MATH 3150 Analysis I, Fall 2016

Arthur J. Parzygnat

These are my personal notes. This is *not* a substitute for Abbott's book. You will *not* be responsible for any **Remark**s in these notes. However, everything else, including what is in Abbott's book (even if it's not here), is fair game for homework, quizzes, and exams. At the end of each lecture, I provide a list of homework problems that can be done after that lecture. I also provide additional exercises which I believe are good to know. You should also browse other books and do other problems as well to get better at writing proofs and understanding the material.

Contents

1 August 30

Real analysis is the study of continuous and smooth functions of a real variable. Historically, analysis was developed assuming several properties of the real number system. We will do this to some degree, but will eventually construct the real numbers from the rational numbers. We will begin with *sets* as one of the structures taken for granted.[1] Sets are collections of elements and are denoted by $\{x, y, z\}$, where x, y, and z denote the elements. Here are several examples.

(a) The people in this room.

(b) Positive even integers $2\mathbb{N} := \{2, 4, 6, 8, \dots\}$.

(c) All natural numbers $\mathbb{N} := \{1, 2, 3, 4, \dots\}$.

(d) All animals on this planet.

There is also a set consisting of no elements. This is called the *empty* set and is denoted by $\varnothing$. Note that the order that the elements are listed in does not matter nor are there any repeats. One common axiom of mathematics is that the empty set is a set, i.e. the empty set exists. Another axiom is that there exists a set with a single element.

Definition 1.1. Given two sets A and B, the *union* of A and B is

$$A \cup B := \{x \ : \ x \in A \text{ or } x \in B\}. \tag{1.2}$$

The right-hand-side (RHS) is read as the set of all elements x such that x is an element of A *or* x is an element of B.

Example 1.3. Let $A = \{\text{cat}, \text{dog}, \text{apple}\}$ and $B = \{\text{apple}, \text{telephone}\}$. Then

$$A \cup B = \{\text{cat}, \text{dog}, \text{apple}, \text{telephone}\}. \tag{1.4}$$

Definition 1.5. The *intersection* of A and B is

$$A \cap B := \{x \ : \ x \in A \text{ and } x \in B\}. \tag{1.6}$$

The RHS is read as the set of all elements x such that x is an element of A *and* x is an element of B.

[1] We reach some point where we must take some concepts and definitions for granted to get anywhere. Thinking of sets as just collections naively can lead to some trouble. An excellent exposition of this is given in Chapter 2 of [8]. Here's a brief summary of some points. One can either specify sets by explicitly listing all the elements or by some characteristic property. The former works for finite sets (perhaps of small number) while the latter can be used for infinite sets as well (such as the natural numbers, as we will see). However, the latter method can lead to some peculiar contradictions if one does not properly define this characteristic. Suppose, say, you consider the set of all trees on the planet. Do you mean all trees that existed, currently exist, or existed during some time period? What do we mean by tree? An even more serious issue occurs in the following example. The barber in a village was asked to shave everyone (the supposed set of people) who does not shave themselves. Should the barber then shave himself (i.e. be included in this supposed set)?

Example 1.7. Let $A = \{2, 4, 6, 8, \dots\}$ be the set of positive even numbers and $B = \{3, 6, 9, 12, \dots\}$ be the set of positive multiples of 3. Then

$$A \cap B = \{6, 12, 18, \dots\} \tag{1.8}$$

is the set of positive multiples of 6.

Definition 1.9. Let A be a set. A *subset* of A is a set B such that every element of B is an element of A. When B is a subset of A, we write

$$B \subseteq A. \tag{1.10}$$

Definition 1.11. Let A be a set. The *power set* of A is the collection of all subsets of A

$$\mathcal{P}(A) := \{B \ : \ B \subseteq A\}. \tag{1.12}$$

In particular, $\varnothing, A \in \mathcal{P}(A)$. One common axiom of mathematics is that the power set of a set is always a set.

Exercise 1.13. Prove the following. Let A be a set consisting of n elements where $n \in \mathbb{N}$. Then $\mathcal{P}(A)$ has 2^n elements. Hint: every subset $B \subseteq A$ is determined by the yes or no question "is the element $x \in A$ in the set B?"

Definition 1.14. Let A be a set and $B \subseteq A$ a subset of A. The *complement* of B in A is

$$B^c := \{x \in A \ : \ x \notin B\} \tag{1.15}$$

and is also occasionally written as[2]

$$A \setminus B. \tag{1.16}$$

The RHS of (1.15) is read as the set of all elements x of A such that x is not in B.

Theorem 1.17 (De Morgan's Laws)**.** *Let A and B be subsets of a set X. Then*[3]

$$(A \cap B)^c = A^c \cup B^c \tag{1.18}$$

and

$$(A \cup B)^c = A^c \cap B^c. \tag{1.19}$$

Before proving this, a picture helps to convince oneself of the validity of this claim—see Figure 1.

Proof. We will prove (1.18) by showing first that $(A \cap B)^c \subseteq A^c \cup B^c$ and then $A^c \cup B^c \subseteq (A \cap B)^c$.

[2] The latter notation is more precise as it shows that one is taking the complement with respect to an ambient set.

[3] Using the more precise notation of set differences, these two facts would be written as $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$ and $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$, respectively.

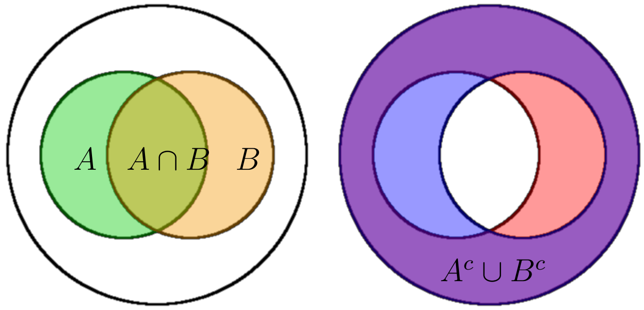


Figure 1: On the left, the sets A and B are colored green and orange, respectively. Their intersection $A \cap B$ is slightly yellow-ish, and the complement $(A \cap B)^c$ in X is everything except the yellow part. On the right, A^c is colored red and B^c is colored blue and the union $A^c \cup B^c$ is the entire shaded region, from which it can be seen that this equals the complement of $A \cap B$.

i. Case $(A \cap B)^c \subseteq A^c \cup B^c$. Let $x \in (A \cap B)^c$. By definition, this means $x \in X$ but $x \notin A \cap B$. In other words, x is not an element of both A and B. Said differently, x is either not an element of A or it is not an element of B, but this exactly means that $x \in A^c \cup B^c$.

ii. Case $A^c \cup B^c \subseteq (A \cap B)^c$. Let $x \in A^c \cup B^c$ and read the argument from above backwards.

The proof of (1.19) is left as an exercise. ■

Definition 1.20. A *function* f from a set A to a set B is an assignment associating to every $x \in A$, a unique element $f(x) \in B$. This is often written as $f : A \to B$ or $A \xrightarrow{f} B$.

Definition 1.21. A function $f : A \to B$ is *one-to-one* (a.k.a. *injective*) if for any two distinct elements $x, y \in A$, then $f(x) \neq f(y)$. f is *onto* (a.k.a. *surjective*) if for any $z \in B$, there exists an $x \in A$ such that $f(x) = z$. A function $f : A \to B$ is a *bijection* if it is one-to-one and onto. Two sets A and B have the same *cardinality* when there exists a bijection from A to B.

Definition 1.22. Let A, B, and C be three sets. Let $f : A \to B$ and $g : B \to C$ be two functions. The *composition* $g \circ f$ is the function from A to C given by $(g \circ f)(x) := g\big(f(x)\big)$ for all $x \in A$.

Definition 1.23. Let A and B be two sets. The *cartesian product* of A and B is the set of ordered pairs

$$A \times B := \{(a, b) \ : \ a \in A, b \in B\}. \tag{1.24}$$

Definition 1.25. A *relation* over the sets A and B is a subset of $A \times B$.

Exercise 1.26. Let A and B be two sets and let $f : A \to B$ be a function. Then the set

$$R_f := \big\{\big(a, f(x)\big) \ : \ a \in A\big\} \tag{1.27}$$

is a relation. Furthermore, any other relation $R \subseteq A \times B$ satisfying the conditions that[4]

$$a \in A \qquad \Rightarrow \qquad \exists\, b \in B \text{ such that } (a, b) \in R \tag{1.28}$$

[4]This is read as "if a is an element of A, then there exists an element b of B such that (a, b) is an element of R."

and[5]

$$(a,b),(a,b') \in R \qquad \Rightarrow \qquad b = b', \tag{1.29}$$

then there exists a unique function $f : A \to B$ such that $R = R_f$.

The previous exercise indicates how functions are related to relations satisfying certain conditions.

Definition 1.30. Let A be a set. An *equivalence relation* on A is a relation $R \subseteq A \times A$ satisfying the following three conditions.

(a) $(a,a) \in R$ for all $a \in A$.

(b) If $(a,a') \in R$, then $(a',a) \in R$.

(c) If $(a,a'),(a',a'') \in R$, then $(a,a'') \in R$.

Proposition 1.31. *Cardinality is an equivalence relation,*[6] *i.e. the following three facts hold.*

(a) Every set A has the same cardinality as itself, i.e. there exists a bijection from A to A.

(b) If a set A has the same cardinality as a set B, then B has the same cardinality as A, i.e. if there exists a bijection from A to B, then there exists a bijection from B to A.

(c) If a set A has the same cardinality as a set B and B has the same cardinality as a set C, then A has the same cardinality as C, i.e. if there exists a bijection from A to B and a bijection from B to C, then there exists a bijection from A to C.

Proof.

(a) The identity function is a bijection.

(b) Let $f : A \to B$ be the bijection. Define a function $f^{-1} : B \to A$ by sending $b \in B$ to the unique $a \in A$ such that $f(a) = b$. Then f^{-1} is a bijection.

(c) Let $f : A \to B$ and $g : B \to C$ be such bijections. Then $g \circ f : A \to C$ is a bijection.

You, the reader, should check these claims. ■

Definition 1.32. The function $f^{-1} : B \to A$ corresponding to a bijection $f : A \to B$ in the proof of Proposition 1.31 is called the *inverse* of f.

Definition 1.33. A set is *finite* if it is empty or has the cardinality of $\{1,2,\ldots,n\}$ for some $n \in \mathbb{N}$. A set is *countable* if it has the cardinality of $\mathbb{N}$. A set is *uncountable* otherwise.

[5] This is read as "if (a,b) and (a,b') are two elements of R, then b and b' are equal."

[6] By definition, an equivalence relation must be defined on a set. Since there is no set of all sets (see [8]), this statement does not even make sense. Instead, we can view the explanation as the statement if the claim noting the resemblance to the definition of an equivalence relation.

Proposition 1.34. *Let B be a countable set and $A \subseteq B$ a subset of B. Then A is finite or countable.*

Proof. See Abbott. ∎

Theorem 1.35 (Cantor's theorem)**.** *Let A be a set. Then there does not exist a surjective function $A \to \mathcal{P}(A)$.*

Proof. See Abbott. ∎

The student is encouraged to read the NY Times article http://opinionator.blogs.nytimes.com/2010/05/09/the-hilbert-hotel/?_r=1 on infinities and the references therein. In particular, the first few paragraphs (including the first few propositions and proofs) of Terence Tao's blog https://terrytao.wordpress.com/2009/11/05/the-no-self-defeating-object-argument/ will be helpful for our next class. Cantor's theorem is also there appearing as proposition 5. More on sets at an intuitive level can be read about in Vilenkin's fantastic book *Stories about Sets* [8]. This book also contains some material that will be useful to get a better idea of what's going on in this class. I highly recommend checking it out. When you are ready for something more advanced, the first article I recommend checking out is Leinster's "Rethinking set theory" [5]. This one is especially nice. It describes the natural numbers $\mathbb{N}$ from a seemingly drastically different perspective. A more standard treatment of set theory is in Halmos' book [4].

After this lecture, it is recommended the student works through problems 1, 2 (parts (c), (d), and (e)), 9, and 10 on HW #1. Additional recommended exercises include exercises 1.2.3, 1.2.7 (b) & (d), 1.2.8, 1.2.9 (b) (but replace "$g : \mathbb{R} \to \mathbb{R}$" with "$g : X \to Y$, where X and Y are two sets" and replace "$A, B \subseteq \mathbb{R}$" with "$A, B \subseteq X$"), 1.2.13, 1.5.1, and 1.5.3 in [2] (or equivalently exercises 1.2.2, 1.2.6 (b) & (d), N/A, 1.2.7 (b) with same comment, 1.2.12, 1.4.7, and 1.4.8 in [1]).

