\}$ has infinite range, then its range has a limit point p. Now we can construct a subsequence $\{p_{n_k}\}$ such that $d(p_{n_k}, p) < \frac{1}{k}$, and that sequence converges to p (we have to be sure to pick $n_k > n_{k-1}$). If $\{p_n\}$ has finite range, then a value in its range is repeated infinitely many times. Take the subsequence which consists only of that value.

(2) Apply (1) to a k-cell containing $\{p_n\}$.

$\square$

Theorem 10.7. *p is a subsequential limit of $\{p_n\}$ $\iff$ every neighborhood of p contains infinitely many points in $\{p_n\}$.*

Theorem 10.8. *The set of subsequential limits of a sequence $\{p_n\}$ is closed.*

For sequences in $\mathbb{R}$, we've seen that convergence implies boundedness, which implies the existence of a convergent subsequence. None of the reverse directions of these claims is true.

Definition 10.9. A sequence of real numbers $\{s_n\}$ is monotonically increasing if $s_n \leq s_{n+1}$ $\forall n \in \mathbb{N}$, and monotonically decreasing if $s_n \geq s_{n+1}$ $\forall n \in \mathbb{N}$. A sequence is monotonic if it's either monotonically increasing or decreasing.

From here on, we'll take $\{s_n\}$ to mean a sequence of real numbers.

Theorem 10.10 (Monotone Convergence)**.** *A monotone sequence $\{s_n\}$ converges if and only if it is bounded.*

Proof. The range of this sequence has a supremum - call it α. Since $\alpha - \epsilon$ isn't an upper bound of the range, there's some $s_N \geq \alpha - \epsilon$. Since the sequence is monotonic, all $s_{n'}$ for $n' \geq N$ exceed s_N, and are thus within distance ϵ of α. So α is the limit of s_n. The other direction follows from the fact that convergent sequences are always bounded. $\square$

Definition 10.11. We'll say $s_n \to \infty$ if $\forall M \in \mathbb{R}$, $\exists N \in \mathbb{N}$ s.t. $n \geq N \implies s_n > M$. Similarly, $s_n \to -\infty$ if this holds with $s_n < M$.

It's important to note that we still consider s_n satisfying the above conditions to be divergent.

Given $\{s_n\}$ (as always, in $\mathbb{R}$), let E consist of $x \in \mathbb{R} \cup \{\pm\infty\}$ such that there exists a subsequence $s_{n_k} \to x$. Note that E is never empty, because if s_n is bounded it must contain a subsequential limit, by sequential compactness. If s_n isn't bounded, then it has either ∞ or $-\infty$ as a subsequential limit.

Now define $s^* = \sup E =: \limsup s_n$ and $s_* = \inf E = \liminf s_n$. These are sometimes called the upper and lower limits of a sequence, and are meant to capture the idea of a sequence's 'eventual bounds.'

Theorem 10.12. *(1) There exists a subsequence $s_{n_k} \to s^*$.*

(2) If s^* is real (i.e. not $\pm\infty$), then $\forall \epsilon > 0$, $\exists N$ s.t. $\forall n \geq N$, $s_n < s^* + \epsilon$.

And s^* is uniquely characterized by these properties.

Proof Sketch. (1) If E is bounded above, then because it's closed it contains its supremum. If E isn't bounded, $s^* = \infty$ and indeed s_n has ∞ as a subsequential limit.

(2) If this weren't the case, you could construct a subsequence of s_n with limit strictly greater than s^*. The full proof appears in Rudin.

□

Example 10.13. (1) $s_n = (-1)^n \frac{n+1}{n}$ has $\liminf = -1$ and $\limsup = 1$. Note that these values don't bound the sequence but are eventual bounds, up to subtraction/addition by ϵ.
(2) $\lim s_n = s$ if and only if $\limsup s_n = \liminf s_n = s$.
(3) Take s_n to be a sequence enumerating $\mathbb{Q}$. Then every real number is a subsequential limit - this sequence has uncountably many subsequences with distinct limits!

Theorem 10.14. *If* $s_n \leq t_n$ $\forall n$ *(or* $\forall n \geq N$*), then* $\liminf s_n \leq \liminf t_n$ *and* $\limsup s_n \leq \limsup t_n$.

11. Lecture 11 — March 5, 2019

The midterm is next Tuesday - you're allowed to bring a copy of Rudin, but if you're leafing through it to remember definitions you'll probably run out of time. The exam will be a mix of proofs and examples, but you shouldn't expect to have to recreate a page-long proof that we saw in class, because there's just not enough time for that.

Now about sequences:

Theorem 11.1. *The following hold for real sequences:*

(1) For $p \in \mathbb{R}_+$, $\lim_{n\to\infty} \frac{1}{n^p} = 0$
(2) For $p \in \mathbb{R}_+$, $\lim_{n\to\infty} p^{1/n} = 1$
(3) $\lim_{n\to\infty} n^{1/n} = 1$
(4) If $|x| < 1$, $\lim_{n\to\infty} x^n = 0$
(5) If $|x| < 1$, $p \in \mathbb{R}$, *then* $\lim_{n\to\infty} n^p x^n = 0$

We won't be going over this proof, but it appears in Rudin.

Definition 11.2. In $\mathbb{R}, \mathbb{C}, \mathbb{R}^k$, one can associate to a sequence $\{a_n\}$ a new sequence $s_n = \sum_{k=1}^n a_k$ of partial sums of the series $\sum_{n=1}^\infty a_n$. This infinite sum is only a symbol, which may not equal any element in $\mathbb{R}, \mathbb{C}, \mathbb{R}^k$. The limit s of $\{s_n\}$, if it exists, is the sum of the series, and we write $\sum_{n=1}^\infty a_n = s$.

Because it's often difficult to calculate the limit s, abstract convergence criteria for $\{s_n\}$, which don't make use of a known limit, are useful. We've already seen the Cauchy criterion for convergence, which states the convergence in $\mathbb{R}, \mathbb{C}$, and $\mathbb{R}^k$ is equivalent to being Cauchy. Restating this in the language of series, we arrive at the following result.

Proposition 11.3 (Cauchy criterion). $\sum_{i=1}^\infty a_n$ *converges if and only if* $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ *such that* $\forall m \geq n \geq N$, $|\sum_{k=n}^m a_k| \leq \epsilon$.

Proof. $|\sum_{k=n}^m a_k| = |s_m - s_{n-1}|$. Invoke Cauchy criterion for sequence convergence. □

Taking the case $m = n$, we arrive at a necessary condition for series convergence.

Theorem 11.4. *If $\sum_{i=1}^{\infty} a_n$ converges, then $a_n \to 0$.*

Proof. Given $\epsilon > 0$, the Cauchy criterion implies the existence of an N such that $\forall n = n \geq N$, $|a_n| < \epsilon$. This is precisely convergence of a_n to 0. $\square$

Example 11.5. To see that the above condition is necessary but not sufficient for series convergence, consider the series $\sum_{i=1}^{\infty} \frac{1}{n}$. It diverges, despite the fact that $\frac{1}{n} \to 0$.

"The terms need to go to zero, but they need to go to zero in a friendly enough way." - Dr. Auroux. Once again, we'll use a result about sequences to arrive at a result about series for free - this time it'll be monotone convergence (that bounded, monotone sequences converge) rather than the Cauchy criterion.

Theorem 11.6. *A series in $\mathbb{R}$ with $a_n \geq 0$ converges if and only if its partial sums form a bounded sequence.*

Proof. Because $a_n \geq 0$, the sequence of partial sums is monotone. Because they're bounded, monotone convergence guarantees us the existence of a limit. $\square$

Starting now, we'll get a bit lazy and use $\sum a_n$ to mean $\sum_{n=1}^{\infty} a_n$.

Theorem 11.7 (Comparison test)**.** *(1) If $|a_n| \leq c_n$ for all $n \geq N$ and $\sum c_n$ converges, then $\sum a_n$ converges.*
(2) If $a_n \geq d_n \geq 0$ for all $n \geq N$ and $\sum d_n$ diverges, then $\sum a_n$ diverges.

Proof. Under the conditions of (2), if $\sum a_n$ were to converge then $\sum d_n$ would converge by (1), producing contradiction. So (1) $\implies$ (2). To see that (1) holds, note that the Cauchy criterion for $\sum c_n$ implies the Cauchy criterion for $\sum a_n$. In particular,

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

Since the rightmost side becomes arbitrarily small for n, m greater than appropriately large N, so does the leftmost side. Thus, by the Cauchy criterion for series, $\sum a_n$ converges. $\square$

Theorem 11.8. *If $|x| < 1$, then $\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$. If $|x| \geq 1$, then $\sum x^n$ diverges.*

Proof. If $|x| < 1$, $s_n = 1 + x + \cdots + x^n = \frac{1-x^{n+1}}{1-x}$, which converges to $\frac{1}{1-x}$. If $|x| \geq 1$, then x^n does not have limit 0, so the series doesn't converge. $\square$

Theorem 11.9. *$\sum \frac{1}{n^p}$ converges if $p > 1$ and diverges if $p \leq 1$.*

Proof. First a lemma - a series $\sum a_n$ of weakly decreasing, non-negative terms converges if and only if $\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \ldots$ converges. Since $a_n \geq a_{n+1}$, we have that $\sum_{n=1}^{2^m-1} a_n \leq \sum_{k=0}^{m} 2^k a_{2^k}$. So if the new, weird sequence converges, the original sequence does as well, because its partial sums are smaller and, by monotone convergence, convergence is equivalent to bounded partial sums. In the other direction, suppose that the original sequence converges and note that $a_1 + a_2 + \cdots + a_{2^k} \geq \frac{1}{2}a_1 + a_2 + 2a_4 + 4a_3 +$

$\cdots + 2^{k-1}a_{2^k}$. The left hand side is a partial sum of the original sequence and the right hand side is $\frac{1}{2}$ of a partial sum of the new, weird sequence. So if the new weird, sequence were unbounded, the original sequence would be as well. We conclude that the weird sequence converges, as desired.

Now we can begin the main proof. If $p \leq 0$, then $\frac{1}{n^p}$ doesn't converge to 0, so the sum diverges. Suppose $p > 0$ - applying the lemma with $a_n = \frac{1}{n^p}$, we see that $2^k a_{2^k} = 2^k \frac{1}{2^{kp}} = \frac{1}{2^{k(p-1)}}$. This is a geometric series, and it converges if and only if $|\frac{1}{2^{p-1}}| < 1$, which happens iff $p > 1$. $\square$

Theorem 11.10. $\sum \frac{1}{n(\log n)^p}$ *converges if and only if* $p > 1$.

The proof uses our previous lemma about the sequence $\sum 2^k a_{2^k}$.

Definition 11.11. $e = \sum_{n=0}^{\infty} \frac{1}{n!}$

Theorem 11.12. $\lim_{n\to\infty}(1 + \frac{1}{n})^n = e$

A careful proof of this theorem appears in Rudin, but it comes down to lots of technicalities with the binomial formula and bounds - it's not very enlightening.

Theorem 11.13. *e is not rational. In fact, it's not even algebraic.*

12. LECTURE 12 — MARCH 7, 2019

Today we'll talk more about series - this stuff won't appear on the midterm, but it will appear on the final (and it's pretty cool).