2 September 1

In the previous lecture, we introduced the natural numbers $\mathbb{N} := \{1, 2, 3, 4, \dots\}$. A closely related set is the set of all integers

$$\mathbb{Z} := \{\cdots, -3, -2, -1, 0, 1, 2, 3, \dots\}. \tag{2.1}$$

The set of rational numbers is

$$\mathbb{Q} := \left\{ \frac{p}{q} \ : \ p \in \mathbb{Z}, q \in \mathbb{Z} \setminus \{0\} \right\}. \tag{2.2}$$

where for a set A and a subset $B \subseteq A$ of A,

$$A \setminus B := \{x \in A \ : \ x \notin B\}. \tag{2.3}$$

Remark 2.4. The natural numbers $\mathbb{N}$ can be constructed from the empty set $\varnothing$. The integers $\mathbb{Z}$ can be obtained from the natural numbers as the set of equivalence classes of elements $(a, b) \in \mathbb{Z} \times \mathbb{Z}$ where (a, b) is equivalent to (a', b') if and only if there exists a $c \in \mathbb{Z}$ such that $a + b' + c = a' + b + c$. The reason for this equivalence relation is that one thinks of the pair (a, b) as $a - b$. For reference, this is known as the *Grothendieck group construction*. The rational numbers $\mathbb{Q}$ can also be obtained from the integers "algebraically" as the *ring of quotients*. This is typically discussed in a course on algebra [3]. However, the real numbers are quite different.

Proposition 2.5. *$\sqrt{2}$ is not a rational number.*

An easy fact helps to prove this.[7]

Lemma 2.6. *A natural number $p \in \mathbb{N}$ is odd if and only if p^2 is odd.*[8]

Proof. If p is odd, i.e. of the form $p = 2n + 1$ with $n \in \mathbb{N} \cup \{0\}$, then $(2n + 1)^2 = 4n^2 + 4n + 1 = 2(2n^2 + 2n) + 1$, which is odd. Conversely, suppose p^2 is odd and suppose to the contrary that p is even, i.e. of the form $p = 2m$ with $m \in \mathbb{N}$. Then $p^2 = 4m^2 = 2(2m^2)$ contradicting the assumption that p^2 is odd. ∎

Proof of Proposition 2.5. Suppose, to the contrary, that $\sqrt{2}$ is a rational number, i.e. $\sqrt{2} = \frac{p}{q}$ for two integers p and q, which can be chosen so that p and q have no common factor. Squaring gives $2q^2 = p^2$, which shows that p^2 is even, which by Lemma 2.6 is even if and only if p is even. Hence, $p = 2n$ for some $n \in \mathbb{N}$, so that $2q^2 = 4n^2$ giving $q^2 = 2n^2$ showing that q^2, and hence q, is even. Thus, 2 is a common factor of p and q contradicting the assumption that p and q have no common factor. ∎

[7] This is cheating a bit. One does not perform a proof in this backwards manner by first concocting a fact that is useful to prove a claim. This fact was only realized after the proof was begun. When trying to prove statements, one typically asks what the goal is and what all the assumptions and data are. Then one tries to use these assumptions to achieve the goal. During the process, certain facts might be needed. Therefore, if one is trying to learn how to prove statements, I recommend first reading the proof of the proposition and *then* referring to the lemma.

[8] Recall, "statement A if and only if statement B" means that "statement A implies statement B" *and* "statement B implies statement A."

The following sequence of *rational numbers* is a better and better approximation to the value corresponding to $\sqrt{2}$.

$$1, 1.4, 1.41, 1.414, 1.4142, 1.41421, \ldots. \tag{2.7}$$

Notice that the rational numbers are getting bigger in this case (there are also sequences that better approximate $\sqrt{2}$ that are not necesserily increasing). The set of rational numbers is an example of an ordered field, and to better understand what real numbers are, we need to review this concept.

Definition 2.8. A *field* is a set F together with addition and multiplication operations satisfying the following conditions.[9]

(a) (commutativity of addition) $x + y = y + x$ for all $x, y \in F$.

(b) (commutativity of multiplication) $xy = yx$ for all $x, y \in F$.

(c) (associativity of addition) $(x + y) + z = x + (y + z)$ for all $x, y, z \in F$.

(d) (associativity of multiplication) $(xy)z = x(yz)$ for all $x, y, z \in F$.

(e) (unit for addition) There exists an element 0 such that $x + 0 = x$ for all $x \in F$.

(f) (unit for multiplication) There exists an element 1 such that $x1 = x$ for all $x \in F$.

(g) (inverses for addition) For every $x \in F \setminus \{0\}$, there exists an element $-x \in F$ such that $x + (-x) = 0$.

(h) (inverses for multiplication) For every $x \in F \setminus \{0\}$, there exists an element $x^{-1} \in F$ such that $xx^{-1} = 1$.

(i) (distributive law) $x(y + z) = xy + xz$ for every $x, y, z \in F$.

Example 2.9. As stated above, the rational numbers $\mathbb{Q}$ with their usual definitions of addition, multiplication, and identities, is a field.

Example 2.10. Let p be a prime number. For another integer $q \in \mathbb{Z}$, the number (known as the *remainder*)

$$r := q \mod p \tag{2.11}$$

is the unique integer in $\{0, 1, \ldots, p - 1\}$ satisfying the condition that there exists an integer $n \in \mathbb{Z}$ such that

$$q = np + r. \tag{2.12}$$

Then $\mathbb{Z}_p := \{0, 1, \ldots, p - 1\}$ together with addition and multiplication defined via modular arithmetic, namely

$$x + y := (x + y) \mod p \qquad \forall\, x, y \in \mathbb{Z}_p \tag{2.13}$$

[9]Mathematical objects, such as fields, are often defined by specifying data/structure (in this case, a set F together with two binary functions $+, \cdot : F \times F \to F$) and conditions. Simply writing a definition does not guarantee such objects exist. One should also have a variety of examples.

and

$$xy := (xy) \mod p \qquad \forall\, x, y \in \mathbb{Z}_p \tag{2.14}$$

is a field. The identity for addition is 0 and the identity for multiplication is 1. The only non-trivial thing to check is that every nonzero element $q \in \mathbb{Z}_p$ has a multiplicative inverse, meaning that there exists a $q^{-1} \in \mathbb{Z}_p$ such that

$$qq^{-1} \mod p = 1 \mod p. \tag{2.15}$$

I leave showing the existence of such an inverse as an exercise (you may want to consult a book on algebra or number theory [3]).

Note that neither $\mathbb{N}$ nor $\mathbb{Z}$ are fields with their usual definitions of addition, multiplication, and identities. Even more surprisingly, $\mathbb{Z}_n$ is not a field if n is not prime.

Definition 2.16. An *ordering* on a set K is a relation $R \subseteq K \times K$, with $(x, y) \in R$ written as $x \leq y$, satisfying the following conditions.

(a) For any pair of elements $x, y \in K$, either $x \leq y$ or $y \leq x$ (or both).

(b) If $x \leq y$ and $y \leq x$, then $x = y$.

(c) If $x \leq y$ and $y \leq z$, then $x \leq z$.

If $x \leq y$ and $x \neq y$, then one often writes $x < y$. A set with an ordering is called an *ordered set*.

Definition 2.17. An *ordered field* is a field F together with an ordering $\leq$ satisfying the following.

(a) If $y \leq z$ with $y, z \in F$, then $x + y \leq x + z$ for all $x \in F$.

(b) If $0 \leq x$ and $0 \leq y$, then $0 \leq xy$.

Example 2.18. $\mathbb{Q}$ is an ordered field.

Although $\mathbb{Z}_p$ for p a prime number is a field, it is *not* an ordered field.

Exercise 2.19. Show why $\mathbb{Z}_4$ is not a field.

Exercise 2.20. Show why $\mathbb{Z}$ is not a field.

Exercise 2.21. Show why $\mathbb{Z}_m$ is not a field if $m \in \mathbb{N}$ is *composite*, i.e. there exist $a, b \in \mathbb{N} \setminus \{1\}$ such that $m = ab$.

Exercise 2.22. (Challenge)[10] Show that $\mathbb{Z}_p$ is a field if and only if p is prime.

Don't forget, there is a Quiz on Tuesday on all the material we have covered! This includes Sections 1.1, 1.2, and most of the subsection called "Field and Order Properties" from Section 8.4 (pages 245–246) in [2]. Because we ended class before getting to suprema and infema, the above exercises are meant to give you additional preparation so you might want to think about them.

[10] You will not be responsible for being able to prove this on quizzes or exams.

3 September 6

Definition 3.1. Let K be an ordered set. A subset $A \subseteq K$ is *bounded (from) above* (in K) if there exists an element $y \in K$ such that for all $x \in A$, $x \leq y$. In this case, y is called an *upper bound* for A.

Note that y need not be an element of A in the definition of an upper bound for A. It is also helpful to imagine the bounded from above condition $x \leq y$ for all $x \in A$ as an arrow

$$x \to y \quad \forall\, x \in A. \tag{3.2}$$

Example 3.3. Let

$$A := \{r \in \mathbb{Q} \;:\; r^2 < 2\}. \tag{3.4}$$

Then $1.4143 \in \mathbb{Q}$ is an upper bound for A. $1.4142 \in \mathbb{Q}$ is *not* an upper bound for A because $1.4142 < 1.41421 \in A$.

Definition 3.5. Let K be an ordered set and $A \subseteq K$ a subset. An element $y \in K$ is a *least upper bound* (a.k.a. *supremum*) for A if y is an upper bound for A and if for any other upper bound $z \in K$ of A, then $y \leq z$. In this case, the supremum is denoted by

$$\sup A = y \tag{3.6}$$

(or more appropriately $\sup_K A = y$). If such a least upper bound y exists and $y \in A$, then y is said to be a *maximum* of A.

Analogous definitions can be made for a *greatest lower bound* (a.k.a. *infimum*) and a *minimum*. The infimum of A is denoted by $\inf A$. In terms of arrows, we may sometimes use the notation

$$\begin{tikzcd} x \arrow[dd] \arrow[dr] & \\ & z \\ y \arrow[ur, dashed] & \end{tikzcd} \tag{3.7}$$

to denote the property of the least upper bound. We will later see that this is an example of a *universal property*, and we therefore refer to it as the universal property of the least upper bound of a set.

Remark 3.8. It might also be helpful to think of the supremum of A in K in the following diagrammatical way. Imagine we could draw all the elements of A

$$\bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet \longrightarrow \bullet\,, \tag{3.9}$$

where the arrows denote that the element at the tail of an arrow is less than or equal to the element on the head of an arrow. An upper bound of A is an element $\star$ in K (it might be in A) such that

$$\begin{tikzcd} \bullet \arrow[r] \arrow[drrr] & \bullet \arrow[r] \arrow[drr] & \bullet \arrow[r] \arrow[dr] & \bullet \arrow[r] \arrow[d] & \bullet \arrow[r] \arrow[dl] & \bullet \arrow[r] \arrow[dll] & \bullet \arrow[dlll] \\ & & & \star & & & \end{tikzcd} \tag{3.10}$$

meaning that $\star$ is greater than or equal to every element of A. A *least* upper bound is an element $\odot$ in K such that it is an upper bound of A (meaning it is one of the $\star$'s from before) and that for any other upper bound $\star$ of A,

$$\odot \longrightarrow \star \tag{3.11}$$

In other words, the least upper bound of A is the "closest" upper bound of A.

Proposition 3.12. *Let K be an ordered set and $A \subseteq K$. If $\sup_K A$ exists, it is unique.*

Proof. See Abbott (this is a good exercise, so try it on your own). ■

Example 3.13. Even though the set $A := \{r \in \mathbb{Q} \ : \ r^2 < 2\}$ is bounded from above, A does not have a supremum in $\mathbb{Q}$.

This example also motivates the need for a number system that contains the least upper bound of a bounded set in $\mathbb{Q}$. But first, we look at one more example.

Example 3.14. Let

$$A := \left\{ \frac{m}{n} \ : \ m, n \in \mathbb{N} \text{ and } m < n \right\}. \tag{3.15}$$

The set of such elements looks like

	$m=1$	$m=2$	$m=3$	$m=4$
$n=1$				
$n=2$	$\frac{1}{2}$			
$n=3$	$\frac{1}{3}$	$\frac{2}{3}$		
$n=4$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{3}{4}$	
$n=5$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$

(3.16)

On the left column, we have a subset of rational numbers given by $\frac{1}{n}$ for all natural numbers $n > 1$. As n increases, these numbers tend to 0. On the diagonal, the rational numbers are increasing and are of the form $\frac{m}{m+1}$ for all natural numbers m. As m increases, these numbers tend to 1. Thus

$$\sup A = 1 \qquad \& \qquad \inf A = 0. \tag{3.17}$$

Exercise 3.18. Prove these claims (Hint: you will be better equipped to do this after Lemma 3.31).

Theorem 3.19. *There exists an ordered field $\mathbb{R}$ satisfying the conditions*

(a) every nonempty subset of $\mathbb{R}$ that is bounded above has a least upper bound and

(b) $\mathbb{R}$ contains $\mathbb{Q}$ as a subfield.

We will not prove this theorem until possibly after we have introduced Cauchy sequences in a few weeks. We will simply take this fact for granted. Condition (a) is referred to as the *Axiom of Completeness*. We will treat the ordered field $\mathbb{R}$ as we normally have done before you signed up for this class.

Remark 3.20. The ordered field $\mathbb{R}$ in Theorem 3.19 is unique in a certain sense. To state this uniqueness property rigorously, one needs to introduce the notion of a morphism of ordered fields. A morphism of ordered fields is a function from one field to another satisfying certain properties. If there are two fields $\mathbb{R}$ and $\mathbb{R}'$ satisfying the conditions of Theorem 3.19, then there exist a morphism $f : \mathbb{R} \to \mathbb{R}'$ and a morphism $g : \mathbb{R}' \to \mathbb{R}$ such that $f \circ g = \mathrm{id}_{\mathbb{R}'}$ and $g \circ f = \mathrm{id}_{\mathbb{R}}$. In other words, the two fields are isomorphic. In fact, the uniqueness of $\mathbb{R}$ is even better than this, but to explain that further, one should introduce limits in categories, which we probably will not, at least not yet. See https://ncatlab.org/nlab/show/real+number#the_complete_ordered_field and http://math.stackexchange.com/questions/839848/category-theoretic-description-of-the-real-numbers for further discussion.

Definition 3.21. A (finite) *interval* in $\mathbb{R}$ is a subset of $\mathbb{R}$ of one of the following forms

$$[a, b] := \{x \in \mathbb{R} \ : \ a \le x \le b\} \tag{3.22}$$
$$(a, b) := \{x \in \mathbb{R} \ : \ a < x < b\} \tag{3.23}$$
$$[a, b) := \{x \in \mathbb{R} \ : \ a \le x < b\} \tag{3.24}$$
$$(a, b] := \{x \in \mathbb{R} \ : \ a < x \le b\} \tag{3.25}$$

Here $a, b \in \mathbb{R}$ with $a \le b$. The first is called a *closed interval*, the second is called an *open interval*, and either of the last two is called a *half-open interval*.

Theorem 3.26 (Nested Interval Property)**.** *Let*

$$I_n := [a_n, b_n] \tag{3.27}$$

be a sequence of closed intervals in $\mathbb{R}$ *satisfying*

$$I_1 \supseteq I_2 \supseteq I_3 \supseteq \cdots . \tag{3.28}$$

Then

$$\bigcap_{n=1}^{\infty} I_n \ne \varnothing. \tag{3.29}$$

Proof. The goal is to construct an element of $\bigcap_{n=1}^{\infty} I_n$. First note that the set $A := \{a_1, a_2, a_3, \ldots\}$ is bounded above. In fact, every b_n serves as one such upper bound. By the Axiom of Completeness, $a := \sup A$ exists. We claim that $a \in \bigcap_{n=1}^{\infty} I_n$, i.e. $a \in I_n$ for all $n \in \mathbb{N}$. Since a is an upper bound of A, $a \ge a_n$. Furthermore, $a \le b_n$ by the universal property of least upper bounds (since $a_n \le b_n$ and $a_n \le a$, it follows that $a \le b_n$). Hence, $a \in I_n$. ■

It should be clear that there is no smallest positive rational number nor is there any largest rational number. For instance, to see the former, let $\frac{p}{q}$ be the supposed smallest number. Then $\frac{p}{q+1}$ is smaller still and rational. This is slightly less clear in $\mathbb{R}$ because we have added several new elements, but it is nevertheless true.

Definition 3.30. Let $\mathbb{R}^+$ denote the set of strictly positive real numbers, $\mathbb{R}^+ := \{r \in \mathbb{R} \ : \ r > 0\}$.

Lemma 3.31. *Let $A \subseteq \mathbb{R}$ be a bounded subset of $\mathbb{R}$ and let $s \in \mathbb{R}$ be such an upper bound. Then $s = \sup A$ if and only if for every $\epsilon \in \mathbb{R}^+$, there exists an element $a \in A$ such that $s - \epsilon < a$.*

Proof. See Lemma 1.3.8 in [2] (or Lemma 1.3.7 in [1]). ■

Proposition 3.32 (Arithmetic of sup). *Let A and B be nonempty bounded subsets of $\mathbb{R}$. Define the sets*

$$\begin{aligned} A + B &:= \{a + b \ : \ a \in A, b \in B\} \\ AB &:= \{ab \ : \ a \in A, b \in B\}. \end{aligned} \tag{3.33}$$

Then

$$\sup(A + B) = \sup(A) + \sup(B). \tag{3.34}$$

Furthermore, if A and B are subsets of nonnegative numbers, then

$$\sup(AB) = \sup(A)\sup(B). \tag{3.35}$$

Proof. The first part of this is listed as Exercise 1.3.6 in [2]. The second part is also left to the student as an exercise.[11] ■

After this lecture, it is recommended the student works through problems 2, 3, 4, 5, and 6 on HW #1. Additional recommended exercises include exercises 1.2.7, 1.3.5, 1.3.6, 1.3.7, 1.3.8, 1.3.11, and 1.4.4 from [2] (or equivalently exercises 1.2.6, 1.3.5, 1.3.9, 1.3.7, 1.3.6, 1.3.4, and N/A from [1]).

[11] These proofs are nontrivial!

4 September 8

Theorem 4.1 (Archimedean Property of $\mathbb{R}$)**.**

(a) For every $x \in \mathbb{R}$, there exists an $n \in \mathbb{N}$ satisfying $n > x$.

(b) For every $y \in \mathbb{R}^+$, there exists an $n \in \mathbb{N}$ satisfying $\frac{1}{n} < y$.

Abbott has a great discussion following this theorem, please read it.

Proof.

(a) Suppose, to the contrary, that $x \geq n$ for all $n \in \mathbb{N}$, i.e. $\mathbb{N}$ is a bounded subset of $\mathbb{R}$. By the Axiom of Completeness, a least upper bound of $\mathbb{N}$ exists. Hence, let $s := \sup \mathbb{N}$ in $\mathbb{R}$. Because $\mathbb{R}$ is an ordered field, $s - 1 \in \mathbb{R}$. Since $s - 1 < s$, $s - 1$ is not an upper bound of $\mathbb{N}$. Hence, there exists an $n \in \mathbb{N}$ such that $s - 1 < n$, i.e. $s < n + 1$, which shows that s cannot be an upper bound of $\mathbb{N}$, contradicting our assumption.

(b) By (a), there exists an $n \in \mathbb{N}$ such that $n > \frac{1}{y}$, i.e. $\frac{1}{n} < y$.

■

Theorem 4.2 (Density of $\mathbb{Q}$ and $\mathbb{I}$ in $\mathbb{R}$)**.**

(a) For any two real numbers $a, b \in \mathbb{R}$ with $a < b$, there exists a rational number $r \in \mathbb{Q}$ satisfying $a < r < b$.

(b) For any two real numbers $a, b \in \mathbb{R}$ with $a < b$, there exists an irrational number $t \in \mathbb{Q}$ satisfying $a < t < b$.

Proof.

(a) By assumption,[12] $b - a > 0$. By the Archimedian Property of $\mathbb{R}$ (Theorem 4.1), there exists an n satisfying

$$\frac{1}{n} < b - a. \tag{4.3}$$

It therefore suffices to find an integer m such that $na < m < nb$ or equivalently, $a < \frac{m}{n} < b$. Again by the Archimedean Property of $\mathbb{R}$, there exists an integer M satisfying $M > na$. Let m be the smallest such integer. Hence, $m - 1 \leq na$ and $na < m$. The first inequality can be rewritten as

$$\begin{aligned} m &\leq na + 1 \\ &< n\left(b - \frac{1}{n}\right) + 1 \qquad \text{by (4.3)} \\ &= nb. \end{aligned} \tag{4.4}$$

This proves $na < m < nb$ and hence $a < \frac{m}{n} < b$.

[12] The idea is to first find a rational number smaller than the spacing between a and b. Then, using this rational number, we will scale it by successive positive integers until the value lies within (a, b). Abbot has a nice picture for this.

(b) Note that

$$\frac{b}{\sqrt{2}} - \frac{a}{\sqrt{2}} = \frac{b-a}{\sqrt{2}} > 0 \tag{4.5}$$

since $b - a > 0$. Hence, by part (a), there exists a rational number $\frac{m}{n}$ satisfying

$$\frac{a}{\sqrt{2}} < \frac{m}{n} < \frac{b}{\sqrt{2}}. \tag{4.6}$$

Multiplying by $\sqrt{2}$ throughout gives

$$a < \frac{m}{n}\sqrt{2} < b. \tag{4.7}$$

The claim then follows from from the fact that $\frac{m}{n}\sqrt{2}$ is irrational.[13]

■

Theorem 4.8. [14] *For every positive real number $x \in \mathbb{R}^+$ and every natural number $n \in \mathbb{N}$, there exists a unique positive real number $y \in \mathbb{R}$ such that $y^n = x$. This number y is written as $x^{1/n}$ or $\sqrt[n]{x}$.*

Before we prove this, let's have a discussion.

(a) What are our data/assumptions? These are a real number $x > 0$ and a positive integer $n > 0$.

(b) What are our outputs/goals? This is the construction of a real number y.

(c) What are the conditions? The conditions required of y are that it satisfies $y^n = x$ and that there is no other $z > 0$ such that $z^n = x$.

The idea of the following proof will be to consider the set of all real numbers t such that $t^n < x$. If this set is nonempty and bounded from above, we can use the Axiom of Completeness for $\mathbb{R}$ to construct our desired y. Then we will have to prove that y satisfies the required condition $y^n = x$ and that there is no other z such that $z^n = x$. A proper and clear proof is given in Rudin's book [6]. The following proof hopefully will provide a little more intuition and explanation.

Proof of Theorem 4.8. Let

$$E := \{t \in \mathbb{R} \ : \ t^n < x\}. \tag{4.9}$$

E is nonempty because the number $t := \frac{x}{1+x}$ satisfies

$$0 < t < 1 \qquad \& \qquad t < x \tag{4.10}$$

[13] You should check this. In fact, you should do Exercise 1.4.2 in [1] (Exercise 1.4.1 in [2]), which is a more general fact.

[14] I was unable to prove Theorem 4.8 and Corollary 4.25 during the class time.

and therefore satisfies[15]

$$t^n < t < x. \tag{4.11}$$

Let $s := 1 + x$ so that $s > 1$ and $s > x$. Hence,[16] $s^n > s > x$ showing that s is an upper bound for E. Hence, by the Axiom of Completeness, the supremum exists. Let

$$y := \sup E. \tag{4.12}$$

To show that $y^n = x$, we will prove that $y^n < x$ and $y^n > x$ both lead to contradictions.

Assume $y^n < x$. The goal is to find an integer $m > 0$ such that $y + \frac{1}{m}$ is still in E, contradicting that y is an upper bound. To show the existence of such an integer m, we will work backwards to obtain a condition on m such that

$$\left(y + \frac{1}{m}\right)^n < x. \tag{4.13}$$

Note that subtracting y^n from this desired inequality turns it into

$$\left(y + \frac{1}{m}\right)^n - y^n < x - y^n. \tag{4.14}$$

Now, the right-hand-side is a positive quantity by assumption.[17] Meanwhile, the left-hand-side is of the form $b^n - a^n$ with $b > a > 0$, which can be factored into[18]

$$b^n - a^n = (b-a)\sum_{k=1}^{n} b^{n-k}a^{k-1} \tag{4.15}$$

Since $b > a$,

$$b^n - a^n < (b-a)\sum_{k=1}^{n} b^{n-k}b^{k-1} = (b-a)\sum_{k=1}^{n} b^{n-1} = (b-a)nb^{n-1}. \tag{4.16}$$

[15]It's not immediately obvious why this is true. We will prove this by induction by first proving the base case and then doing the induction step. By assumption, $0 < a < 1$, i.e. $0 < a$ and $a < 1$. Rearranging the latter inequality gives $0 < 1 - a$. By Axiom (b) of the definition of $\mathbb{R}$ being an ordered field, $0 \le a(1-a)$, which upon rearranging proves that $a^2 \le a$. Now, $a^2 \ne a$ because if this were true, dividing by a would give $a = 1$, which contradicts our initial assumption. Now suppose that $0 < a^n < a < 1$ for some integer n. Then by Axiom (b) of the definition of $\mathbb{R}$ being an ordered field, $0 \le a^n(1-a)$, which shows $a^{n+1} \le a^n < a$, proving the induction step. The reason this proof is so complicated is because we are technically assuming only that $\mathbb{R}$ is an ordered field containing $\mathbb{Q}$ as a subfield and there are many common facts we have not yet proved about $\mathbb{R}$. We will often avoid these technicalities for now, but the preceding argument indicates what one might do to proceed carefully.