Theorem 12.1 (Root test)**.** *Given* $\sum a_n$*, let* $\alpha = \limsup \sqrt[n]{|a_n|}$*. Then*

(1) If $\alpha < 1$*, then* $\sum a_n$ *converges.*
(2) If $\alpha > 1$*, then* $\sum a_n$ *diverges.*
(3) If $\alpha = 1$*, the test is inconclusive.*

Proof. If $\alpha < 1$, take $\alpha < \beta < 1$. Since β is strictly greater than all the subsequential limits of $\sqrt[n]{|a_n|}$, $\exists N$ s.t. $\sqrt[n]{|a_n|} \leq \beta$ for $n \geq N$. Otherwise, one could construct a subsequence with limit at least β. So for $n \geq N$, $|\alpha_n| \leq \beta^n$, and β^n is a convergent geometric series, so $\sum \alpha_n$ converges by comparison.

If $\alpha > 1$, take $\alpha > \beta \geq 1$. By definition of α, there are infinitely many terms in $\sqrt[n]{|a_n|}$ which exceed β. These terms have $|a_n| \geq \beta^n$. So the series doesn't converge (formally, because the a_n don't converge to 0). $\square$

Theorem 12.2 (Ratio test)**.** *Fix a series* $\sum a_n$*:*

(1) If $\limsup |\frac{a_{n+1}}{a_n}| < 1$*, then* $\sum a_n$ *converges.*
(2) If $|\frac{a_{n+1}}{a_n}| \geq 1$ *for all* $n \geq N$*, then* $\sum a_n$ *diverges.*

Proof. For (1), take $\limsup |\frac{a_{n+1}}{a_n}| < \beta < 1$, then $\exists N$ such that $\forall n \geq N$, $|\frac{a_{n+1}}{a_n}| \leq \beta$. Then $|a_n| \leq c\beta^n$ for some c. Since $\beta < 1$, that series converges and $\sum a_n$ converges by comparison. For (2), $|\frac{a_{n+1}}{a_n}| \geq 1$ after N, so $|a_N| \leq |a_{N+1}| \leq \dots$, so the a_n don't even converge to 0. $\square$

It turns out that the root test is stronger than the ratio test. The reason why is that for any sequence of positive real numbers c_n, $\limsup \sqrt[n]{c_n} \leq \limsup \frac{c_{n+1}}{c_n}$. The reason why we use the ratio test is that it's often easier to compute $\frac{c_{n+1}}{c_n}$ than $\sqrt[n]{c_n}$.

Example 12.3. Consider $\sum a_n$ where $a_n = \begin{cases} \frac{1}{2^n} & n \text{ even} \\ \frac{1}{3^n} & n \text{ odd} \end{cases}$. This series converges because it's less than $\sum \frac{1}{2^n}$. The root test detects this, because $\sqrt[n]{|a_n|} \in \{\frac{1}{2}, \frac{1}{3}\}$. The ratio test, however, fails to detect this, because $\frac{a_{n+1}}{a_n}$ exceeds 1 at even terms, where $\frac{3^n}{2^{n+1}} = \frac{1}{2}(\frac{3}{2})^n$. And $\limsup \frac{a_{n+1}}{a_n} = \infty$, as $\frac{1}{2}(\frac{3}{2})^n$ grows arbitrarily large.

Given a sequence $\{c_n\}$ of complex numbers, $\sum c_n z^n = c_0 + c_1 z + c_2 z^2 + \ldots$ forms a power series. This is a fairly natural generalization of the polynomial, but whether it actually makes sense as a quantity depends on the convergence of the series. For now, we'll think of $z \in \mathbb{C}$ as a number, but later in the course we'll think of it as a variable and consider the differentiability of these functions.

Theorem 12.4. *Let $\alpha = \limsup \sqrt[n]{|c_n|}$, and let $R = \frac{1}{\alpha}$[1]. Then $\sum c_n z^n$ converges if $|z| < R$ and diverges if $|z| > R$. R is referred to as the radius of convergence of $\sum c_n z^n$.*

Proof. Let $a_n = c_n z^n$, and apply the root test, noting that $\sqrt[n]{|a_n|} = \sqrt[n]{|c_n|}|z|$ and $\limsup \sqrt[n]{|a_n|} = \frac{|z|}{R}$. □

Notice that we haven't said anything about what happens when $|z| = R$. In that case, it's difficult to say anything without considering the particular power series at hand.

Example 12.5. $\sum z_n$ has $c_n = 1\ \forall n$, so $R = 1$. On the other hand, $\sum n^n z^n$ has $c_n = n^n$; since $\sqrt[n]{|c_n|} = n \to \infty$, $R = 0$. And $\sum \frac{1}{n!} z^n = e^z$ has $R = \infty$.

An alternating series is one whose terms have alternating signs. More explicitly, either all its odd terms are positive and its even terms are negative, or vice versa. An example is $\sum (-1)^n a_n$ where $a_n > 0; \forall n$.

Theorem 12.6. *Suppose $\{a_n\} \in R$ is an alternating series where $|a_1| \geq |a_2| \geq |a_3| \geq \ldots$ and $a_n \to 0$. Then $\sum a_n$ converges.*

Proof. Let $s_n = \sum_{k=1}^{n} a_k$. Then, because the sequence alternates and $|a_{k+1}| \geq |a_k|$,

$$s_2 \leq s_4 \leq s_6 \cdots \leq s_5 \leq s_3 \leq s_1$$

So s_{2m} and s_{2m+1} are monotonic, bounded sequences, meaning they converge. They converge to the same thing because $s_{2m+1} - s_{2m} = a_{2m+1} \to 0$. □

The above theorem is pretty remarkable, because it's a rare case in which convergence is not dependent upon the rate at which the terms of the series converge to 0.

[1] If $\alpha = \infty$, we define $\frac{1}{\alpha} = 0$. Similarly, if $\alpha = 0$, we define $\frac{1}{\alpha} = \infty$

Definition 12.7. $\sum a_n$ converges absolutely if $\sum |a_n|$ converges.

Proposition 12.8. *Absolute convergence implies convergence.*

Proof. Use the Cauchy criterion. $|\sum_{k=n}^{m} a_k| \leq \sum_{k=n}^{m} |a_k|$, and the right hand side gets arbitrarily small for sufficiently large n because $\sum |a_n|$ converges. $\square$

The rest of the class will be dedicated to operations on series which are seemingly safe but in fact require the condition of absolute convergence in order to be safe.

Theorem 12.9. *If $\sum a_n = A$ and $\sum b_n = B$, then their sum $\sum(a_n + b_n) = A + B$.*

Proof. Look at the partial sums. $\sum_{k=1}^{n}(a_k + b_k) = \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k$. We learned a few weeks ago that the limit of the sum is the sum of the limits, so this converges to $A + B$. $\square$

So defining addition of series is not so hard, and is as well-behaved as we'd like it to be[2]. Defining multiplication is much less obvious, however. Consider

$$(a_0 + a_1 + a_2 + \dots)(b_0 + b_1 + b_2 + \dots)$$

One way to make sure all terms hit each other is to group them by the sum of their indices, and define the sum like so:

$$a_0 b_0 + (a_1 b_0 + a_0 b_1) + (a_2 b_0 + a_1 b_1 + a_0 b_2) + \dots$$

Definition 12.10. The product of $\sum a_n$ and $\sum b_n$ is $\sum c_n$ where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$.

We've defined a product but this doesn't mean anything yet, as we don't know anything about the behavior of this operation on series. Unfortunately, it turns out that it does not in general send convergent series to convergent series.

> **Example 12.11.** By our theorem for alternating series, $\sum \frac{(-1)^n}{\sqrt{n+1}}$ converges. Its product with itself, however, is a sequence of positive terms which does not converge. One can check that $|c_n|$ does not converge to 0, and is in fact always at least 2.

Fortunately, things are nicer with the assumption of absolute convergence, though we aren't going to prove why right now.

Theorem 12.12. *If $\sum a_n = A$ converges absolutely and $\sum b_n = B$ converges, then their product converges to AB.*

Definition 12.13. Let $\{n_k\}$ be a sequence of positive integers in which every positive integer appears exactly once. Then the series $\sum_{k=1}^{\infty} a_{n_k}$ is a rearrangement of the series $\sum_{k=1}^{\infty} a_k$.

Theorem 12.14 (Riemann). *Let $\sum a_n$ be a series of real numbers which converges but does not converge absolutely. Then for any $\alpha, \beta \in \mathbb{R}$ with $\alpha \leq \beta$, there exists a rearrangement $\sum a'_n$ whose partial sums s'_n satisfy $\liminf s'_n = \alpha_1$, $\limsup s'_n = \beta$.*

[2]Multiplying series by constants is also not so hard, as $\sum c a_n = c \sum a_n$ when $\sum a_n$ converges.

This is an insane theorem, and shows that rearrangements are in general not at all safe. In particular, they can be used to warp the upper and lower limits of any convergent but not absolutely convergent series of real numbers to anything you want. Fortunately, there's something of an antithesis to this theorem.

Theorem 12.15. *If $\sum a_n$ converges absolutely to A, then all of its rearrangements converge to A.*

13. Lecture 13 — March 14, 2019

:(

14. Lecture 14 — March 26, 2019

Guest lecture by Dr. Williams!

Definition 14.1. For X, Y metric spaces, $f : X \to Y$ is continuous at $p \in X$ if $\forall \epsilon > 0$, $\exists \delta > 0$ such that $\forall x \in X$, $\rho_X(x, p) < \delta \implies \rho_Y(f(x), f(p)) < \epsilon$.

We say that f is continuous if it's continuous at all points in its domain. Intuitively, this mean that we can restrict output by restricting input.

Remark 14.2. If p is not an isolated point, continuity of f at p is equivalent to the statement $\lim_{x\ top} f(x) = f(p)$.

Theorem 14.3. *$f : X \to Y$ is continuous if and only if for all open $V \subseteq Y$, $f^{-1}(V) \subseteq X$ is open.*

Theorem 14.4 (Main Theorem)**.** *If $f : X \to Y$ is continuous and X is compact, then $f(X)$ is compact.*

Proof. Let $\{V_\alpha\}$ be an open cover of $f(X)$, meaning $\cup_\alpha V_\alpha \supseteq f(X)$. Since f is continuous, the $f^{-1}(V_\alpha)$ are open, and since anything in x has image in one of the v_α, the $f^{-1}(V_\alpha)$ cover X. Since X is compact, this reduces to a finite subcover $f^{-1}(V_{\alpha_1}), \dots, f^{-1}(V_{\alpha_n})$. Then $V_{\alpha_1}, \dots, V_{\alpha_n}$ form a finite subcover of $f(X)$, as desired. □

We'll see that this is really a generalization of the extreme value theorem, and it'll be quite useful for the rest of the lecture. Let's examine some of its corollaries.

Corollary 14.5. *If $f : X \to Y$ is continuous and X is compact, then $f(X)$ is closed and bounded.*

Corollary 14.6. *If $F : X \to \mathbb{R}$ is continuous and X is compact, then $\exists p \in X$ such that $f(p) = \sup_{x \in X} f(x)$ and $\exists q \in X$ such that $f(q) = \inf_{x \in X} f(x)$. In words, f achieves maximal/minimal values.*

Proof. By the main theorem, $f(X) \subset \mathbb{R}$ is compact, so it's closed and bounded. Since it's bounded, $\sup f(x)$ and $\inf f(x)$ exist in $\mathbb{R}$, and since it's closed, these values are elements of $f(X)$, as desired. □

When $X = [a, b] \subset \mathbb{R}$, this is precisely the extreme value theorem.

Theorem 14.7. *If $f : X \to Y$ is a continuous bijection and X is compact, then $f^{-1} : Y \to X$ is also continuous.*

Proof. By our characterization of continuity, we need to show that f sends open set to sets (meaning pre-images of open sets under f^{-1} are open). Given $V \subseteq X$ open, $V^c \subseteq X$ is a closed subset of a compact set, so it's compact. By our main theorem, $f(V^c)$ must also be compact in X. Since f is a bijection, $f(V^c) = f(V)^c$. So $f(V)^c$ is compact, which means it's closed. Thus $f(V)$ is open, as desired. $\square$

Definition 14.8. $f : X \to Y$ is <u>uniformly continuous</u> if $\forall \epsilon > 0$, $\exists \delta > 0$ s.t. $\forall x, p \in X$, $\rho_x(x, p) < \delta \implies \rho_Y(f(x), f(p)) < \epsilon$.

The crucial difference between uniform continuity and continuity is that continuity allows for δ to be selected as a function of ϵ and p, whereas uniform continuity only allows δ to be selected as a function of ϵ. As we'll soon see, continuous functions may not permit choices of δ which sufficiently restrict output across the entirety of their domains, meaning uniform continuity is stronger than continuity.

Example 14.9. $f : \mathbb{R}_+ \to \mathbb{R}_+$, $f(x) = \frac{1}{x}$ is continuous but not uniformly continuous. To see why, take $\epsilon = \frac{1}{2}$. Regardless of how small $\delta > 0$ is selected, we can find sufficiently large $n \in \mathbb{N}$ such that $|\frac{1}{n} - \frac{1}{n+1}| < \delta$ but $|f(\frac{1}{n}) - f(\frac{1}{n+1})| = 1 > \epsilon$.

Theorem 14.10. *If $f : X \to Y$ is continuous and X is compact, then f is in fact uniformly continuous.*

Proof. Fix $\epsilon > 0$. Since f is continuous, for all $p \in X$, $\exists \delta(p)$ such that $\rho_X(x, p) < \delta(p) \implies \rho_Y(f(x), f(p)) < \epsilon/2$ for $x \in X$. Let $V_p = N_{\delta(p)/2}(p)$. The V_p collectively cover X, so they reduce to a finite subcover $V_{p_1}, \ldots, V_{p_n}$. Now let $\delta = \frac{1}{2}\min(\delta(p_1), \ldots, \delta(p_n))$. Now consider p, x with $\rho_X(p, x) < \delta$. Then there's a V_{p_i} with $\rho_X(p, p_i) < \delta(p_i)/2$. Since $\delta < \delta(p_i)/2$, by the triangle inequality we have $\delta_X(x, p_i) < \delta(p_i)$. Then, by definition of $\delta(p_i)$, $\rho_Y(f(p), f(p_i) < \epsilon/2$ and $\rho_Y(f(x), f(p_i)) < \epsilon/2$. So, again by the triangle inequality, $d_Y(f(x), f(p)) < \epsilon$. Thus δ restricts the behavior of X on its entire domain, and f is uniformly continuous. $\square$

Theorem 14.11. *If $E \subseteq \mathbb{R}$ is not compact, then*

(a) $\exists f : E \to \mathbb{R}$ which is continuous such that $f(E)$ is not bounded.
(b) $\exists f : E \to R$ which is continuous and bounded, but which has no maximum.
(c) If E is also bounded, $\exists f : E \to \mathbb{R}$ which is continuous but not uniformly continuous.

Finally, here's a theorem we'll see next time:

Theorem 14.12. *If $f : X \to Y$ is continuous and $E \subseteq X$ is connected, then $f(E)$ is connected.*

15. Lecture 15 — March 28, 2019

We'll start with a theorem we saw last time, which we weren't able to prove.

Theorem 15.1. *If $f : X \to Y$ is continuous and $E \subseteq X$ is connected, then $f(E)$ is connected.*

Proof. Suppose $f(E)$ is not connected, meaning it can be witnessed as the union of non-empty, separated sets $f(E) = A \cup B$. Now consider $G = E \cap f^{-1}(A)$ and $H = E \cap f^{-1}(B)$. These sets are disjoint because pre-images preserve disjointness (i.e. there could not be an $x \in E$ with $f(x) \in A$ and $f(x) \in B$). It remains to show that $\overline{G} \cap H = \emptyset$ and $G \cap \overline{H} = \emptyset$. First, we claim $f(\overline{G}) \subseteq \overline{A}$. Any $x \in \overline{G}$ appears as the limit of a sequence in G x_n. By continuity of f, $f(x_n) \to f(x)$, so $f(x)$ is the limit of a sequence in A, so it's in $\overline{A}$. So $\overline{G} \cap H = \emptyset$, since $f(\overline{G}) \subseteq \overline{A}$, $f(H) \subseteq B$, and $\overline{A} \cap B = \emptyset$. By symmetry, $G \cap \overline{H} = \emptyset$, so we've demonstrated that E is disconnected, producing contradiction. □

"Did I just prove a homework problem? Being a mathematician is less dangerous than being a surgeon or a pilot, but there are some occupational risks." - Dr. Auroux.

Given this theorem, we get the Intermediate Value theorem more or less for free.

Theorem 15.2 (Intermediate Value Theorem)**.** *If $f : [a, b] \to \mathbb{R}$ is continuous, then $\forall c \in \mathbb{R}$ such that $f(a) < c < f(b)$, c lies in the image of f.*

Proof. We've seen that $E \subseteq R$ is connected if and only if $x, y \in E$ means $z \in E$ for any $x < y < z$. So $[a, b]$ is connected, and by the previous theorem, $f([a, b])$ is connected. By this characterization of connected sets, the result holds. □

In previous classes you may have heard or said something along the lines of 'as x goes to infinity, $f(x)$ goes to infinity'. We don't currently have the tools to formalize this idea, because infinity can't live in a metric space. By further developing limits of functions and defining neighborhoods in the extended reals, we can handle these cases.

We've seen that $\lim_{x\to p} f(x) = q$ if and only if $\forall \epsilon > 0$, $\exists \delta > 0$ such that $0 < |x - p| < \delta$ implies $|f(x) - q| < \epsilon$. Equivalently, for every sequence x_n converging to p with $x_n \neq p$, $f(x_n) \to q$. This is also equivalent to saying that for every neighborhood U of q, there exists a neighborhood V of p with $x \in V \implies f(x) \in U$.

In the extended real number system $\mathbb{R} \cup \{-\infty, \infty\}$, declare neighborhoods of ∞ to be intervals (c, ∞) (and neighborhoods of $-\infty$ to be intervals $(-\infty, -c)$).

Example 15.3. With the neighborhood-based definition of functional limits, $\lim_{x\to\infty} f(x) = \infty$ means that for any $C > 0$ (we're taking the neighborhood (C, ∞)), there exists $A > 0$ (we're taking the neighborhood (A, ∞)) such that $f(x) > C$ when $x > A$. Equivalently, for any $x_n \to \infty$, $f(x_n) \to \infty$.

Now we'll consider one-sided limits, which you also may have seen previously in a calculus class.

Definition 15.4. $\lim_{p^+} f(x) = q \iff \forall \epsilon > 0\ \exists \delta > 0$ such that $p < x < p + \delta$ means $|f(x) - q| < \epsilon$. Equivalently, for any $x_n \to p$ with $x_n > p$, $f(x_n) \to q$.

The definition of $\lim_{x\to p^-} f(x)$ is analogous. When these limits exist, Rudin writes them as $f(p^+)$ and $f(p^-)$, respectively.

Proposition 15.5. $\lim_{x\to p} f(x) = q \iff f(p^-) = f(p^+) = q$.

Proof sketch. The forward direction is immediate, and the backward direction involves taking the minimum of the δ's you get from the definitions of $f(p^-)$ and $f(p^+)$. □

It's important to note that it's possible that $f(p^-) = f(p^+) = q$ but $f(p) \neq q$.

Definition 15.6. We say $f : \mathbb{R} \to \mathbb{R}$ has a simple discontinuity at p if it is not continuous at p but $f(p^-)$ and $f(p^+)$ exist. This is also called a discontinuity of first kind, while all other discontinuities are called discontinuities of the second kind.

Example 15.7. $f(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$ has a simple discontinuity at 0.

$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & \text{else} \end{cases}$ has discontinuities of the second kind at every point, as neither of the one-sided limits exist.

Definition 15.8. For $E \subseteq \mathbb{R}$, $f : E \to \mathbb{R}$ is monotonically increasing if $x < y \implies f(x) \leq f(y)$, and monotonically decreasing if $x < y \implies f(x) \geq (y)$. A monotonic function is either monotonically increasing or monotonically decreasing.

Theorem 15.9. *If f is monotonically increasing on (a, b), then $\forall x \in (a, b)$, $f(x^-)$ and $f(x^+)$ exist. In fact,* $\sup_{t \in (a,x)} f(t) = f(x^-) \leq f(x) \leq f(x^+) = \inf_{t \in (x,b)} f(t)$.

Proof. $\{f(t) | t \in (a, x)\} \subseteq \mathbb{R}$ is non-empty and bounded above by $f(x)$, so it has a least upper bound A. We'd like to show $f(x^-) = \lim_{t \to x^-} f(t) = A$. Fix $\epsilon > 0$. Since $A - \epsilon$ is not a least upper bound for $\{f(t) | t \in (a, x)\}$, $\exists \delta > 0$ such that $x - \delta \in (a, x)$ and $f(x - \delta) > A - \epsilon$. Now, for $t \in (x - \delta, x)$, $A - \epsilon < f(x - \delta) \leq f(t) \leq A$. So $|f(t) - A| < \epsilon$ and $\lim_{t \to x^-} f(t) = A$. By an almost identical argument, the statement for one-sided limits from the right holds. $\square$

So the discontinuities of monotonic functions are fairly reasonable.

Corollary 15.10. *A monotonic function has at most countably many discontinuities.*

Proof. If f is monotonically increasing, then a discontinuity at x means $f(x^-) < f(x^+)$. There exists a rational in $(f(x^-), f(x^+))$, and there are only countably many rationals, so there can only be countably many of these jumps. Likewise if f is monotonically decreasing. $\square$

One example of a monotonic function which realizes infinitely many discontinuities is the function $\lceil x \rceil$, which outputs the smallest integer greater than its input and is discontinuous at each integer.

16. Lecture 16 — April 2, 2019

Definition 16.1. The derivative of $f : [a, b] \to \mathbb{R}$ at $x \in [a, b]$ is $\lim_{t \to x} \frac{f(t) - f(x)}{t - x}$, if it exists.

When the above value exists, we write it as $f'(x)$, and say that the function f is differentiable at x.

Theorem 16.2. *If f is differentiable at x, then it is continuous at x.*

Proof. To show that $\lim_{t\to x} f(t) = f(x)$ amounts to proving that $\lim_{t\to x} f(t) - f(x) = 0$. We have

$$\begin{aligned}\lim_{t\to x} f(t) - f(x) &= \lim_{t\to x} \frac{f(t)-f(x)}{t-x}(t-x) \\ &= f'(x)\lim_{t\to x}(t-x) \\ &= f'(x)\cdot 0 \\ &= 0\end{aligned}$$

$\square$

Theorem 16.3. *If $f, g : [a,b] \to R$ are differentiable at x, then so are $f+g$, fg, and (provided $g(x) \neq 0$), f/g. Moreover,*

1. $(f+g)'(x) = f'(x) + g'(x)$
2. $(fg)'(x) = f'(x)g(x) + g'(x)f(x)$
3. $(f/g)'(x) = \frac{g(x)f'(x)-f(x)g'(x)}{g(x)^2}$

Proof. The first claim follows from the fact that the limit of a sum is the sum of limits - formally, $\lim(s_n + t_n) = \lim s_n + \lim t_n$, where $s_n = \frac{f(t)-f(x)}{t-x}$ and $t_n = \frac{g(t)-g(x)}{t-x}$.