[16]A similar proof to the previous footnote applies here by considering $\frac{1}{s}$.

[17]It would be nice if we could use the Archimedean Property here, but that would require an inequality of the form $\frac{1}{m}u < x - y^n$, where u only depends on x, y, and n but does not depend on m. We must therefore find an explicit expression for u to achieve this goal.

[18]You should check this yourself, but here is the calculation:

$$(b-a)\sum_{k=1}^{n} b^{n-k}a^{k-1} = \sum_{k=1}^{n} b^{n-(k-1)}a^{k-1} - \sum_{k=1}^{n} b^{n-k}a^{k} = b^n + \left(\sum_{k=1}^{n-1} b^{n-k}a^{k} - b^{n-k}a^{k}\right) - a^n = b^n - a^n.$$

In our case, this means[19]

$$\left(y+\frac{1}{m}\right)^n - y^n < \frac{n}{m}\left(y+\frac{1}{m}\right)^{n-1} < \frac{n(y+1)^{n-1}}{m} \tag{4.17}$$

as long as $m > 1$. In order to prove our desired inequality (4.14), we should choose m so that

$$\frac{n(y+1)^{n-1}}{m} < x - y^n \tag{4.18}$$

but rearranging this inequality gives

$$\frac{1}{m} < \frac{x-y^n}{n(y+1)^{n-1}}. \tag{4.19}$$

Such an integer $m > 1$ exists by the Archimedean Property of $\mathbb{R}$ since the right-hand-side is a positive quantity.

Now assume $y^n > x$. The goal is to find an integer $m > 0$ such that $y - \frac{1}{m}$ is still an upper bound for E, contradicting that y is the least such upper bound. Again, we work backwards restricting m to satisfy $\left(y - \frac{1}{m}\right)^n > x$. This time, multiplying by -1 and adding y^n to both sides gives

$$y^n - \left(y-\frac{1}{m}\right)^n < y^n - x. \tag{4.20}$$

This is again in the form $b^n - a^n$ so the same trick should work. Expanding out the left-hand-side and using the earlier inequality gives

$$y^n - \left(y-\frac{1}{m}\right)^n < \frac{n}{m}\left(y-\frac{1}{m}\right)^{n-1} < \frac{ny^{n-1}}{m} \tag{4.21}$$

as long as $m > 1$ *and* $y > \frac{1}{m}$ (once we show an m exists, we can choose a larger m if necessary so that this latter condition can be satisfied by using the Archimedean Property again). In order to prove our desired inequality (4.20), we should choose m so that

$$\frac{ny^{n-1}}{m} < y^n - x, \tag{4.22}$$

but rearranging this inequality gives

$$\frac{1}{m} < \frac{y^n - x}{ny^{n-1}}. \tag{4.23}$$

Such an integer $m > 1$ exists by the Archimedean Property of $\mathbb{R}$ since the right-hand-side is a positive quantity.

Finally, now that we have shown an n-th root exists, we must show it is unique. Suppose that y_1 and y_2 are two *positive* numbers that satisfy $y_1^n = y_2^n = x$. Then either $y_1 < y_2$, $y_2 < y_1$, or $y_1 = y_2$. The first is not possible because

$$y_1^n < y_1^{n-1}y_2 < \cdots < y_1y_2^{n-1} < y_2^n. \tag{4.24}$$

A similar argument shows that the second is impossible. Thus, $y_1 = y_2$ and the n-th root is unique. ■

[19] Realizing that the previous inequality is one way to prove our claim is the least obvious step in the proof and requires some insight. If you have a simpler proof or a clear explanation for how to easily see this, please let me know!

Corollary 4.25. *For every pair of positive real numbers $a, b \in \mathbb{R}^+$ and every positive integer $n \in \mathbb{N}$,*

$$(ab)^{1/n} = a^{1/n} b^{1/n}. \tag{4.26}$$

Proof. Since the n-th root of any positive real number exists by the previous theorem, it makes sense to set $\alpha := a^{1/n}$ and $\beta := b^{1/n}$. Then

$$\begin{tikzcd}[column sep=small] & \alpha^n\beta^n \arrow[dl, equal] \arrow[dr, equal] & \\ ab & & (\alpha\beta)^n \end{tikzcd} \tag{4.27}$$

Taking the n-th root of both ends gives

$$(ab)^{1/n} = \alpha\beta = a^{1/n} b^{1/n}. \tag{4.28}$$

By the uniqueness of such roots by the same theorem, the right-hand-side is the unique n-th root of ab. ■

Theorem 4.29.

(a) The set $\mathbb{Q}$ of rational numbers is countable.

(b) The set $\mathbb{R}$ of real numbers is uncountable.

Proof.

(a) Let $A_0 := \{0\}$ and set

$$A_n := \left\{ \pm\frac{p}{q} \;:\; \text{where } p, q \in \mathbb{N} \text{ have no common factors and satisfy } p + q = n \right\} \tag{4.30}$$

for each $n \in \mathbb{N}$. Note that the rational number $\frac{p}{q}$ (with p and q having no factors in common) appears in the set $A_{|p|+|q|}$ showing that

$$\mathbb{Q} \subseteq \bigcup_{n=0}^{\infty} A_n. \tag{4.31}$$

This calculation also shows that $\frac{p}{q}$ appears in A_n for a *unique* $n \in \mathbb{N}$. Hence, $A_n \cap A_m = \varnothing$ if $n \neq m$. Conversely, $A_n \subset \mathbb{Q}$ for all $n \in \mathbb{N}$ so that

$$\bigcup_{n=0}^{\infty} A_n \subseteq \mathbb{Q}. \tag{4.32}$$

This shows that

$$\bigcup_{n=0}^{\infty} A_n = \mathbb{Q}. \tag{4.33}$$

Since the left-hand-side is a countable union of finite sets, it is countable.[20]

[20] In fact, a countable union of countable sets is countable. This is claimed in Theorem 1.5.8 in [2] with a proof outlined in Exercise 1.5.3 in [2]. I recommend you work this proof out (you should have actually already done it since I suggested it on day 1). A proof of this is also explained in the references provided at the end of the first lecture.

(b) Suppose, to the contrary, that $\mathbb{R}$ is countable. Then there exists a bijection $f : \mathbb{N} \to \mathbb{R}$ and let $x_n := f(n)$ for all $n \in \mathbb{N}$. Let $I_1 := [a_1, b_1]$ be a finite closed interval that does not contain x_1. Suppose that $I_n := [a_n, b_n]$ has been inductively defined to satisfy the condition that $I_n \subseteq I_{n-1}$ and $x_n \notin I_n$. If $x_{n+1} \notin I_n$, then one can set $I_{n+1} = I_n$. If $x_{n+1} \in I_n$, then $a_n \leq x_{n+1} \leq b_n$ with a strict inequality on at least one side. Without loss of generality, suppose that $a_n < x_{n+1}$. Then by the Density of $\mathbb{Q}$ and $\mathbb{I}$ in $\mathbb{R}$, there exists a real number b_{n+1} satisfying $a_n < b_{n+1} < x_{n+1}$. Hence, set $I_{n+1} := [a_n, b_{n+1}]$. A similar argument can be done if $x_{n+1} < b_n$. Thus, the collection $\{I_n\}_{n \in \mathbb{N}}$ satisfies

$$I_{n+1} \subseteq I_n \qquad \& \qquad x_{n+1} \notin I_{n+1} \qquad \forall\, n \in \mathbb{N}. \tag{4.34}$$

By the Nested Interval Property,

$$J := \bigcap_{n=1}^{\infty} I_n \neq \varnothing. \tag{4.35}$$

Let x be such an element. By assumption that $\mathbb{R}$ is countable, there exists an $m \in \mathbb{N}$ such that $f(m) = x$. But by construction, $x \notin I_m$ and therefore cannot be in J, which is a contradiction.

■

After this lecture, it is recommended the student works through the rest of the problems on HW #1. Additional recommended exercises include 1.4.5, 1.4.6, and 1.4.7 from [2] (or equivalently exercises 1.4.3, N/A, and 1.4.6. from [1]). Any other exercises in Abbot that have not already been mentioned are also good. In particular, go over Cantor's argument for the uncountability of $\mathbb{R}$. Don't forget that HW #1 is due the next class!

5 September 13

Abbott has some great motivation for some of the weird properties possessed by infinite sums in Section 2.1. (you should read it). Given a countable set of real numbers $\{a_1, a_2, a_3, \dots\}$, what is the meaning of $\sum_n a_n$? Abbott discusses how the *order* in which the sum is taken may alter the value of such a sum so that it doesn't make sense to define the sum of the elements of an arbitrary countable *set* such as $\{a_1, a_2, a_3, \dots\}$.[21] However, perhaps the sum may be defined if the order is specified. This leads one to the notion of a sequence.

Definition 5.1. A *sequence* of real numbers is a function $\mathbb{N} \to \mathbb{R}$.

Definition 5.2. A set A is *at most countable* if A is either finite or countable.

Proposition 5.3. *A sequence determines, and is determined by, an at most countable subset $A \subseteq \mathbb{R}$ together with a bijection $\{1, \dots, N\} \to A$ if A is finite and has cardinality $N \in \mathbb{N}$ or $\mathbb{N} \to A$ if A is countable.*

Proof.
($\Rightarrow$) Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence. Then the image of a, namely $a(\mathbb{N})$, is an at most countable subset of $\mathbb{R}$.[22]
($\Leftarrow$) Let $A \subseteq \mathbb{R}$ be an at most countable subset of $\mathbb{R}$. If A is countable, let $a : \mathbb{N} \to A$ be a bijection. Since $A \subseteq \mathbb{R}$, this precisely defines a sequence $\mathbb{N} \to A \hookrightarrow \mathbb{R}$.[23] If A has cardinality $N \in \mathbb{N}$, let $a : \{1, \dots, N\} \to A$ be a bijection. Set

$$a(n) := a(N) \qquad \forall\, n > N. \tag{5.4}$$

This defines a function $a : \mathbb{N} \to A \hookrightarrow \mathbb{R}$ and hence a sequence. ∎

Hence, a sequence can be viewed as an at most countable subset of real numbers with additional data, namely a particular order in which the elements are specified. Note that the sequence associated to an at most countable subset of $\mathbb{R}$ is not unique. Because of Proposition 5.3, we can write a sequence $a : \mathbb{N} \to \mathbb{R}$ as $\{a_n\}_{n \in \mathbb{N}}$ or more abusively as $\{a_n\}$. It is also common to write a sequence as $(a_1, a_2, a_3, \dots)$ or more succinctly as $(a_n)_{n \in \mathbb{N}}$ or abusively as (a_n). I personally find most of these a bit ambiguous and will mostly use $a : \mathbb{N} \to \mathbb{R}$ or $(a_n)_{n \in \mathbb{N}}$ to avoid as much confusion as possible.

Definition 5.5. Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence[24] of real numbers, whose value at $n \in \mathbb{N}$ is written as a_n. The *partial sums* of a is the sequence $S : \mathbb{N} \to \mathbb{R}$ given by

$$\mathbb{N} \ni m \mapsto S_m := \sum_{n=1}^{m} a_n. \tag{5.6}$$

[21] Remember, elements in a set are not ordered in any particular way. The sets $\{\text{cat}, \text{dog}\}$ and $\{\text{dog}, \text{cat}\}$ are exactly the same.

[22] This is left as an exercise: Let A be a countable set, X a set, and $f : A \to X$ a function. Then $f(A)$ is an at most countable subset of X.

[23] The notation $A \hookrightarrow \mathbb{R}$ is used to describe the inclusion function associated to a subset $A \subseteq \mathbb{R}$ and is often used for more general subsets of arbitrary sets.

[24] An at most countable subset instead of a sequence would not be enough data to unambiguously define the partial sums.

If all the a_n are non-negative, one might try to define

$$\sum_{n=1}^{\infty} a_n \overset{?}{:=} \sup\{S_1, S_2, \dots\} \tag{5.7}$$

if the supremum is defined. In fact, we will define $\sum_{n=1}^{\infty} a_n$ more generally for arbitrary sequences and *prove* that (5.7) is true. This result will follow from the Monotone Convergence Theorem and will be the subject of the next lecture.

Definition 5.8. A sequence $a : \mathbb{N} \to \mathbb{R}$ *converges* to a real number $\lim a$ if for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_n - \lim a| < \epsilon \qquad \forall\, n \geq N. \tag{5.9}$$

If a sequence converges, it is said to be *convergent*. If a sequence does not converge to any real number, it is said to be *divergent*.