To prove the second claim, we creatively add zero.

$$\begin{aligned}(fg)'(x) &= \lim_{t\to x} \frac{f(t)g(t)-f(x)g(x)}{t-x} \\ &= \lim_{t\to x} \frac{f(t)g(t)-f(t)g(x)+f(t)g(x)-f(x)g(x)}{t-x} \\ &= \lim_{t\to x} f(t)\frac{g(t)-g(x)}{t-x} + g(x)\frac{f(t)-f(x)}{t-x} \\ &= f(x)g'(x) + g(x)f'(x)\end{aligned}$$

Note that we used continuity of f - which followed from its differentiability - to conclude that $\lim_{t\to x} f(t) = f(x)$. We won't prove the claim for f/g here. $\square$

Theorem 16.4 (Chain rule)**.** *Suppose f is continuous on $[a,b]$ and differentiable at $x \in [a,b]$, and g is defined on an interval containing $f([a,b])$ and differentiable at $f(x)$. Then $h(t) = g \circ f(t)$ is defined on $[a,b]$ and differentiable at x, with $h'(x) = g'(f(x))f'(x)$.*

Proof. Write $f(t) - f(x) = (t-x)(f'(x) + u(t))$ for $u(t)$ an error term with limit 0 as $t \to x$. Likewise, taking $y = f(x)$ for ease of notation, write $g(s) - g(y) = (s-y)(g'(y) + v(s))$, for $v(s)$ an error determ with limit 0 as $s \to y$. Then

$$\begin{aligned}g(f(t)) - g(f(x)) &= (f(t)-f(x))(g'(f(x)) + v(f(t))) \\ g(f(t)) - g(f(x)) &= (t-x)(f'(x)+u(t))(g'(f(x)) + v(f(t))) \\ \frac{g(f(t)) - g(f(x))}{t-x} &= (f'(x)+u(t))(g'(f(x)) + v(f(t)))\end{aligned}$$

Taking the limit as $t \to x$ proves the claim. $\square$

Example 16.5. Consider $f(x) = \begin{cases} x\sin(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$. f is continuous at 0, as $|f(x) - f(0)| = |x\sin(\frac{1}{x})| \leq |x|$, which approaches 0 as x approaches 0. It's easier to see that it's continuous on $\mathbb{R} \setminus \{0\}$, using the fact that products, quotients, and compositions of continuous functions are continuous. One can also see that f is differentiable on $\mathbb{R} \setminus \{0\}$, but it fails to be differentiable at 0, as $\frac{f(x)-f(0)}{x-0} = \frac{x\sin(\frac{1}{x})}{x} = \sin(\frac{1}{x})$, which does not have a limit as $x \to 0$.

The following theorems will be quite useful for the remainder of the course. First, a familiar definition.

Definition 16.6. A function f has a <u>local maximum</u> at p if $\exists \delta > 0$ such that $|x-p| < \delta \implies f(x) \leq f(p)$.

Theorem 16.7. *If $f : [a,b] \to \mathbb{R}$ has a local maximum at $x \in (a,b)$ and is differentiable at x, then $f'(x) = 0$.*

Proof. Consider approaching x from the right and left side (note that we're making use of the fact that x is in the interior of f's domain). By assumption, $\lim_{t\to x} \frac{f(t)-f(x)}{t-x}$ exists. When $t - x > 0$, then $\frac{f(t)-f(x)}{t-x} \leq 0$, as $f(t) - f(x) \leq 0$. Similarly, when $t - x < 0$, $\frac{f(t)-f(x)}{t-x} \geq 0$. It follows that the limit must be zero. $\square$

Theorem 16.8 (Mean Value). *Let $f, g : [a,b] \to \mathbb{R}$ be differentiable on (a,b). Then $\exists x \in (a,b)$ such that $(f(b) - f(a))g'(x) = f'(x)(g(b) - g(a))$.*

Proof. Let $h(t) = (f(b) - f(a))g(t) - f(t)(g(b) - g(a))$. Then h is continuous on $[a,b]$ and differentiable on (a,b). The problem reduces to proving that $h'(t) = 0$ for some $t \in (a,b)$. Note that

$$h(a) = f(b)g(a) - f(a)g(b) = h(b)$$

If h is constant, then its derivative is everywhere zero, and the claim follows. If h is not constant, then - by the extreme value theorem - it reaches a maximum or minimum at an interior point t. By the previous theorem, $h'(t) = 0$, proving the claim. $\square$

Corollary 16.9. *The previous statement of the Mean Value theorem may appear foreign, but it implies the more familiar one. In particular, taking g to be the identity proves the existence of an $x \in (a,b)$ for which $f(b) - f(a) = (b-a)f'(x)$.*

Theorem 16.10. *Let f be a real-valued function differentiable on (a,b).*

1. *If $f'(x) \geq 0\ \forall x \in (a,b)$, then f is monotonically increasing on (a,b).*
2. *If $f'(x) \leq 0\ \forall x \in (a,b)$, then f is monotonically decreasing on (a,b).*
3. *If $f'(x) = 0\ \forall x \in (a,b)$, then f is constant on (a,b).*

Proof. Suppose we are in case 1, and fix $x, y \in (a,b)$ with $x < y$. Then, by the Mean Value theorem, $f(y) - f(x) = f'(t)(y-x)$. The right hand side is the product of two nonnegative numbers, so it's nonnegative. Then $f(y) - f(x) \geq 0$ and $f(y) \geq f(x)$, as desired. The remaining cases follow similarly. $\square$

17. Lecture 17 — April 4, 2019

Last time we looked at the Mean value theorem, which states that the mean value of a function's rate of change is achieves somewhere. In particular, for $f : [a, b] \to \mathbb{R}$, there exists an $x \in (a, b)$ such that $f'(x) = \frac{f(b)-f(a)}{b-a}$. The generalization is that for $f, g : [a, b] \to \mathbb{R}$, there exists an $x \in (a, b)$ with $\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(x)}{g'(x)}$.

Theorem 17.1 (L'Hopital's rule)**.** *Let $f, g : (a, b) \to \mathbb{R}$ be differentiable, $g'(x) \neq 0\ \forall x$, and suppose $\lim_{x\to a} \frac{f'(x)}{g'(x)} = A$, for $A \in \mathbb{R} \cup \{\pm\infty\}$. If either*

(1) $f(x), g(x) \to 0$ as $x \to a$, or
(2) $g(x) \to \infty$ as $x \to a$

then $\frac{f(x)}{g(x)} \to A$ as $x \to a$. Likewise for b.

Proof. To show $\frac{f(x)}{g(x)} \to A$ as $x \to a^+$, we show

(A) $\forall q \in \mathbb{R}$ s.t. $A < q$, $\exists c \in (a, b)$ s.t. $x \in (a, c) \implies \frac{f(x)}{g(x)} \leq q$.
(B) $\forall h \in \mathbb{R}$ s.t. $A > h$, $\exists c' \in (a, b)$ s.t. $x \in (a, c') \implies \frac{f(x)}{g(x)} \geq q$.

First suppose we obey (1). Then, because $\frac{f'(x)}{g'(x)} \to A$ as $x \to a$, $\exists c \in (a, b)$ s.t. $x \in (a, b) \implies \frac{f'(x)}{g'(x)} < q$. Then for $a < y < x < c$, the generalized MVT provides the existence of $t \in (y, x)$ such that $\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(t)}{g'(t)} < q$. Keeping x fixed, as $y \to a$, $f(y) \to 0$ and $g(y) \to 0$. Then $\lim \frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f(x)}{g(x)} \leq q$. So we've shown (A).

If we obey (2), then again $\exists c$ with $x \in (a, b) \implies \frac{f'(x)}{g'(x)} < q$. Again by the generalized MVT, $a < x < c \implies \exists t \in (x, c)$ s.t. $\frac{f(x)-f(c)}{g(x)-g(c)} = \frac{f'(t)}{g'(t)} < q$. So, as $x \to a^+$, because $g(x) \to \infty$ and $g(c)$ is a constant, $\frac{f(c)}{g(x)-g(c)} \to 0$ and $\frac{g(x)-g(c)}{g(x)} \to 1$. So

$$\frac{f(x)}{g(x)} = \left(\frac{f(x) - f(c)}{g(x) - g(c)} + \frac{f(c)}{g(x) - g(c)}\right)\frac{g(x) - g(c)}{g(x)}$$

Where we've shown that the rightmost term approaches 1, the middle term approaches 0, and the leftmost term is bounded by q. If we had done this with some $q' < q$, then we could have concluded that $\exists c' \in (a, c)$ s.t. $a < x < c' \implies \frac{f(x)}{g(x)} < q$. So we've proven (A) in both cases, and (B) follows similarly. $\square$

Theorem 17.2 (Taylor's theorem)**.** *For $f : [a, b] \to \mathbb{R}$ and $n \geq 1$, suppose $f^{(n-1)}$ is continuous on $[a, b]$ and $f^{(n)}$ exists on (a, b). Let α, β be distinct in $[a, b]$. The $(n-1)$th Taylor polynomial of f at α is*

$$P(t) = f(\alpha) + f'(\alpha)(t - \alpha) + \frac{f''(\alpha)}{2}(t - \alpha)^2 + \cdots + \frac{f^{(n-1)}}{(n-1)!}(t - \alpha)^{n-1}$$

And there exists $x \in (\alpha, \beta)$ with $f(\beta) = P(\beta) + \frac{f^{(n)}(x)}{n!}(\beta - \alpha)^n$.

Proof. Let M be the constant such that $f(\beta) = P(\beta) + M(\beta - \alpha)^n$, and let $g(t) = f(t) - P(t) - M(t-\alpha)^n$. Then $g^{(n)}(t) = f^{(n)}(t) - n!M$. We'd like to show $\exists x$ with $g^{(n)}(x) = 0$, which would imply $M = \frac{f^{(n)}(x)}{n!}$. Suppose, without loss of generality, that $\alpha < \beta$. Since P has the same derivative as f, $g(\alpha)$ and the first $n-1$ derivatives of g at α are zero. We also have that $g(\beta) = 0$. Then, by MVT, $g'(x_1) = 0$ for some $x_1 \in (\alpha, \beta)$. Again using the MVT (with $g'(\alpha) = g'(x_1) = 0$), we have that $g''(x_2) = 0$ for some $x_2 \in (\alpha, x_1)$. Proceeding in this way, we arrive at the existence of an x_n with $g^{(n)}(x_n) = 0$. ☐

This statement of Taylor's theorem is nice, because the $\frac{f^{(n)}(x)}{n!}(\beta-\alpha)^n$ allows us to bound our errors, by considering the $x \in (\alpha, \beta)$ with greatest nth derivative. So we're done with derivatives, and we're off to Riemann integrals.

Definition 17.3. A <u>partition</u> P of $[a,b] \subseteq \mathbb{R}$ is a finite set $x_0, \ldots, x_n \in \mathbb{R}$ such that $a = x_0 \leq x_1 \leq \cdots \leq x_n = b$. We write $\Delta x_i = x_i - x_{i-1}$, $i = 1, \ldots, n$.