If we use the notation $(a_n)_{n\in\mathbb{N}}$ to describe a sequence, the limit is often written as $\lim\limits_{n\to\infty} a_n$. Note that N depends on ϵ in (5.9). If we wanted to be a bit more clear, we might write N_ϵ since a different choice of ϵ may require a different choice of N.

Proposition 5.10. *If a sequence $a : \mathbb{N} \to \mathbb{R}$ converges, then the number it converges to is unique.*

Proof. Suppose a converges to both real numbers x and y. Then for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_n - x| < \frac{\epsilon}{2} \qquad \& \qquad |a_n - y| < \frac{\epsilon}{2} \qquad \forall\, n \geq N \tag{5.11}$$

(because such an N exists for both x and y, we can take the largest of the two values). Hence, for any $\epsilon > 0$,

$$0 \leq |x - y| = |x - a_n + a_n - y| \leq |x - a_n| + |a_n - y| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \tag{5.12}$$

for all $n \geq N$ by the triangle inequality.[25] Thus $|x - y| = 0$ (by Lemma 3.31 for infima, which was part of your HW #1) so that $x = y$. ∎

Example 5.13. Abbott considers the sequence

$$\mathbb{N} \ni n \mapsto \frac{1}{\sqrt{n}} \tag{5.14}$$

and shows that it converges to 0. Please look at the proof and understand how given any $\epsilon > 0$, the required N_ϵ is constructed.

An example of a sequence that does *not* converge to any real number is the sequence

$$\mathbb{N} \ni n \mapsto (-1)^n. \tag{5.15}$$

Exercise 5.16. Show that

$$\lim_{n\to\infty} \ln\left(1 + \frac{1}{n}\right) = 0. \tag{5.17}$$

Recall, the logarithm to the base e was defined in HW #1.

[25] The triangle inequality was an exercise in Chapter 1 of Abbot.

Example 5.18. Fix $p > 1$. Consider the sequence of partial sums

$$\mathbb{N} \ni n \mapsto a_n := \sum_{k=1}^{n} \frac{1}{p^k} \tag{5.19}$$

and find the limit $\lim_{n\to\infty} a_n$. To do this, it helps to write out the first few such partial sums

$$a_1 = \frac{1}{p}, \quad a_2 = \frac{p+1}{p^2}, \quad a_3 = \frac{p^2+p+1}{p^3}, \quad a_4 = \frac{p^3+p^2+p+1}{p^4} \tag{5.20}$$

and more generally

$$a_n = \frac{p^{n-1}+p^{n-2}+\cdots+p+1}{p^n}, \tag{5.21}$$

which can be rewritten as

$$a_n = \frac{(p-1)(p^{n-1}+p^{n-2}+\cdots+p+1)}{(p-1)p^n} = \frac{p^n-1}{(p-1)p^n} \tag{5.22}$$

(the second equality follows from a formula we used in Lecture 4). We claim (as you might now guess)

$$\lim_{n\to\infty} a_n = \frac{1}{p-1}. \tag{5.23}$$

To check this, we take the difference

$$\left| a_n - \frac{1}{p-1} \right| = \frac{1}{(p-1)p^n} \tag{5.24}$$

Given an $\epsilon > 0$, we want to find an $N \in \mathbb{N}$ such that

$$\frac{1}{(p-1)p^n} < \epsilon \tag{5.25}$$

for all $n \geq N$. Hence, we want to solve this equation for n. Rearranging gives

$$\frac{1}{\epsilon(p-1)} < p^n. \tag{5.26}$$

Applying the base p logarithm (which was defined in HW #1) gives

$$\log_p\left(\frac{1}{\epsilon(p-1)}\right) < n. \tag{5.27}$$

By the Archimedean Property of $\mathbb{R}$, there exists an N_ϵ such that

$$\log_p\left(\frac{1}{\epsilon(p-1)}\right) < N_\epsilon. \tag{5.28}$$

Hence, we have found the desired $N_\epsilon \in \mathbb{N}$. We should check that this N_ϵ in fact works. Therefore, let $\epsilon > 0$ be fixed. Then, choosing N_ϵ as in (5.28),

$$\begin{aligned} \left| a_n - \frac{1}{p-1} \right| &= \frac{1}{(p-1)p^n} \\ &\leq \frac{1}{(p-1)p^{\log_p\left(\frac{1}{\epsilon(p-1)}\right)}} \qquad \forall\, n \geq N_\epsilon \\ &= \epsilon \end{aligned} \tag{5.29}$$

proving (5.23).

Proving this result would have been much simpler if we could prove that

$$\lim_{n\to\infty} \frac{p^n - 1}{(p-1)p^n} = \frac{\lim\limits_{n\to\infty} \left(1 - \dfrac{1}{p^n}\right)}{\lim\limits_{n\to\infty} (p-1)}. \tag{5.30}$$

This is in fact true, but we need to prove that we can manipulate limits in this fashion. Before stating the general fact (Theorem 5.35), we need a definition and a result to help us.

Definition 5.31. A sequence $a : \mathbb{N} \to \mathbb{R}$ is <u>*bounded*</u> if there exists a real number $M > 0$ such that $|a_n| \leq M$ for all $n \in \mathbb{N}$ (equivalently, if the subset $a(\mathbb{N})$ is bounded above by M and below by $-M$).

Lemma 5.32. *Every convergent sequence is bounded.*

Proof. Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence that converges with limit $\lim a \in \mathbb{R}$. Fix some $\epsilon > 0$. Then, by definition of a converging to $\lim a$, there exists an $N \in \mathbb{N}$ such that

$$|a_n - \lim a| < \epsilon \qquad \forall\, n \geq N \tag{5.33}$$

Set

$$M := \max\big\{|a_1|, |a_2|, \ldots, |a_{N-1}|, |\lim a| + \epsilon\big\}. \tag{5.34}$$

Then the sequence a is bounded by M. ■

Theorem 5.35 (Algebraic Limit Theorem for Sequences)**.** *Let $a, b : \mathbb{N} \to \mathbb{R}$ be two sequences. Then*

(a) $\lim\limits_{n\to\infty} (xa_n) = x \lim\limits_{n\to\infty} a_n$ *for all* $x \in \mathbb{R}$,

(b) $\lim\limits_{n\to\infty} (a_n + b_n) = \lim\limits_{n\to\infty} a_n + \lim\limits_{n\to\infty} b_n$,

(c) $\lim\limits_{n\to\infty} (a_n b_n) = \left(\lim\limits_{n\to\infty} a_n\right)\left(\lim\limits_{n\to\infty} b_n\right)$, *and*

(d) $\lim\limits_{n\to\infty} \left(\dfrac{a_n}{b_n}\right) = \dfrac{\lim\limits_{n\to\infty} a_n}{\lim\limits_{n\to\infty} b_n}$, *provided that*[26] $\lim\limits_{n\to\infty} b_n \neq 0$.

We will prove this in the next lecture.

Exercise 5.36. (Challenge). Make rigorous sense of the expression

$$\sqrt{1 - \sqrt{2 - \sqrt{3 - \sqrt{4 - \sqrt{5 \cdots}}}}} \tag{5.37}$$

and determine if it has a (real) numerical value. You do not need to determine this value if it exists. Email your solution to Professor Stuart Sidney by 3:00 PM September 22, 2016 for a prize.

After this lecture, it is recommended the student works through problems 1, 2, and 3 on HW #2. You may use the Algebraic Limit Theorem for any problems unless it is explicitly stated in the problem that you should not. Additional recommended exercises include exercises 2.2.4, 2.2.5, 2.2.7, 2.3.1, 2.3.6, 2.3.7, 2.3.9, 2.3.10, and 2.3.12 in [2] (I'll upload the corresponding exercise in [1] shortly).

[26] If $b_n = 0$ for some n, the meaning of $\frac{a_n}{b_n}$ is meant only for sufficiently large n, so that $b_n \neq 0$.

6 September 15

Proof of the Algebraic Limit Theorem.

(a) Notice that the claim is obviously true when $x = 0$.[27] Hence, suppose that $x \neq 0$. Let $\epsilon > 0$. By assumption, there exists an $N \in \mathbb{N}$ such that[28]

$$|a_n - \lim a| < \frac{\epsilon}{|x|} \qquad \forall\, n \geq N. \tag{6.1}$$

Multiplying throughout by $|x|$ gives

$$|x||a_n - \lim a| = |xa_n - x \lim a| < \epsilon \qquad \forall\, n \geq N. \tag{6.2}$$

(b) Let $\epsilon > 0$. By assumption, there exist $N_a, N_b \in \mathbb{N}$ such that[29]

$$|a_n - \lim a| < \frac{\epsilon}{2} \qquad \forall\, n \geq N_a \qquad \& \qquad |b_n - \lim b| < \frac{\epsilon}{2} \qquad \forall\, n \geq N_b. \tag{6.3}$$

Let $N := \max\{N_a, N_b\}$. Then, by the triangle identity

$$\begin{aligned} \big|(a_n + b_n) - (\lim a + \lim b)\big| &= \big|a_n - \lim a + b_n - \lim b\big| \\ &\leq |a_n - \lim a| + |b_n - \lim b| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \qquad \forall\, n \geq N \\ &= \epsilon \qquad \forall\, n \geq N. \end{aligned} \tag{6.4}$$

(c) First consider the case $\lim b \neq 0$. Since convergent sequences are bounded (Lemma 5.32), there exists an $M > 0$ such that

$$|a_n| \leq M \qquad \forall\, n \in \mathbb{N}. \tag{6.5}$$

Let $\epsilon > 0$. Since $(b_n)_{n \in \mathbb{N}}$ converges to $\lim b$, there exists an $N_b \in \mathbb{N}$ such that

$$|b_n - \lim b| < \frac{\epsilon}{2M} \qquad \forall\, n \geq N_b. \tag{6.6}$$

Since $(a_n)_{n \in \mathbb{N}}$ converges to $\lim a$, there exists an $N_a \in \mathbb{N}$ such that

$$|a_n - \lim a| < \frac{\epsilon}{2|\lim b|} \qquad \forall\, n \geq N_a. \tag{6.7}$$

[27] I normally dislike saying that things are 'obvious' because it presupposes some superiority. Nevertheless, this really is obvious if you think about it: the sequence constructed is constant, so the only thing to show is that a constant sequence converges to that same constant. It's good practice to check your understanding of the definition of convergence by writing these few details out explicitly.

[28] The choice of $\frac{\epsilon}{|x|}$ was made by first thinking about the desired conclusion, namely, the calculation in (6.2).

[29] Again, the choice of $\frac{\epsilon}{2}$ was made by first thinking about the desired conclusion, namely, the calculation in (6.4). In practice, one works backwards. This should not be too unfamiliar. When you learn about integrals in calculus, usually you think about functions whose derivatives are what you start with. With practice, you get better and better at 'guessing.'

Putting these two conclusions together, using the triangle inequality, and setting $N := \max\{N_a, N_b\}$ gives

$$
\begin{aligned}
|a_n b_n - \lim a \lim b| &= \big|a_n b_n - a_n \lim b + a_n \lim b - \lim a \lim b\big| \\
&= \big|a_n(b_n - \lim b) + \lim b(a_n - \lim a)\big| \\
&\le \big|a_n(b_n - \lim b)\big| + \big|\lim b(a_n - \lim a)\big| \\
&= |a_n||b_n - \lim b| + |\lim b||a_n - \lim a| \\
&< M\frac{\epsilon}{2M} + |\lim b|\frac{\epsilon}{2|\lim b|} \qquad \forall\, n \ge N \\
&= \epsilon \qquad \forall\, n \ge N.
\end{aligned}
\tag{6.8}
$$

Now consider the case $\lim b = 0$. Let M be as above. Then, since $\lim b = 0$, there exists an $N' \in \mathbb{N}$ such that

$$
|b_n| < \frac{\epsilon}{M} \qquad \forall\, n \ge N'. \tag{6.9}
$$

Therefore,

$$
|a_n b_n| = |a_n||b_n| < M\frac{\epsilon}{M} = \epsilon \qquad \forall\, n \ge N'. \tag{6.10}
$$

(d) Because $\lim b \ne 0$, there exists an $N_1 \in \mathbb{N}$ and a $\delta > 0$ such that

$$
|b_n| > \delta \qquad \forall\, n \ge N_1. \tag{6.11}
$$

Since $(b_n)_{n\in\mathbb{N}}$ converges, there exists an $N_2 \in \mathbb{N}$ such that

$$
|b_n - \lim b| < \epsilon\delta \lim b \qquad \forall\, n \ge N_2. \tag{6.12}
$$

Setting $N := \max\{N_1, N_2\}$ and putting these two together gives

$$
\begin{aligned}
\left|\frac{1}{b_n} - \frac{1}{\lim b}\right| &= \left|\frac{\lim b - b_n}{b_n \lim b}\right| \\
&= \frac{1}{|b_n|}\frac{1}{|\lim b|}|b_n - \lim b| \\
&< \frac{1}{\delta}\frac{1}{|\lim b|}\epsilon\delta \lim b \qquad \forall\, n \ge N \\
&= \epsilon \qquad \forall\, n \ge N,
\end{aligned}
\tag{6.13}
$$

which proves that

$$
\lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{\lim b}. \tag{6.14}
$$

The claim $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = \frac{\lim a}{\lim b}$ then follows from the fact that $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = \lim\limits_{n\to\infty} \left(a_n \frac{1}{b_n}\right)$ and part (c) of this Theorem.