Given a bounded function $f : [a,b] \to R$ and a partition $P = \{x_0, \ldots, x_n\}$, let $M_i = \sup\{f(x), x_{i-1} \leq x \leq x_i\}$ and $m_i = \inf\{f(x), x_{i-1} \leq x \leq x_i\}$. Set $U(P,f) = \sum_{i=1}^n M_i \Delta x_i$ and $L(P,f) = \sum_{i=1}^n m_i \Delta x_i$. Then the upper Riemann integral is $\overline{\int_a^b} f dx = \inf\{U(P,f) : P \text{ a partition of}[a,b]\}$ and the lower Riemann integral is $\underline{\int_a^b} = \sup\{L(P,f) : P \text{ a partition of}[a,b]\}$.

Theorem 17.4. *For any partition P, $L(P,f) \leq \underline{\int_a^b} f dx \leq \overline{\int_a^b} f dx \leq U(P,f)$.*

Definition 17.5. f is Riemann integrable if $\underline{\int_a^b} f dx = \overline{\int_a^b} f dx$.

Remark 17.6. The upper and lower integrals always exist for bounded f.

18. Lecture 18 — April 9, 2019

Last time we talked about Riemann integrals, and decided to call a function Riemann-integrable on [a,b] if $\sup_P L(P,f) = \underline{\int_a^b} f dx = \overline{\int_a^b} f dx = \inf_P U(P,f)$. We then write these quantities as $\int_a^b f(x) dx$, and write $f \in R$ to denote that f is integrable.

If f is bounded, then we're guaranteed the existence of all lower and upper sums $L(P,f) = \sum_i^n m_i \Delta x_i$, $U(P,f) = \sum_{i=1}^n M_i \Delta x_i$, as these quantities are bounded by $m(b-a)$ and $M(b-a)$, respectively, where $m \leq f(x) \leq M$.

Definition 18.1. A <u>refinement</u> of the partition $P = \{x_0, \ldots, x_n\}$ is a partition $P^* = \{x_0^*, \ldots, x_N^*\}$ with $\{x_0, \ldots, x_n\} \subseteq \{x_0^*, \ldots, x_N^*\}$.

Proposition 18.2. *For P^* a refinement of P,*

$$L(P,f) \leq L(P^*,f) \leq U(P^*,f) \leq U(P,f)$$

Proof sketch. $L(P,f) \leq L(P^*,f)$ because when $[x_{i-1}^*, x_i^*] \subseteq [x_{j-1}, x_j]$, $m_i^* \geq m_j$. Likewise for $U(P^*,f) \leq U(P,f)$. ☐

Then, given any partition P_1, P_2, there exists a refinement of both P_1 and P_2, say P^*, which cuts the interval whenever P_1 or P_2 do. By the above proposition, we arrive at

$$L(P_1, f) \leq L(P^*, f) \leq U(P^*, f) \leq U(P_2, f)$$

Taking the supremum over all P_1, keeping P_2 fixed, we arrive at

$$\underline{\int_a^b} f dx = \sup_{P_1} L(P_1, f) \leq U(P_2, f)$$

Now varying P_2 taking the infimum, we conclude:

$$\underline{\int_a^b} f dx \leq \inf_{P_2} U(P_2, f) = \overline{\int_a^b} f dx$$

Theorem 18.3.

- $\underline{\int_a^b} f dx \leq \overline{\int_a^b} f dx$
- *$f \in R$ if and only if $\forall \epsilon > 0$, $\exists$ partition P with $U(P, f) - L(P, f) \leq \epsilon$.*

Proof. We've already proven the first claim. To show the second claim, first suppose $f \in \mathbb{R}$. By the definition of inf and sup, given any $\epsilon > 0$, $\exists P_1, P_2$ such that $L(P_1, f) \geq \int_a^b f dx - \epsilon/2$ and $U(P_2, f) \leq \int_a^b f dx + \epsilon/2$. Taking P^* to be a refinement of the P_i, we have $U(P^*, f) - L(P^*, f) \leq \epsilon$. Conversely, if $\forall \epsilon > 0$ $\exists P$ such that $U(P, f) - L(P, f) \leq \epsilon$, then $\overline{\int} - \underline{\int} \leq \epsilon$. Since this holds for any ϵ, $\overline{\int} = \underline{\int}$. □

Remark 18.4. If $U(P, f) - L(P, f) \leq \epsilon$, then $\forall s_i, t_i \in [x_{i-1}, x_i]$, $\sum_{i=1}^n |f(s_i) - f(t_i)| \Delta x_i \leq \epsilon$ (as this quantity is bounded by $\sum (M_i - m_i) \Delta x_i = U(P, f) - L(P, f)$). So, also assuming $f \in R$, one can conlude

$$\left| \sum_{i=1}^n f(s_i) \Delta x_i - \int_a^b f(x) dx \right| \leq \epsilon$$

Theorem 18.5. *If f is continuous on $[a, b]$, then $f \in R$.*

Proof. Fix ϵ. We'd like to build P such that $U(P, f) - L(P, f) \leq \epsilon$. So we'd like to ensure that $M_i - m_i \leq \frac{\epsilon}{b-a}$. Since f is continuous on a compact set, it's uniformly continuous. Thus, there exists δ for which $|x - y| < \delta \implies |f(x) - f(y)| < \frac{\epsilon}{b-a}$ $\forall x, y \in [a, b]$. Now pick a partition P of $[a, b]$ with N equal steps of width $\delta x_i = \frac{b-a}{N}$ such that $\frac{b-a}{N} < \delta$. For any $s, t \in [x_{i-1}, x_i]$, $|f(s) - f(t)| < \frac{\epsilon}{b-a}$. So $M_i - m_i \leq \frac{\epsilon}{b-a}$. Thus

$$\begin{aligned} U(P, F) - L(P, f) &= \sum_{i=1}^n (M_i - m_i) \Delta x_i \\ &\leq \frac{\epsilon}{b-a} \sum_{i=1}^N \Delta x_i \\ &= \epsilon \end{aligned}$$

□

Theorem 18.6. *If f is monotonic on $[a, b]$, it's integrable.*

Proof. Without loss of generality, assume f is monotonically increasing. Fixing $\epsilon > 0$, take P such that all Δx_i are equal and are weakly less than $\frac{\epsilon}{f(b)-f(a)}$. Because f is monotonic, $M_i = f(x_i)$ and $m_i = f(x_{i-1})$. So $L(P, f) = \sum_{i=1}^{n} f(x_{i-1})\Delta x_i =) f(x_0) + \cdots + f(x_{n-1}))\Delta x_i$ and likewise $U(P, f) = (f(x_1 + \cdots + f(x_n))\Delta x_i$. Thus

$$U(P, f) - L(P, f) = (f(b) - f(a))\Delta x_i \leq \epsilon$$

□

Theorem 18.7. *If f is bounded on $[a, b]$ and has finitely many discontinuities, then f is integrable.*

Proof sketch. Take increasingly narrow intervals around the discontinuities, and integrate the rest using the argument for continuous fucntions. □

Theorem 18.8. *If f is integrable and bounded on $[a, b]$, i.e. $m \leq f \leq M$, and φ is continuous on $[m, M]$, then $\varphi \circ f$ is integrable on $[a, b]$.*

Proof. See Rudin. □

Theorem 18.9.
(a) If f_1, f_2 are integrable on $[a, b]$, then $f_1 + f_2 \in \mathbb{R}$ and $\int_a^b (f_1 + f_2)dx = \int_a^b f_1 dx + \int_a^b f_2 dx$. Likewise, $\forall v \in \mathbb{R}$, cf is integrable with $\int_a^b (cf)dx = c\int_a^b f dx$.
(b) If $f_1(x) \leq f_2(x)$ are integrable on [a,b], then $\int_a^b f_1 dx \leq \int_a^b f_2 dx$.
(c) If f is integrable on $[a, b]$ and $a < c < b$, then f is also integrable on $[a, c]$ and $[c, b]$, and

$$\int_a^b f dx = \int_a^c f dx + \int_b^c f dx$$

(d) If f is integrable and $|f(x)| \leq M$ on [a,b], then $|\int_a^b f dx| \leq M(b - a)$
(e) If f and g are integrable, then fg is integrable.
(f) If f is integrable, then $|f|$ is as well, and $|\int f dx| \leq \int |f| dx$.

Proof. (a) Note that $L(f_1 + f_2, P) \leq L(f_1, P) + L(f_2, P)$. Observing the analogous result for upper sums, the result holds.

□

19. Lecture 19 — April 11, 2019

Recall that f is integrable (which we write $f \in R$) if $\underline{\int} = \overline{\int}$ or, equivalently, for any $\epsilon > 0$, there exists a partition P with $U(P, f) - L(P, f) < \epsilon$. Last time we saw that continuous functions, piece-wise continuous functions with finitely many discontinuities, and monotonic functions are integrable (on sets of the form $[a, b]$). We also saw that R is closed under addition and multiplication, meaning sums and products of integrable functions are integrable.

Today we'll be talking about change of variables - sometimes called u-substitution in calculus courses - and this appears as 6.17 and 6.19 in Rudin.

Theorem 19.1. *Say f is integrable on $[a, b]$ and φ is a strictly increasing function which surjects from $[A, B]$ to $[a, b]$. Assume $\varphi' \in R$. Then $g(y) = f(\varphi(y))\varphi'(y)$ on $[A, B]$ is integrable, and furthermore $\int_A^B g(y)dy = \int_A^B f(\varphi(y))\varphi'(y)dy = \int_a^b f(x)dx$.*

Proof. To each partition $P = \{x_0, \dots, x_n\}$ of $[a,b]$ we can associate a partition $Q = \{y_0, \dots, y_n\}$ of $[A,B]$ such that $x_i = \varphi(y_i)$. Given such P, Q, let

$$
\begin{aligned}
m_i(f) &= \inf_{x\in[x_{i-1},x_i]} f(x) = \inf_{y\in[y_{i-1},y_i]} f\circ\varphi(y) \\
m_i(\varphi') &= \inf_{y\in[y_{i-1},y_i]} \varphi'(y) \\
\min(m_i(f)m_i(\varphi'), m_i(f)M_i(\varphi')) \le m_i(g) &= \inf_{y\in[y_{i-1},y_i]} g(y)
\end{aligned}
$$

So $L(P,f) = \sum_i m_i(f)(x_i - x_{i-1}) = \sum_i m_i(f)\varphi'(y_i^*)(y_i - y_{i-1})$, by the Mean Value theorem and $\varphi(y_i) = x_i, \varphi(y_{i-1}) = x_{i-1}$. Since ϕ' is integrable, $\forall\epsilon > 0$, $\exists P, Q$ s.t. $\forall y_i^* \in [y_{i-1}, y_i]$, $\sum_i \max(|\varphi'(y_i^*) - m_i(\varphi')|, |M_i(\varphi') - \varphi'(y_i^*)|)\Delta y_i \le \epsilon$. And

$$
\begin{aligned}
L(Q,G) &= \sum_i m_i(g)\Delta y_i \\
&\ge \sum_i \min(m_i(f)m_i(\varphi'), m_i(f)M_i(y'))\Delta y_i \\
&\ge \sum m_i(f)\varphi'(y_i^*)\Delta y_i - \sum_i |m_i(f)| \max(|\varphi'(y_i^*) - m_i(\varphi')|, |M_i(\varphi') - \varphi'(y_i^*)|)\Delta y_i \\
&\ge L(P,f) - (\max_i |m_i(f)|)\epsilon
\end{aligned}
$$

Since f is bounded - it's a continuous function on a compact set - the last line is of the form $L(P,f) - c\epsilon$ for some constant c. So, taking sufficiently fine partitions, $L(P,f)$ can be brought arbitrarily close to $L(Q,g)$, meaning $\underline{\int_A^B} g(y)dy = \sup L(Q,g) \ge \int_a^b f(x)dx$. Performing an almost identical procedure with upper sums, we have that $\overline{\int_A^B} g(y)dy \le \int_a^b f dx$. We conclude that $\int_A^B g(y)dy = \int_a^b f(x)dx$, as desired. ∎

This is a pretty messy proof, but one of the crucial steps was applying the Mean Value theorem to lower/upper sums of f in order to witness them as sums of scaled values of φ', rather than of f.