■

Theorem 6.15. *[Order Limit Theorem] Let $a, b : \mathbb{N} \to \mathbb{R}$ be two convergent sequences with limits $\lim a$ and $\lim b$, respectively, satisfying $a_n \le b_n$ for all $n \in \mathbb{N}$. Then $\lim a \le \lim b$.*

Proof. See Abbott. ∎

Definition 6.16. A sequence $a : \mathbb{N} \to \mathbb{R}$ is *non-decreasing* if $a_{n+1} \geq a_n$ for all $n \in \mathbb{N}$ and *non-increasing* if $a_{n+1} \leq a_n$ for all $n \in \mathbb{N}$. A sequence $a : \mathbb{N} \to \mathbb{R}$ is *monotone* if it is either non-decreasing or non-increasing.

Theorem 6.17 (Monotone Convergence Theorem)**.** *If a sequence $a : \mathbb{N} \to \mathbb{R}$ is monotone and bounded, then it converges. In fact, if a is non-decreasing, then $\lim a = \sup\{a_1, a_2, \dots\}$. If a is non-increasing, the $\lim a = \inf\{a_1, a_2, \dots\}$.*

Note that the claim in the proof makes sense because the supremum and infimum of a bounded set always exists in $\mathbb{R}$ by the Axiom of Completeness. Also note that this confirms our earlier guess from the beginning of lecture 5 about how we should define infinite sums whose terms are all non-negative.

Proof of Theorem 6.17. We will prove the statement for a non-decreasing. Set $s := \sup\{a_1, a_2, \dots\}$ and let $\epsilon > 0$. Since s is the supremum, $s - \epsilon$ is not an upper bound of $a(\mathbb{N})$. Hence, there exists an $N \in \mathbb{N}$ such that $s - \epsilon < a_N$. Since a is non-decreasing, $a_N \leq a_n$ for all $n \geq N$. Since s is an upper bound of $a(\mathbb{N})$, $a_n \leq s$ for all $n \in \mathbb{N}$. Putting all these inequalities together gives

$$s - \epsilon < a_N \leq a_n \leq s < s + \epsilon \qquad \forall\, n \geq N. \tag{6.18}$$

Subtracting s from every term gives

$$-\epsilon < a_n - s < \epsilon \qquad \forall\, n \geq N, \tag{6.19}$$

i.e.

$$|a_n - s| < \epsilon \qquad \forall\, n \geq N. \tag{6.20}$$

Hence $\lim a = \sup\{a_1, a_2, \dots\}$. ∎

Definition 6.21. Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence and let $S : \mathbb{N} \to \mathbb{R}$ be the associated sequence of partial sums (see Definition 5.5). Then the expression[30]

$$\sum_{n=1}^{\infty} a_n := \lim S \tag{6.22}$$

is called an *infinite series*. The infinite series $\sum_{n=1}^{\infty} a_n$ is said to *converge* if $\lim S$ exists.

Example 6.23. For every $n \in \mathbb{N}$, let $a_n := \frac{1}{n^2}$. Then the infinite series

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \tag{6.24}$$

associated to a converges. This follows from a slick calculation showing that the partial sums are bounded by 2 (see Abbott). Because the partial sum sequence is non-decreasing, the Monotone

[30]There is some abuse of notation in (6.22) because the limit $\lim S$ need not exist.

Convergence Theorem shows that this sequence converges. Calculating the limit is another story though and is quite involved. Time permitting, we may prove

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \tag{6.25}$$

later in this course.

After this lecture, it is recommended the student works through problems 4 and 5 on HW #2. Additional recommended exercises include exercises 2.4.1, 2.4.6, and 2.4.7 in [2] (I'll upload the corresponding exercise in [1] shortly).

7 September 20

Last week, we learned that convergent sequences are bounded. The converse, however, is false: a bounded sequence need not converge. Abbott provides several examples. Nevertheless, a bounded sequence does have convergent *subsequences*. But before we discuss subsequences, let us talk about sequences that we know converge, but for which we might not know the limit. Recall, to show that a sequence $a : \mathbb{N} \to \mathbb{R}$ converges, we needed to know the limit $\lim a$ to prove that for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $|a_n - \lim a| < \epsilon$ for all $n \geq N$. If we did not know this limit, given our current state, we might not be able to prove it converges. The only method we have to prove it converges to *something* is if we prove it is bounded and monotone. These are somewhat stringent conditions! But there are other ways. What if the distance between elements far out in the sequence tends to zero very quickly? This is the motivation behind Cauchy sequences.[31]

Definition 7.1. A sequence $a : \mathbb{N} \to \mathbb{R}$ is a <u>*Cauchy sequence*</u> if for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_n - a_m| < \epsilon \qquad \forall\, n, m \geq N. \tag{7.2}$$

Cauchy sequences are closely related to convergent sequences.

Theorem 7.3. *Every convergent sequence is a Cauchy sequence.*

Proof. This is one of your homework problems. ∎

Cauchy sequences are incredibly useful because of the following phenomenal fact showing that the converse is true.

Theorem 7.4 (Cauchy Criterion)**.** *A sequence converges if and only if it is a Cauchy sequence.*

Remark 7.5. The sequence must take values in $\mathbb{R}$ or a suitable space for this theorem to hold. When/if you learn about sequences in arbitrary topological spaces, this theorem may fail. Nevertheless, it is true in a plethora of topological spaces.

The usefulness of this theorem is paramount. It can be used to prove that certain solutions to rather complicated partial differential equations exist. More generally, it can be used to prove many interesting properties of operators in analysis, which themselves tell us important results in quantum mechanics. The reason is because if you have an *approximate* solution to some problem but you do not know the actual solution, and you can construct a *sequence* of solutions that are better and better approximations, you might be able to check if these solutions converge to something by checking that this sequence (of functions) is Cauchy. If they are, then by Theorem 7.4, that means the solution exists! We will prove Theorem 7.4 today after introducing several concepts and facts, which may have otherwise seemed un-motivating.

Definition 7.6. Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence. A <u>*subsequence*</u> of a is a sequence of the form $\mathbb{N} \xrightarrow{f} \mathbb{N} \xrightarrow{a} \mathbb{R}$, where $f : \mathbb{N} \to \mathbb{N}$ is a non-decreasing one-to-one sequence of natural numbers.The value of $a \circ f$ at $n \in \mathbb{N}$ is often written as $a_{f(n)}$.

[31] We are doing this slightly out of order with Abbott's book. Instead of talking about subsequences (section 1.5), we are first going over Cauchy sequences (section 1.6) for motivation.

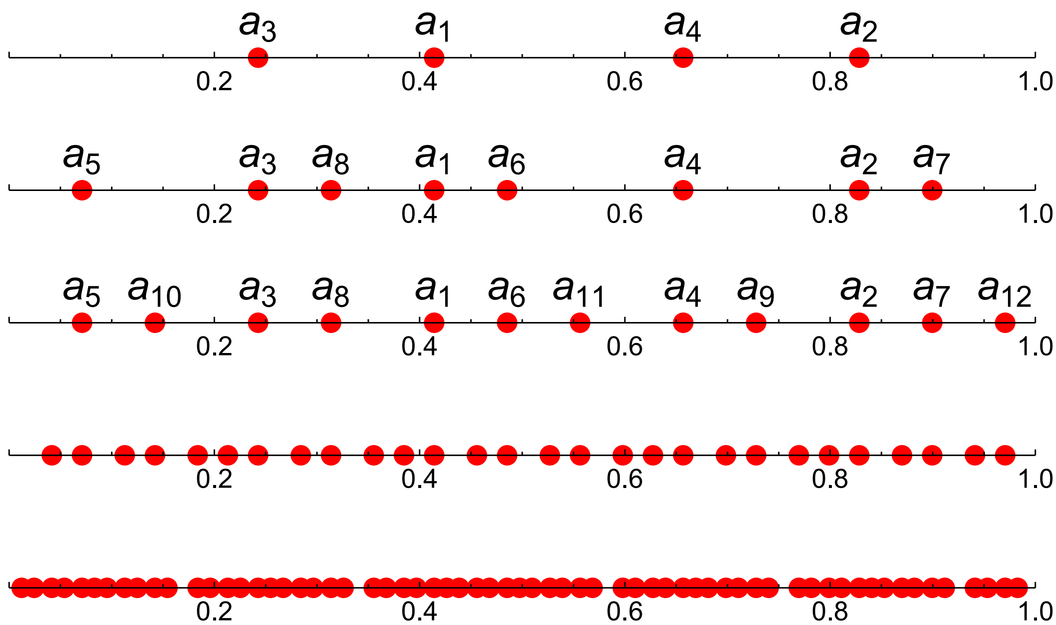


Figure 2: Visualizations of the first n elements in the sequence a for $n = 4, 8, 12, 28$, and 64, with $\alpha = \sqrt{2} - 1$ from Example 7.7.

Example 7.7. Fix $\alpha \in [0,1] \cap \mathbb{I}$. Let $a : \mathbb{N} \to \mathbb{R}$ be the sequence defined by

$$\mathbb{N} \ni n \mapsto a_n := n\alpha - \lfloor n\alpha \rfloor. \tag{7.8}$$

The first few elements of this sequence are depicted in Figure 2. On the one hand, this sequence seems to be "filling up" the entire interval. On the other hand, the successive elements in the sequence jump around somewhat sporadically. An amazing fact about this sequence is the following: "For any $r \in [0,1]$, there exists a subsequence $b : \mathbb{N} \to \mathbb{R}$ of a such that $\lim b = r$." We won't prove this—for a proof, take a course in ergodic theory. If you think about it, it's nuts! We will see a similar example when we construct $\mathbb{R}$ from $\mathbb{Q}$ next week (you might see a similar example in your homework!).

Remark 7.9. The previous example shows up in ergodic theory, which includes the study of systems of large numbers of particles such as gasses and fluids. It is a simple example that illustrates some features of ergodic systems.

Theorem 7.10. *Subsequences of a convergent sequence converge to the same limit as the original sequence.*

Proof. See Abbott. ∎

Theorem 7.11. *[Bolzano-Weierstrass Theorem] Every bounded sequence contains a convergent subsequence.*

Proof. Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence bounded by M. Thus, $|a_n| \leq M$ for all $n \in \mathbb{N}$. Then at least one of the intervals $[-M, 0]$ and $[0, M]$ contains an infinite number of terms in the sequence a.

Choose one of these intervals, denote it by I_1, and let $f_1 : \mathbb{N} \to \mathbb{N}$ be the non-decreasing 1-to-1 function corresponding to the subsequence of a all of whose elements are contained in I_1, i.e. $a \circ f_1 : \mathbb{N} \to \mathbb{R}$ with $(a \circ f_1)(\mathbb{N}) \subset I_1$. Similarly, let I_2 be an interval of length $\frac{M}{2}$ in I_1 such that one of the endpoints of I_2 agrees with an endpoint of I_1 and contains an infinite number of terms in the sequence $a \circ f_1$. Let $f_2 : \mathbb{N} \to \mathbb{N}$ be the non-decreasing 1-to-1 function corresponding to this subsequence, $a \circ f_1 \circ f_2$. Inductively, let I_n be the closed interval of length $\frac{M}{2^{n+1}}$ satisfying the conditions

(a) one of the endpoints of I_n agrees with one of the endpoints of I_{n-1} and

(b) I_n contains an infinite number of terms in the sequence $a \circ f_1 \circ \cdots \circ f_{n-1}$.

Let $f_n : \mathbb{N} \to \mathbb{N}$ be the non-decreasing 1-to-1 function corresponding to the infinite subsequence $a \circ f_1 \circ \cdots \circ f_{n-1} \circ f_n$ of a contained in I_n. Define a new sequence $\alpha : \mathbb{N} \to \mathbb{R}$ by

$$\mathbb{N} \ni n \mapsto \alpha(n) := \Big(a \circ f_1 \circ \cdots \circ f_{n-1} \circ f_n\Big)(1) \equiv a_{f_1(\cdots(f_{n-1}(f_n(1)))\cdots)}. \tag{7.12}$$

Then α is a subsequence of a. To show this sequence converges, note that

$$\bigcap_{n=1}^{\infty} I_n \neq \varnothing \tag{7.13}$$

by the Nested Interval Property. In fact, since these intervals are such that there is only a single element in this intersection,[32] call it x. Let $\epsilon > 0$ and choose $N \in \mathbb{N}$ such that

$$N > \log_2\left(\frac{M}{\epsilon}\right) + 1. \tag{7.14}$$

Then, because

$$\begin{aligned} |\alpha_n - x| &\le \frac{M}{2^{n-1}} && \text{since } \alpha_n, x \in I_n \\ &< \frac{M}{2^{\log_2\left(\frac{M}{\epsilon}\right)+1-1}} && \forall\, n \ge N \\ &= \epsilon \qquad \forall\, n \ge N, \end{aligned} \tag{7.15}$$

the sequence α converges to x. ■

Lemma 7.16. *Cauchy sequences are bounded.*

The proof is similar to the proof that convergent sequences are bounded.

Proof. Let $a : \mathbb{N} \to \mathbb{R}$ be a Cauchy sequence and let $\epsilon > 0$. By definition of a being a Cauchy sequence, there exists an $N \in \mathbb{N}$ such that

$$|a_n - a_m| < \epsilon \qquad \forall\, n, m \ge N. \tag{7.17}$$

Let

$$S := \max\left\{|a_1|, |a_2|, \dots, |a_{N-1}|, |a_N| + \epsilon\right\}. \tag{7.18}$$

Then $|a_n| \le S$ for all $n \in \mathbb{N}$ showing that a is bounded. ■

[32] To prove this, suppose that $x, y \in \bigcap_{n=1}^{\infty} I_n$ with $x \neq y$ and show that there is a contradiction.

We can now prove the Cauchy Criterion.

Proof of Theorem 7.4.
($\Rightarrow$) If a sequence converges, it is Cauchy by Theorem 7.3.
($\Leftarrow$) Let $a : \mathbb{N} \to \mathbb{R}$ be a Cauchy sequence. By Lemma 7.16, a is bounded. By the Bolzano-Weierstrass Theorem (Theorem 7.11), there exists a convergent subsequence, $a \circ f$, of a with $f : \mathbb{N} \to \mathbb{N}$ non-decreasing and 1-to-1. Denote the limit of this subsequence by

$$x := \lim(a \circ f). \tag{7.19}$$

The claim is that $x = \lim a$. To prove this, let $\epsilon > 0$. Since a is a Cauchy sequence, there exists an $N_a \in \mathbb{N}$ such that

$$|a_n - a_m| < \frac{\epsilon}{2} \qquad \forall\, n, m \geq N_a. \tag{7.20}$$

Since α is a converging sequence, there exists an $N_\alpha \in \mathbb{N}$ such that

$$|a_{f(n)} - x| < \frac{\epsilon}{2} \qquad \forall\, n \geq N_\alpha. \tag{7.21}$$

Let

$$N := \max\{N_a, N_\alpha\}. \tag{7.22}$$

Then, by the triangle inequality

$$\begin{aligned} |a_n - x| &= |a_n + a_{f(n)} - a_{f(n)} - x| \\ &\leq |a_n - a_{f(n)}| + |a_{f(n)} - x| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \qquad \forall\, n \geq N \\ &= \epsilon \qquad \forall\, n \geq N. \end{aligned} \tag{7.23}$$

The third line follows from the fact that $f(n) \geq n$. ■

After this lecture, it is recommended the student works through problems 6 and 7 on HW #2. Additional recommended exercises include exercises 2.5.2, 2.5.3, 2.5.9, 2.6.2, and 2.6.4 in [2] (I'll upload the corresponding exercise in [1] shortly).

8 September 22

We can now discuss many aspects of series that have motivated our discussions of sequences. Some of these facts follow [6].

Theorem 8.1. *The series associated to the sequence*

$$\mathbb{N} \ni n \mapsto \frac{1}{(n-1)!}, \tag{8.2}$$

where $n! := \prod_{k=1}^{n} k$ *and* $0! := 1$, *converges. This limit is denoted by*

$$e := \sum_{n=0}^{\infty} \frac{1}{n!} \tag{8.3}$$

and is called <u>Euler's constant</u>.

Proof. The partial sums S are given by

$$\begin{aligned} \mathbb{N} \ni n \mapsto S(n) &:= \sum_{k=0}^{n} \frac{1}{k!} \\ &= 1 + \frac{1}{1} + \frac{1}{1\cdot 2} + \frac{1}{1\cdot 2\cdot 3} + \frac{1}{1\cdot 2\cdot 3\cdot 4} + \cdots + \frac{1}{1\cdot 2\cdot \cdots \cdot n} \\ &< 1 + \frac{1}{2^0} + \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots + \frac{1}{2^{n-1}} \\ &= 1 + 1 + \frac{2^{n-1}-1}{(2-1)2^{n-1}} \\ &= 3 - \frac{1}{2^{n-1}} \\ &< 3 \end{aligned} \tag{8.4}$$

by Example 5.18, where we looked into the geometric series. Because S is bounded and increasing, it converges by the Monotone Convergence Theorem (Theorem 6.17). ■

Theorem 8.5 (Cauchy Criterion for Series)**.** *Let* $a : \mathbb{N} \to \mathbb{R}$ *be a sequence. The series* $\sum_{n=1}^{\infty} a_n$ *converges if and only if for any* $\epsilon > 0$, *there exists an* $N \in \mathbb{N}$ *such that*

$$|a_{m+1} + a_{m+2} + \cdots + a_n| < \epsilon \qquad \forall\, n, m \in \mathbb{N} \text{ with } n > m \geq N. \tag{8.6}$$

Proof. Let $S : \mathbb{N} \to \mathbb{R}$ be the associated sequence of partial sums of a. Then S converges if and only if S is Cauchy by the Cauchy Criterion for sequences (Theorem 7.4). By definition, S is Cauchy if and only if for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$|S_n - S_m| < \epsilon \qquad \forall\, n, m \geq N. \tag{8.7}$$

But the left-hand-side of this inequality is precisely

$$|S_n - S_m| = |a_{m+1} + a_{m+2} + \cdots + a_n| < \epsilon \qquad \forall\, n, m \in \mathbb{N} \text{ with } n > m \geq N, \tag{8.8}$$

supposing, without loss of generality, that $n > m$. ■

Theorem 8.9 (Absolute Convergence Test)**.** *Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence. If the series $\sum_{n=1}^{\infty} |a_n|$ converges, then the series $\sum_{n=1}^{\infty} a_n$ converges.*

Proof. By the Cauchy Criterion for Series (Theorem 8.5), $\sum_{n=1}^{\infty} |a_n|$ converges if and only if for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$\Big| |a_{m+1}| + |a_{m+2}| + \cdots + |a_n| \Big| < \epsilon \qquad \forall\, n, m \in \mathbb{N} \text{ with } n > m \geq N. \tag{8.10}$$

But the left-hand-side of this inequality satisfies

$$\begin{aligned} |a_{m+1} + a_{m+2} + \cdots + a_n| &\leq |a_{m+1}| + |a_{m+2}| + \cdots + |a_n| \\ &= \Big| |a_{m+1}| + |a_{m+2}| + \cdots + |a_n| \Big| \\ &< \epsilon \qquad \forall\, n, m \in \mathbb{N} \text{ with } n > m \geq N \end{aligned} \tag{8.11}$$

by the triangle inequality. Again, by the Cauchy Criterion for Series, $\sum_{n=1}^{\infty} a_n$ converges. ■

This theorem motivates the following definition.

Definition 8.12. Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence. The associated series $\sum_{n=1}^{\infty} a_n$ *converges absolutely* if the series $\sum_{n=1}^{\infty} |a_n|$ converges. The series $\sum_{n=1}^{\infty} a_n$ *converges conditionally* if the series $\sum_{n=1}^{\infty} |a_n|$ diverges but $\sum_{n=1}^{\infty} a_n$ converges.

Theorem 8.13. *Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence. If the series $\sum_{n=1}^{\infty} a_n$ associated to a converges, then* $\lim a = 0$.

Proof. By the Cauchy Criterion for Series (Theorem 8.5), for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_{m+1} + a_{m+2} + \cdots + a_n| < \epsilon \qquad \forall\, n, m \in \mathbb{N} \text{ with } n > m \geq N. \tag{8.14}$$

In particular, when $n := m + 1$, this implies

$$|a_{m+1}| < \epsilon \qquad \forall\, m \geq N \tag{8.15}$$

so that $\lim a = 0$. ■

The converse of this statement is not true! We will give an example later. However, the converse is true if an additional sufficient condition holds.

Theorem 8.16 (Alternating Series Test)**.** *Let $a : \mathbb{N} \to \mathbb{R}$ be a non-increasing sequence with* $\lim a = 0$. *Then the series*

$$\sum_{n=1}^{\infty} (-1)^n a_n \tag{8.17}$$

converges.

Proof. Exercise 2.7.1 in [2] asks the reader to supply *three* different proofs of this theorem. You should try to produce them all! ■

Definition 8.18. Let $a : \mathbb{N} \to \mathbb{R}$ be a sequence. A *rearrangement* of a is a sequence of the form $a \circ f$ with $f : \mathbb{N} \to \mathbb{N}$ a bijection.

Warning: If $f : \mathbb{N} \to \mathbb{N}$ is a bijection and $a : \mathbb{N} \to \mathbb{R}$ converges, this does *not* mean that $a \circ f$ converges! In fact, you may have guessed this by the examples provided in Section 2.1 of Abbott's book. This is particularly relevant for series associated to a whose partial sums are given by

$$\mathbb{N} \ni n \mapsto S(n) := \sum_{k=1}^{n} a_k. \tag{8.19}$$

The partial sums associated to $a \circ f$ are given instead by

$$\mathbb{N} \ni n \mapsto S'(n) := \sum_{k=1}^{n} a_{f(k)}. \tag{8.20}$$

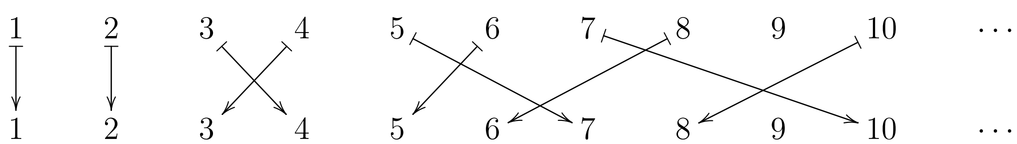
Example 8.21. Let $a : \mathbb{N} \to \mathbb{R}$ be the sequence given by

$$\mathbb{N} \ni n \mapsto a_n := \frac{(-1)^{n+1}}{n} \tag{8.22}$$

and let S be the associated sequence of partial sums. By the Alternating Series Test (Theorem 8.16), S converges. In fact, it converges to $\ln 2$ where $\ln := \log_e$ and e is Euler's constant defined earlier in this lecture. Let S denote the sequence of partial sums of the sequence a. Let $f : \mathbb{N} \to \mathbb{N}$ be the function defined by

$$\mathbb{N} \ni n \overset{f}{\longmapsto} \begin{cases} n + \lfloor \frac{n}{2} \rfloor & \text{if } n \text{ odd} \\ n - \lfloor \frac{n}{4} \rfloor & \text{if } n \text{ even} \end{cases} \tag{8.23}$$

For the first few natural numbers, this looks like



(8.24)

Then, the partial sums associated to the sequence $a \circ f$ are given by

$$\mathbb{N} \ni m \mapsto \sum_{n=1}^{m} a_{f(n)} = \sum_{n \text{ odd}}^{m} a_{n + \lfloor \frac{n}{2} \rfloor} + \sum_{n \text{ even}}^{m} a_{n - \lfloor \frac{n}{4} \rfloor} \tag{8.25}$$

The first few terms of this rearranged series are

$$\begin{aligned} \underbrace{1 - \frac{1}{2}}_{\frac{1}{2}} - \frac{1}{4} + \underbrace{\frac{1}{3} - \frac{1}{6}}_{\frac{1}{6}} - \frac{1}{8} + \underbrace{\frac{1}{5} - \frac{1}{10}}_{\frac{1}{10}} - \frac{1}{12} + \frac{1}{7} + \cdots &= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \cdots \\ &= \frac{1}{2} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \end{aligned} \tag{8.26}$$

This shows that upon a particular rearrangement, the value of the series was halved!

Theorem 8.27. *If the series associated to a sequence converges absolutely, then any rearrangement of the sequence gives an associated series that converges to the same limit.*

In other words, the "sum" in an absolutely convergent series is commutative.