Theorem 19.2. *Let f be integrable on $[a,b]$ and define $F(x) = \int_a^b f(t)dt$ for $x \in [a,b]$. Then F is continuous on $[a,b]$ and if f is continuous at x, then F is differentiable at x with derivative $F'(x) = f(x)$.*

Proof. Since f is integrable, it's bounded, so say $|f(t)| \le M$ for $t \in [a,b]$. Then for $a \le x \le y \le b$, $|F(y) - F(x)| = |\int_x^y f(t)dt| \le M|y - x|$. So F is continuous, by taking $\delta = \frac{\epsilon}{M}$. Now, assuming f is continuous at x, given $\epsilon > 0$ select δ such that $|t - x| < \delta \implies |f(t) - f(x)| < \epsilon$. Then for $s, t \in (x-\epsilon, x+\epsilon)$,

$$|\frac{F(t) - F(s)}{t - s} - f(x)| = |\frac{1}{t-s}\int_s^t (f(u) - f(x))du| \le \epsilon$$

∎

Theorem 19.3 (Fundamental Theorem of Calculus)**.** *If f is integrable on $[a,b]$ and F is differentiable on $[a,b]$ with $F' = f$, then $\int_a^b f(x)dx = F(b) - F(a)$.*

Proof. Given $\epsilon > 0$, integrability of f implies the existence of a partition P of $[a, b]$ with $U(P, f) - L(P, f) \leq \epsilon$, and thus $\forall x_i^* \in [x_{i-1}, x_i]$, $|\sum_i f(x_i^*)\Delta x_i - \int_a^b f(x)dx| \leq \epsilon$. By the Mean Value theorem, $\exists x_i^* \in [x_{i-1}, x_i]$ with $F(x_i) - F(x_{i-1}) = F'(x_i^*)(x_i - x_{i-1}) = f(x_i^*)\Delta x_i$. So

$$\begin{aligned} F(b) - F(a) &= \sum_i (F(x_i) - F(x_{i-1})) \\ &= \sum_i f(x_i^*)\Delta x_i \end{aligned}$$

Which is within ϵ of $\int_a^b f(x)dx$ for arbitrary ϵ. So it's identically $\int_a^b f(x)dx$. ☐

Theorem 19.4 (Integration by parts)**.** *Suppose F, G are differentiable on $[a, b]$ and $F' = f, G' = g$ are integrable on $[a, b]$. Then $\int_a^b F(x)g(x)dx = F(b)G(b) - F(a)G(a) - \int_a^b f(x)G(x)dx$.*

Proof. Let $H(x) = F(g)G(x)$. Then $H' = fG + Fg$ by the product rule. Apply the Fundamental Theorem of Calculus! ☐

20. Lecture 20 — April 16, 2019

The goal is now to discuss sequences and series of functions. In order to define series of functions, we need to be able to speak of sums of functions - which we'll define pointwise - so the codomains of our functions need to support an addition operation. For this reason, we'll restrict ourselves to real-valued functions.

Definition 20.1. A sequence $f_n : (E, d) \to \mathbb{R}$ converges pointwise to $f : E \to \mathbb{R}$ if $\forall x \in E$, $f_n(x) \to f(x)$.

It turns out that sequences of continuous functions *do not* in general converge to continuous functions, and sequences of differentiable functions do not converge to differentiable functions either. For this reason, we'll be considering stronger forms of convergence, like uniform convergence.

Definition 20.2. A series $\sum_{n=0}^n f_n(x)$ converges pointwise if $\forall x \in E$, $\sum f_n(x)$ converges (meaning the sequence of partial sums converges).

Example 20.3. Let $f_n = \frac{x^2}{(1+x^2)^n} : \mathbb{R} \to \mathbb{R}$. The f_n are all continuous. Now consider $f(x) = \sum_{n=0}^\infty f_n(x)$. We have $f(0) = \sum 0 = 0$, as $f_n(0) = 0\ \forall n \in \mathbb{N}$. For $x \neq 0$, $f(x) = x^2 \sum_{n=0}^\infty (\frac{1}{1+x^2})^n$. That's a convergent geometric series, since $x \neq 0$, so

$$\begin{aligned} f(x) &= x^2 \frac{1}{1 - \frac{1}{1+x^2}} \\ &= x^2 \frac{1}{\frac{x^2}{1+x^2}} \\ &= 1 + x^2 \end{aligned}$$

Then $f(0^+) = f(0^-) = 1 \neq 0 = f(0)$, so f isn't continuous.

Example 20.4. Let $f_n = \frac{\sin(nx)}{\sqrt{n}}$. Note that $|f_n(x)| \leq \frac{1}{\sqrt{n}} \to 0$. But $f_n'(x) = \sqrt{n}\cos(nx)$ doesn't converge at all.

Definition 20.5. A sequence of functions $\{f_n\}$ converges uniformly on E to a function f if $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall n \geq N$, $\forall x \in E$, $|f_n(x) - f(x)| < \epsilon$.

Note that this definition differs from pointwise convergence in that we're not allowed to tailor N for each $x \in E$.

Theorem 20.6. *Suppose* $\lim_{n\to\infty} f_n(x) = f(x)$ $\forall x \in E$, *meaning there is pointwise convergence. Let* $M_n = \sup_{x\in E} |f_n(x) - f(x)|$. *Then* $f_n \to f$ *uniformly* $\iff M_n \to 0$.

Theorem 20.7. f_n *converges uniformly on E if and only if* $\forall \epsilon > 0$, $\exists N$ *such that* $\forall m, n \geq N$, $\forall x \in E$, $|f_n(x) - f_m(x)| \leq \epsilon$.

Proof. First suppose f_n converges uniformly to some f. Fixing $\epsilon > 0$, $\exists N$ s.t. $\forall x \in E$, $\forall n \geq N$, $|f_n(x) - f(x)| \leq \epsilon/2$. Then for $m, n \geq N$, by the triangle inequality, $|f_n(x) - f_m(x)| \leq \epsilon$.

Now suppose that $\{f_n\}$ is uniformly Cauchy. Then $\forall x \in E$, $\{f_n(x)\}$ is Cauchy in $\mathbb{R}, \mathbb{C}, \mathbb{R}^k$, meaning it converges to some limit $f(x)$. Then $f_n(x) \to f(x)$ pointwise. And, given any ϵ, $\exists N$ s.t. $\forall m, n \geq N$, $\forall x \in E$, $|f_n(x) - f_m(x)| \leq \epsilon$. Taking the limit as m tends to ∞, for fixed x, n, we have that $|f_n(x) - f(x)| \leq \epsilon$ $\forall n \geq N$, $\forall x \in E$. □

Definition 20.8. A series $\sum f_n$ converges uniformly on E to its sum $s(x) = \sum f_n(x)$ if the sequence of partial sums $s_n(x) = \sum_{k=0}^n f_k(x)$ converges to $s(x)$ uniformly.

Remark 20.9. If $\sum f_n$ converges uniformly, then $\sup_{x\in E} |f_n(x)| \to 0$ as $n \to \infty$. This is a consequence of the Cauchy criterion when $m = n + 1$.

It's important to keep in mind that the above condition is necessary but not sufficient (and not even sufficient if $\sum f_n$ is known to converge pointwise!). Recall the series $\sum \frac{1}{n}$, and observe the series $\sum \frac{x^2}{(1+x^2)^n}$ - it converges pointwise and it's uniformly bounded, but it doesn't converge uniformly.

Theorem 20.10. *If* $|f_n(x)| \leq M_n$ $\forall x \in E$ *and if* $\sum M_n$ *converges, then* $\sum f_n$ *converges uniformly on E.*

Proof. Given $\epsilon > 0$, we'd like to construct N with $m \geq n \geq N$ implies $|s_m(x) - s_n(x)| = |\sum_{k=n+1}^m f_k(x)| \leq \epsilon$. We know that $|\sum_{k=n+1}^m f_k(x)| \leq \sum_{k=n+1}^m |f_k(x)| \leq \sum_{k=n+1}^m M_k$. And the rightmost term can be made arbitrarily small for sufficiently large n, by the Cauchy criterion on convergence of $\sum M_n$. □

21. Lecture 21 — April 18, 2019

Last time we saw that uniform convergence of a sequence of functions $f_n : E \to \mathbb{R}$ to the function $f : E \to \mathbb{R}$ meant that $\forall \epsilon > 0, \exists N \in N$ s.t. $\forall n \geq N, \forall x \in E, |f_n(x) - f(x)| < \epsilon$.Equivalently, $\sup_{x\in E} |f_n(x) - f(x)| = M_n \to 0$. In words, it means that we can be guaranteed that eventually *all* outputs of the f_n are as close as desired to f, rather than needing to tailor the waiting time for each point in the domain.

Theorem 21.1. *Suppose $f_n \to f$ uniformly and $\lim_{t\to x} f_n(t) = A_n$. Then the sequence A_n converges and*

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

Corollary 21.2. *The uniformly limit of a sequence of continuous functions is itself a continuous function.*

Proof. Apply the previous theorem with $A_n = f_n(x)$. By definition, $\lim_{n\to\infty} A_n = f(x)$, so $\lim_{t\to x} f(t) = \lim_{n\to\infty} f_n(x) = f(x)$. Thus f is continuous. $\square$

Theorem 21.3. *Let $f_n : [a,b] \to \mathbb{R}$ be integrable, and assume $f_n \to f$ uniformly on $[a,b]$. Then f is integrable and $\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx$.*

Proof. Let $M_n = \sup_{x\in[a,b]} |f_n(x) - f(x)|$. By uniform convergence, $M_n \to 0$. For any $x \in [a,b]$, we have $f_n(x) - M_n \le f(x) \le f_n(x) + M_n$. Then

$$\int_a^b f_n(x)dx - M_n(b-a) \le \underline{\int_a^b} f(x)dx \le \overline{\int_a^b} f(x)dx \le \int_a^b f_n(x)dx + M_n(b-a)$$

Thus

$$\begin{aligned}\overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx &\le (\int_a^b f_n(x)dx + M_n(b-a)) - (\int_a^b f_n(x)dx - M_n(b-a)) \\ &= 2M_n(b-a) \\ &\to 0\end{aligned}$$

Meaning f is integrable. We also have $|\int_a^b f(x)dx - \int_a^b f_n(x)dx| \le M_n(b-a) \to 0$, so the integrals indeed coincide. $\square$

We now turn our focus to the relationship between convergence of sequences of functions and differentiation. An important observation is that even uniform convergence does not imply that the limit of the derivatives is the derivative of the limit.

Example 21.4. Consider $f_n(x) = \frac{1}{\sqrt{n}}\sin(nx)$. $f_n(x) \to 0$ but $f_n'(x) = \sqrt{n}\cos(nx)$, which does not converge to the derivative of the zero function.

Theorem 21.5. *Suppose f_n are differentiable on $[a,b]$ and $f_n(x_0)$ converges for some $x_0 \in [a,b]$. Suppose also that the f_n' converge uniformly on $[a,b]$. Then the f_n converge uniformly on $[a,b]$ to a limit f, and f is differentiable with $f'(x) = \lim_{n\to\infty} f_n'(x)$.*

Definition 21.6. Let $\mathcal{C}(X)$ denote the space of bounded, continuous functions from a metric space X to $\mathbb{R}$. For $f \in \mathcal{C}(X)$, let $||f|| = \sup_{x\in X} |f(x)|$. The distance function $\rho(f,g) = ||f-g||$ then turns $\mathcal{C}(x)$ into a metric space, as

(i) $||f-g|| = 0 \iff \sup|f(x)-g(x)| = 0 \iff f(x) = g(x)\ \forall x \in X$
(ii) $||f-g|| = \sup|f(x)-g(x)| = \sup|g(x)-f(x)| = ||g-f||$

(iii)

$$\begin{aligned} ||f-h|| &= \sup_{x\in X}|f(x)-h(x)| \\ &\le \sup_{x\in X}(|f(x)-g(x)|+|g(x)-h(x)|) \\ &\le ||f-g||+||g-h|| \end{aligned}$$

Theorem 21.7. *$\mathcal{C}(X)$ is complete.*

Proof. Let f_n be a Cauchy sequence in $\mathcal{C}(X)$. Then for $\forall \epsilon > 0$, $\exists N$ s.t. $\forall m, n \ge N$, $||f(n) - f(m)|| < \epsilon$. Because of our choice of metric, this means that the f_n are uniformly Cauchy. Because $\mathbb{R}$ is complete, the $f_n(x)$ thus converge pointwise to a value for each $x \in X$, Because the convergence is uniform, the limit is itself continuous. □

22. Lecture 22 — April 23, 2019

:(

23. Lecture 23 — April 25, 2019

Let's return to functions defined by power series, $f(x) = \sum_{n=0}^{\infty} c_n x^n$. We've seen that an important quantity here is the radius of convergence $R = \frac{1}{\limsup |c_n|^{1/n}}$, which lives in the extended reals. The root test told us that the series converges (absolutely) for $|x| < R$ and diverges for $|x| > R$. Now that we have the tools to think about series of functions, we can generalize our results.

Theorem 23.1. *Suppose $\sum_{n=0}^{\infty} c_n x^n$ converges for $|x| < R$, and define $f(x) = \sum_{n=0}^{\infty} c_n x^n$ on $(-R, R)$. Then*

1) *The series converges uniformly on $[-R+\epsilon, R-\epsilon]$ $\forall \epsilon > 0$.*
2) *f is differentiable in $(-R, R)$, meaning it is also continuous.*
3) *$f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1}$.*

Proof. We've seen that in the series $\sum g_n$, where $g_n(x) = c_n x^n$, if g_n is bounded by M_n and $\sum M_n$ converges, then $\sum g_n$ converges uniformly. We'd then like to show that $\sup_{x\in E} |f(x) - f_n(x)| = \sup_{x \in E} |\sum_{k=n+1}^{\infty} c_k x^k| \to 0$.

Using this criterion and the fact that, for $|x| \le R - \epsilon$, we have that $|c_n x^n| \le |c_n|(R-\epsilon)^n$, take $M_n = |c_n|(R-\epsilon)^n$. Then $\limsup M_n^{1/n} = \limsup |c_n|^{1/n}(R-\epsilon) = (R-\epsilon)\limsup |c_n^{1/n}$. Since $R - \epsilon < R$, and $\limsup c_n^{1/n} = \frac{1}{R}$, this quantity comes out to less than one. By the root test, $\sum M_n$ converges.

Now, given that $f_n = \sum_{k=0}^{n} c_k x^k \to f$ uniformly on $[-R+\epsilon, R-\epsilon]$ for all ϵ, f is continuous on $[-R+\epsilon, R-\epsilon]$ for all ϵ. So in fact it's continuous on $(-R, R)$. Finally, we've seen that if f_n' converges uniformly to a limit, then f is differentiable and $f'(x) = \lim_{n\to\infty} f_n'(x)$. Now, $f_n' = \sum_{k=1}^{n} k c_k x^{k-1}$ are the partial sums of the power series $h(x) = \sum_{k=1}^{\infty} k c_k x^{k-1}$. Note that $\limsup |kc_k|^{1/k} = \limsup |c_k|^{1/k}$, because $\sqrt[k]{k} \to 1$. So the radius of convergence of h is the same as for f - they converge pointwise on $(-R, R)$ and uniformly on $[-R+\epsilon, R-\epsilon]$. So over $[-R+\epsilon, R-\epsilon]$, $f_n' \to h$ uniformly. Then, by Theorem 7.17

in Rudin, since f is differentiable and $f' = h$ on $[-R+\epsilon, R-\epsilon]$ for any ϵ, $f' - h$ on $(-R, R)$. $\square$

Corollary 23.2. *For f infinitely differentiable on $(-R, R)$ - often called smooth - applying the theorem k times shows $f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)c_n x^{n-k}$. In particular, $c_k = \frac{1}{k!} f^{(k)}(0)$.*

So a given smooth function f is given by a power series near 0 if and only if its Taylor series at 0 converges and equals f. To see that the last condition is not trivial, observe the following example.

Example 23.3. Consider $f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$. One can check that f is smooth, meaning it has derivatives of all orders, but $f^{(n)}(0) = 0\ \forall n$. So the Taylor series at 0 converges, but to the zero function rather than f.

There are several ways to define the exponential function. The first, which is not so pretty, is to define the real number e as $e = \sum_{n=0}^{\infty} \frac{1}{n!} \approx 2.718$, then define e^n for integral n in the natural way, define $e^{p/q}$ as $(e^p)^{1/q}$, and define e^x for x irrational as $\sup\{e^{p/q} : \frac{p}{q} < x\}$.

The second, much nicer way is to define the exponential function as $\exp(x) = \sum_{n=1}^{\infty} \frac{x^n}{n!}$. Its radius of converge is ∞, since $n! > (\frac{n}{2})^{\frac{n}{2}}$, which implies $(n!)^{1/n} \to \infty$. So this function is well defined for all $x \in R$ (and, as it happens, all $x \in \mathbb{C}$). It's continuous, and in fact differentiable, and we have

$$\begin{aligned} \exp(x)\exp(y) &= (\sum_{k=0}^{\infty} \frac{x^k}{k!})(\sum_{m=0}^{\infty} \frac{y^m}{m!}) \\ &= \sum_{n=0}^{\infty} (\sum_{k=0}^{n} \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!}) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} (\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}) \\ &= \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} \\ &= \exp(x+y) \end{aligned}$$

As a special case, $\exp(x)\exp(-x) = \exp(0) = 1$. Since $\exp(x) > 0$ for $x \geq 0$, it follows that $\exp(x) > 0$ for $x < 0$ as well (otherwise $\exp(x)\exp(-x) < 0$). And since $\exp' = \exp > 0$, the function is strictly increasing. Since $\exp(1) = e$, the result also shows that $\exp(n) = \exp(1 + \cdots + 1) = \exp(1)^n = e^n$.

Since exp is a strictly increasing function $\mathbb{R} \to \mathbb{R}_+$, it has an inverse $\log : \mathbb{R}_+ \to \mathbb{R}$ defined by the property that $\exp(\log(y)) = y\ \forall y > 0$ and $\log(\exp(x)) = x\ \forall x \in \mathbb{R}$. We get for free that log is also strictly increasing. We also have that $\lim_{x \to 0} \log(x) = -\infty$, and the chain rules gives us

$$\begin{aligned}
\log(\exp x) &= x \\
\log'(\exp x)\exp' x &= 1 \\
\log'(\exp x) &= \frac{1}{\exp x} \\
\log'(y) &= \frac{1}{y}
\end{aligned}$$

By our results for the exponential, we also have that $\log(uv) = \log u + \log v$ and $\log(1) = 0$, which implies $\log(\frac{1}{u}) = -\log u$. Finally, we define the trigonometric functions to be

$$\begin{aligned}
\cos(x) &= \frac{\exp(ix) + \exp(ix)}{2} \\
\sin(x) &= \frac{\exp(ix) - \exp(-ix)}{2i}
\end{aligned}$$

where $i \in \mathbb{C}$ such that $i^2 = 1$. You can check that $\cos'(x) = -\sin(x)$ and $\sin'(x) = \cos(x)$.

24. Lecture 24 — April 30, 2019

Today we'll be talking about something slightly more applied than most of what we've looked at in the course - Fourier series.

Definition 24.1. A trigonometric polynomial is a finite sum of the form $f(x) = a_0 + \sum_{n=1}^{N} a_n \cos(nx) + b_n \sin(nx)$. Trigonometric polynomials are 2π-periodic and infinitely differentiable (or smooth). A trigonometric series is a series of the form $a_0 + \sum_{n=1}^{N} a_n \cos(nx) + b_n \sin(nx)$.

Recall from last time that $e^{ix} = \cos x + i \sin x$. Then complex generalizations of these definitions are, respectively,

$$f(x) = \sum_{n=-N}^{N} c_n e^{inx}$$

and

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

Though these produce complex outputs in general, they are real-valued if $\overline{c_n} = c_{-n}$ for all n. We'll be working with the complex forms of trigonometric polynomials/series, as they're a bit easier to handle.

Our first result is that for $f(x) = \sum_{-N}^{N} c_n e^{inx}$ and $m \in \{-N, \dots, N\}$, then

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-imx}dx &= \frac{1}{2\pi}\int_0^{2\pi}\sum_{-N}^{N} c_n e^{i(n-m)x}dx \\ &= \frac{1}{2\pi}\sum_{n=-N}^{N} c_n \int_0^{2\pi} e^{i(n-m)x}dx \\ &= c_m\end{aligned}$$

So the coefficients of such a trigonometric polynomial can be detected by performing appropriate integrals. Motivated by the result, we make the following definitions.

Definition 24.2. The mth Fourier coefficient of f, assumed integrable on $[0, 2\pi]$, is

$$c_m(f) = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-imx}dx$$

The Fourier sum of f is

$$s_N(f)(x) = \sum_{-N}^{N} c_n(f)e^{inx}$$

We also have

Definition 24.3. A sequence $\{\phi_n\}$ of complex-valued functiosn on $[a, b]$ is an orthogonal system of functions if $\int_a^b \phi_n(x)\overline{\phi_m(x)}dx = 0$ whenever $n \neq m$. It is an orthonormal system if furthermore $\int_a^b |\phi_n(x)|^2 dx = 1$ for all n.

Example 24.4. Consider $\{\frac{1}{\sqrt{2\pi}}e^{inx}\}_{n\in\mathbb{Z}}$ on $[0, 2\pi]$ (or $[-\pi, \pi]$). This forms an orthonormal system because

$$\begin{aligned}\int_0^{2\pi} \phi_n(x)\overline{\phi_m(x)} &= \frac{1}{2\pi}\int_0^{2\pi} e^{inx}e^{-imx}dx \\ &= \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}\end{aligned}$$

Theorem 24.5. *Let $\{\phi_n\}$ be an orthonormal system on $[a, b]$, and consider the Nth Fourier sum of f - $s_N(x) = \sum_{n=1}^{N} c_n\phi_n(x)$. Then for any $t_N(x) = \sum_{n=1}^{N} d_n\phi_n(x)$,*

$$\int_a^b |f - t_N|^2 dx \geq \int_a^b |f - s_N|^2 dx$$

with equality if and only if $c_n = d_n$ for all n.

Proof. First some intermediate results:

$$\int_a^b f\overline{t_n}dx = \sum_{n=1}^{N} \overline{d_n} \int_a^b f\overline{\phi_n}dx$$
$$= \int_{n=1}^{N} \overline{d_n}c_n$$

And

$$\int_a^b |t_N|^2 dx = \int_a^b (\sum_{n=1}^{N} d_n\phi_n)(\sum_{m=1}^{N} \overline{d_m}\overline{\phi_m})dx$$
$$= \int_a^b \phi_n\overline{\phi_n}$$
$$= \begin{cases} 0 & m \neq n \\ 1 & m = n \end{cases}$$

Note that the second equality used the fact that only terms in which $m = n$ contribute to the sum. Returning to the original problem, we have

$$\int_a^b |f - t_N|^2 dx = \int_a^b |f|^2 - f\overline{t_N} - \overline{f}t_n + |t_N|^2 dx$$
$$= \int_a^b |f|^2 dx - \sum_{n=1}^{N} c_n\overline{d_n} - \sum_{n=1}^{N} \overline{c_n}d_n + \sum_{n=1}^{N} |d_n|^2$$
$$= \int_a^b |f|^2 dx - \sum_{n=1}^{N} |c_n|^2 + \sum_{n=1}^{N} |c_n - d_n|^2$$

The first two terms are irrespective of t_N and the second is minimized when $c_n = d_n$, proving the claim. □

In words, what we've shown is that s_N is the best approximation to f in the least-squares sense.

Corollary 24.6. *$\sum_{n=1}^{\infty} |c_n|^2 \leq \int_a^b |f(x)|^2 dx$, meaning the left hand side converges. In particular, $c_n \to 0$ as $n \to \infty$.*

For the remainder of the class, we'll state some important theorems which we don't have time to prove.

Theorem 24.7. *If for given x, $\exists \delta > 0$, $\exists M > 0$ s.t. $|f(x+t) - f(x)| \leq M|t|$ for $|t| < \delta$, then $\lim_{N\to\infty} s_N(f)(x) = f(x)$.*

In words, the above conditions - which are stronger that continuity but weaker than differentiability - suffice to guarantee that the Fourier sums converge pointwise.

Theorem 24.8 (Stone-Weierstrass)**.** *If f is continuous and 2π-periodic, then f is the uniform limit of a sequence of trigonometric polynomials.*

Theorem 24.9 (Parseval)**.** *If f is integrable and 2π-periodic, with Fourier coefficients c_n and partial sums s_N, then*

- $\lim_{N\to\infty} \frac{1}{2\pi} \int_0^{2\pi} |f(x) - s_N(x)|^2 dx = 0$
- $\sum |c_n|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx$

That's all - congratulations on having (almost) completed Math 112!

25. Review section — May 7, 2019

The final will take place the coming Monday, May 13, from 9am-12pm in Emerson 210. You're allowed to bring a copy of Rudin. There's lots of space on the test for scratch work, so there shouldn't be a need to bring scratch paper. The format of the final is approximately C times the midterm, for $1 < C < 2$. Expect a similar mix of concrete and abstract material as in the midterm, and keep in mind that the standard of detail for proof is slightly relaxed from that expected on homework, given the time limitation. It is cumulative, and will cover the semester's content more or less uniformly.

Chapter 1

We learned about fields, which are sets equipped with operations of addition and multiplication satisfying some properties. Examples were $\mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$. $\mathbb{Q}$ and $\mathbb{R}$ are additionally ordered fields, and $\mathbb{R}$ is made special by the least upper bound property, which states that non-empty sets with upper bounds have least upper bounds. We learned that the rationals (and irrationals) are dense in $\mathbb{R}$, and that $\mathbb{R}$ is complete, meaning its Cauchy sequences converge.

Chapter 2

We talked about unions and intersections of sets, what functions are, and the difference between finite, countable, and uncountable sets. Countable sets are the smallest kind of infinite sets, and they can be enumerated, which means formally that they admit a bijection to $\mathbb{N}$ (or, equivalently, $\mathbb{Z}$). We learned that a countable union of countable sets is countable, and that $\mathbb{R}$ is uncountable.

We learned about metric spaces, including the definitions of neighborhoods, a metric, and open/closed subsets of metric spaces. Recall that a subset of a metric space is open if all of its points are interior points (meaning they admit a neighborhood contained in the subset), and that a subset of a metric space is closed if it contains all its limit points. Equivalently, the subset is closed if limits of convergent sequences in the subset are also in the subset.

A useful result is that subset is closed if and only if its complement is open. Keep in mind that subsets can be neither open nor closed, and that they can also be open and closed. We learned functions between metric spaces are continuous if and only if preimages of open sets in the codomain are open sets in the domain (or, equivalently, preimages of closed sets in the codomain are closed in the domain).

Keep in mind that the openness or closedness of a set is a property of the ambient metric space in which it exists (for instance, $(0, 1]$ is neither closed nor open in the metric space $\mathbb{R}$, but both closed and open in the metric space $(0, 1]$). There is an important correspondence between open sets in a metric space and those in a metric space it contains. In particular, for $E \subset Y \subset X$, E is open in Y if and only if $E = Y \cap U$ for U an open set in X.

A set K is compact if all of its open covers reduce to finite subcovers (i.e. $K \subseteq \cup_{\alpha \in A} G_\alpha$ and G_α all open implies that $K \subseteq \cup_{i=1}^{n} G_{\alpha_i}$). We learned that compactness is equivalent to sequential compactness in metric spaces, which is the property that all infinite subsets have limit points. An equivalent, and perhaps more intuitive, definition of sequential compactness that all sequences in K have a convergent subsequent in K (whose limit is in K).

We learned that compact sets are always closed and bounded, and that the converse is true in $\mathbb{R}^k$ and $\mathbb{C}^K$, but not in general. Continuous functions send compact sets to compact sets, and they also send connected sets to connected sets. Applying the results to continuous functions $f : [a, b] \to \mathbb{R}$, you can conclude that $f([a, b])$ is compact and connected. This implies that the image of $[a, b]$ attains its maximum and maximum and minimum, because it is compact, and that it attains all values between $f(a)$ and $f(b)$, because of our characterizations of connected subsets of $\mathbb{R}$. These are the extreme and intermediate value theorems, respectively.

Chapter 3

We learned about sequences and series. A sequence is the image of a function defined on $\mathbb{N}$, and the limit of a sequence x_n is the point ℓ such that $\forall \epsilon > 0$, $\exists N$ s.t. $n \geq N \implies d(x_n, x) < \epsilon$. Limits are unique in metric spaces, allowing us to use the word 'the' in the previous sentence. For a function f, we write $\lim_{t \to x} f(t) = a$ if $\forall \epsilon > 0$ $\exists \delta > 0$ such that $0 < d(t, x) < \delta \implies d(f(t), a) < \epsilon$.

Convergent sequences are bounded, but bounded sequences need not be convergent. In $\mathbb{R}^k$, bounded sequences have convergent subsequences (as they sit in a compact k-cell, which is also sequentially compact). Another partial converse is that monotonic sequences in $\mathbb{R}$ converge if they're bounded.

A sequence is Cauchy if $\forall \epsilon > 0$, $\exists N$ such that $m, n \geq N \implies d(x_n, x_m) < \epsilon$. In complete metric spaces, like $\mathbb{R}^k$ and $\mathbb{C}^k$, Cauchy sequences converge (in fact, this is the definition of completeness).

Series are said to converge if the sequences of their partial sums converge, which allows us to restate results from sequences as results for series. In particular, the Cauchy criterion for series - which follows from that for sequences - implies that a real series with the following Cauchy property convergences: $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ s.t. $m, n \geq N$ implies $|\sum_{k=n}^{m} a_k| < \epsilon$. The Cauchy property for series implies that $a_n \to 0$, by considering $m = n + 1$. We can also make use of the Monotone convergence theorem for sequences to conclude that a series $\sum a_n$ with $a_n \geq 0$ converges if and only if its partial sums are bounded.

The root test considers $\alpha = \limsup |a_n|^{1/n}$. When $\alpha < 1$, it concludes that $\sum a_n$ converges absolutely (meaning $\sum |a_n|$ converges), and when $\alpha > 1$ it concludes that the series diverges. We also learned about the ratio test. Power series are defined to be series of the form $\sum_{n=0}^{\infty} c_n z^n$ - the root test leads us to define $R = \frac{1}{\limsup |c_n|^{1/n}}$, and conclude that the power series converges when $|z| < R$ and diverges when $|z| > R$. Since we're on the subject of sequences and series of numbers, we're going to jump to sequences and series of functions.

Chapter 7

A sequence of functions f_n converges to f pointwise if $\forall x \in X$, $f_n(x) \to f(x)$. Informally, uniform convergence says that this takes place in a controlled way. A sequence of functions f_n converges to f uniformly if $||f_n - f|| = \sup_{x \in X} |f_n(x) - f(x)|$ converges to 0. Equivalently, $\forall \epsilon > 0$, $\exists N$ such that $\forall x \in X, \forall n \geq N$, $|f_n(x) - f(x)| < \epsilon$. Crucially, uniform convergence does not permit delta to be selected as a function of x.

Note that pointwise and uniform convergence are properties of the functions f_n and of the domains on which they are defined. For instance, power series $\sum_{n=0}^{\infty} c_n z^n$ converge uniformly on any compact subset of $(-R, R)$ but only pointwise on $(-R, R)$.

We learned that uniform limits of continuous functions are continuous, but that the same is not true for pointwise limits. We also learned about how sequences and series of functions interact with integration and differentiation. The Weierstrass theorem states that every continuous functions on $[a, b]$ is the uniform limit of a sequence of polynomials.

Chapter 4

Here we talked about continuous functions and their properties, much of which we've already mentioned. In this chapter we defined $\lim_{t \to x} f(t)$, defined f to be continuous at x if $\forall \epsilon, \exists \delta > 0$ s.t. $\forall y, |x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$. We saw that this is equivalent to the condition that $\lim_{t \to x} f(t) = f(x)$. f is continuous on its domain if it is continuous at all points in its domain, and it is uniformly continuous if it is continuous on its domain and δ can be selected only as a function of ϵ (and not of x).

Continuity on a compact set implies uniform continuity, and continuous functions preserve compact and connected sets. For real functions $f : R \to R$, we talked about left and right limits, $f(x-)$ and $f(x+)$.

Chapter 5

When it exists, the derivative of f at x is defined $f'(x) = \lim_{t \to x} \frac{f(t) - f(x)}{t - x}$. Differentiable functions are continuous, but continuous functions need not be differentiable. Differentiable functions have derivative zero at local minima and maxima. The Mean value theorem implies that a differentiable[3] function $f : [a, b] \to \mathbb{R}$ satisfies $f(b) - f(a) = (b - a) f'(x)$ for some $x \in (a, b)$.

Chapter 6

This chapter was about integration - we used lower and upper sums to define upper and lower integrals, and defined the integral to be equal to the upper and lower integrals when they coincide. We saw that f is Riemann-integrable, written $f \in R$, if $\forall \epsilon$, $\exists P$ s.t. $U(P, f) - L(P, f) < \epsilon$.

Among functions $[a, b] \to \mathbb{R}$, the ladder of nice functions is roughly

[3]Technically we demand that f be continuous on $[a, b]$ and differentiable on (a, b), as we haven't defined the notion of differentiability for endpoints.

$$\{\text{differentiable}\} \subset \{\text{continuous}\} \subset \{\text{finitely many continuities}\} \subset \{\text{integrable}\} \subset \{\text{bounded}\}$$

All of these inclusions are strict. Integration has many of the nice properties that limits do (e.g. integration respects sums and scaling). We saw that $f \geq 0$ implies $\int_a^b f dx \geq 0$ and that $|f| \leq M$ implies $|\int_a^b f dx| \leq \int_a^b |f| dx \leq M(b-a)$.