Proof. See Abbott. ■

Just as there was an Algebraic Limit Theorem for Sequences, there is one for series. However, there are subtleties for the products of two series. Let $a, b : \mathbb{N} \to \mathbb{R}$ be two sequences with associated partial sums S and T, respectively. Then the product of the partial sums S and T is given by the product of the associated partial sums, namely

$$\mathbb{N} \ni n \mapsto (ST)_n \equiv S_n T_n := \left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{n} b_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j. \tag{8.28}$$

Theorem 8.29 (Algebraic Limit Theorem for Series)**.** *Let a and b be two sequences with associated convergent series $A := \sum_{n=1}^{\infty} a_n$ and $B := \sum_{n=1}^{\infty} b_n$. Then the following hold.*

(a) $\sum_{n=1}^{\infty} c a_n = cA$ *for all* $c \in \mathbb{R}$.

(b) $\sum_{n=1}^{\infty}(a_n + b_n) = A + B$.

(c) If either of the series $\sum_{n=1}^{\infty} a_n$ *or* $\sum_{n=1}^{\infty} b_n$ *converge absolutely, then*

$$\lim_{n\to\infty}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j\right) = AB = \lim_{n\to\infty}\left(\sum_{k=1}^{n} a_k b_{n-k}\right). \tag{8.30}$$

Proof. The first two are immediate consequences of the algebraic limit theorem for sequences. The last one is non-trivial. A lengthy proof of the first equality is given in Abbott while a shorter proof of the second equality is given by [6]. ■

After this lecture, it is recommended the student works through problems 8 and 9 on HW #2. Additional recommended exercises include exercises 2.7.1 (I highly recommend this one!), 2.7.2, 2.7.4, and 2.7.8 in [2] (I'll upload the corresponding exercise in [1] shortly).

9 September 27

Today we will construct $\mathbb{R}$ from $\mathbb{Q}$ in groups. The following are my notes on this construction. I will also make available the work of the students once ready.

10 September 29

The next few lectures will focus on topological concepts from $\mathbb{R}$. There are many (equivalent) ways to present this material. The definitions of what it means to be open, closed, and so on may differ from reference to reference. Nevertheless, most definitions are all equivalent. My presentation prefers to choose the definition to be something satisfying a particular condition and then proving that there is a construction for those things. This only slightly differs from the presentation in Abbott's book.

Definition 10.1. Let $a \in \mathbb{R}$ and $\epsilon > 0$. The *ϵ-neighborhood* of a is the set[33]

$$V_\epsilon(a) := \{x \in \mathbb{R} \ : \ |x - a| < \epsilon\} \equiv (a - \epsilon, a + \epsilon). \tag{10.2}$$

Definition 10.3. A subset $A \subseteq \mathbb{R}$ is *open* if for every $a \in A$, there exists an $\epsilon > 0$ such that $V_\epsilon(a) \subseteq A$.

Example 10.4. $\mathbb{Q}$ is not an open subset of $\mathbb{R}$. To see this, let $r \in \mathbb{Q}$ and fix $\epsilon > 0$. Then $V_\epsilon(r) = (r - \epsilon, r + \epsilon)$ and by the density of $\mathbb{I}$ in $\mathbb{R}$ (Theorem 4.2), there exists an irrational number $s \in (r, \epsilon)$. Therefore, $V_\epsilon(r) \nsubseteq \mathbb{Q}$. Since this happens for every $\epsilon > 0$, $\mathbb{Q}$ is not open.

Theorem 10.5.

(a) $\mathbb{R}$ and $\varnothing$ are open subsets of $\mathbb{R}$.

(b) The union of an arbitrary collection of open sets is open.

(c) The intersection of a finite collection of open sets is open.

Proof.

(a) Let $a \in \mathbb{R}$ and $\epsilon > 0$. Then $(a - \epsilon, a + \epsilon) \subseteq \mathbb{R}$ showing that $\mathbb{R}$ is open. $\varnothing$ is open vacuously since there is no element in $\varnothing$ so that the required condition automatically holds.

(b) Let Λ be a set and $\{U_\lambda\}_{\lambda \in \Lambda}$ a collection of open sets in $\mathbb{R}$ indexed by Λ. Let $a \in \bigcup_{\lambda \in \Lambda} U_\lambda$. By definition of the union, $a \in U_\lambda$ for some $\lambda \in \Lambda$. Since U_λ is open, there exists an $\epsilon > 0$ such that $V_\epsilon(a) \subseteq U_\lambda$. Then $V_\epsilon(a) \subseteq \bigcup_{\lambda \in \Lambda} U_\lambda$.

(c) Let $\{U_1, \ldots, U_n\}$ be a finite collection of open sets in $\mathbb{R}$. Let $a \in \bigcap_{i=1}^{\infty} U_i$. Then, by definition of intersection, $a \in U_i$ for all $i \in \{1, \ldots, n\}$. Hence, since each U_i is open, there exists an $\epsilon_i > 0$ such that $V_{\epsilon_i}(a) \subseteq U_i$. Set $\epsilon := \min\{\epsilon_1, \ldots, \epsilon_n\}$. Then $V_\epsilon(a) \subseteq \bigcap_{i=1}^{n} U_i$ because $V_\epsilon(a) \subseteq U_i$ for all $i \in \{1, \ldots, n\}$.

■

[33] The definition of an ϵ-neighborhood seems a little useless at this point because it turns out the just be an open interval. These definitions become a little more useful in higher dimensional Euclidean spaces, abstract manifolds, and even more general spaces, where the notion of an interval is no longer available, but an ϵ-neighborhood still makes sense.

This theorem is the foundations for the study of topology, the subject that studies spaces deformable under continuous transformations. Because it is so important, and closely related to this theorem, we provide the definition.

Definition 10.6. Let X be a set. A *topology* on X is a subset $\tau \subseteq \mathcal{P}(X)$ of the power set of X satisfying the following conditions.

(a) $X, \varnothing \in \tau$.

(b) For any collection of sets $U : \Lambda \to \tau$ with $U_\lambda := U(\lambda)$, $\bigcup_{\lambda \in \Lambda} U_\lambda \in \tau$.

(c) For any finite collection of sets $\{U_1, \ldots, U_n\}$ with $U_i \in \tau$ for all $i = 1, \ldots, n$, $\bigcap_{i=1}^{n} U_i \in \tau$.

It took many years to come up with the above definition as a robust enough approach for a study of continuity.

Definition 10.7. Let $A \subseteq \mathbb{R}$ be a set and $\tau_{\mathbb{R}}$ the set of open sets in $\mathbb{R}$. An *open cover* of A is a collection of open sets $U : \Lambda \to \tau_{\mathbb{R}}$ such that

$$A \subseteq \bigcup_{\lambda \in \Lambda} U_\lambda. \tag{10.8}$$

Theorem 10.5 and Definition 10.7 have interesting consequences. Even though we know that $\mathbb{Q}$ is not an open set, we can still cover it by open sets. A naive guess would think that every open cover of $\mathbb{Q}$ would contain every irrational number since the rational numbers are spread throughout all of $\mathbb{R}$ in such a way that between any two irrationals is a rational. Nevertheless, this turns out to be false.

Example 10.9. Let $\varphi : \mathbb{N} \to \mathbb{Q}$ be an enumeration of the rationals. Fix $p > 1$ and consider the sequence of open intervals given by

$$\mathbb{N} \ni n \mapsto J_n := \left(\varphi(n) - \frac{1}{2p^n}, \varphi(n) + \frac{1}{2p^n} \right). \tag{10.10}$$

Then the set of all $\{J_n\}_{n \in \mathbb{N}}$ form an open cover of $\mathbb{Q}$,

$$\mathbb{Q} \subseteq \bigcup_{n=1}^{\infty} J_n. \tag{10.11}$$

In particular, $\bigcup_{n=1}^{\infty} J_n$ is an open set by Theorem 10.5 and it contains $\mathbb{Q}$. Even more surprisingly, the length (a.k.a. *measure*) of $\bigcup_{n=1}^{\infty} J_n$ is finite and bounded from above by[34]

$$\sum_{n=1}^{\infty} \frac{1}{p^n} = \frac{1}{p-1}. \tag{10.12}$$

Although closed sets in $\mathbb{R}$ are usually thought of as closed intervals, the latter are just a special case of the former. Closed sets are sets which contain limits of sequences, whenever they exist, that are contained in them.

[34] It is bounded from above by this value because any two of these open sets might intersect, and this sum overcounts the length due to the overlaps.

Definition 10.13. Let $A \subseteq \mathbb{R}$. A point $x \in \mathbb{R}$ is a *limit point* of A if for any $\epsilon > 0$,

$$V_\epsilon(x) \cap (A \setminus \{x\}) \neq \varnothing. \tag{10.14}$$

A point $a \in A$ is an *isolated point* of A if it is not a limit point of A.

Example 10.15. Let $a : \mathbb{N} \to \mathbb{R}$ be a convergent sequence with limit $\lim a \notin A$, where $A := a(\mathbb{N})$ is the image of the sequence a as a subset of $\mathbb{R}$. Then $\lim a$ is the only limit point of A, i.e. $\lim a$ is a limit point of A and a_n is an isolated point of A for all $n \in \mathbb{N}$.

To see the first claim, note that by definition of $\lim a$, for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$|a_n - \lim a| < \epsilon \qquad \forall\, n \geq N. \tag{10.16}$$

In other words, $a_n \in V_\epsilon(\lim a)$ for all $n \geq N$. Since $a_n \neq \lim a$ for all $n \in \mathbb{N}$, this shows that $\lim a$ is a limit point of $a(\mathbb{N})$.

To see that a_n is an isolated point[35] of $a(\mathbb{N})$ for any $n \in \mathbb{N}$, let

$$\epsilon_n := \frac{|a_n - \lim a|}{2}. \tag{10.17}$$

Note that $\epsilon_n > 0$ since $a_n \neq \lim a$ by assumption. Since a converges to $\lim a$, there exists an $M \in \mathbb{N}$ such that

$$|a_m - \lim a| < \epsilon_n \qquad \forall\, m \geq M. \tag{10.18}$$

Let

$$K := \big\{ k \in \{1, \dots, M-1\} \mid a_k \neq a_n \big\}. \tag{10.19}$$

Since $a_k \neq a_n$ for all $k \in K$, set

$$\epsilon_k := |a_k - a_n|. \tag{10.20}$$

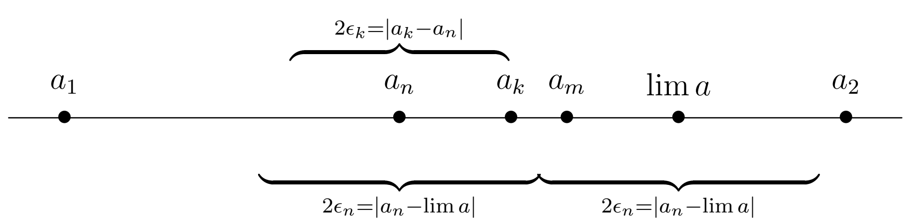
Finally, set

$$\epsilon := \min\Big(\{\epsilon_k \mid k \in K\} \cup \{\epsilon_n\} \Big). \tag{10.21}$$

Then $V_\epsilon(a_n)$ contains a_n but

$$V_\epsilon(a_n) \cap \big(A \setminus \{a_n\}\big) = \varnothing. \tag{10.22}$$

A picture for this proof is given by the following



Definition 10.23. A set $A \subseteq \mathbb{R}$ is *closed* if it contains its limit points.

[35] Kevin Pratt pointed out another (simpler) proof of this fact. Suppose, to the contrary, that a_n is a limit point of a. Then, for all $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $a_N \in V_\epsilon(a_n) \setminus \{a_n\}$, i.e. $V_\epsilon(a_n) \cap a(\mathbb{N}) \setminus \{a_n\} \neq \varnothing$. Setting $\epsilon_m := \frac{1}{m}$ for each $m \in \mathbb{N}$, this shows there exists a subsequence $a \circ f$ of a, with $f : \mathbb{N} \to \mathbb{N}$ one-to-one and nondecreasing, such that $\lim(a \circ f) = a_n$. But since every subsequence of a must converge to $\lim a$ since a is convergent by Theorem 7.10, this contradicts that $a_n \neq \lim a$.

Theorem 10.24. *Let $A \subseteq \mathbb{R}$. A point $x \in \mathbb{R}$ is a limit point of A if and only if there exists a sequence $a : \mathbb{N} \to A \setminus \{x\}$ with $\lim a = x$.*

Proof.
($\Rightarrow$) Let x be a limit point of A and let $\epsilon : \mathbb{N} \to \mathbb{R}$ be the sequence

$$\mathbb{N} \ni n \mapsto \epsilon_n := \frac{1}{n}. \tag{10.25}$$

Since x is a limit point of A, there exists an $a_n \in V_{\epsilon_n}(x)$, with $a_n \neq x$, for all $n \in \mathbb{N}$. Then the associated sequence $a : \mathbb{N} \to \mathbb{R}$ has $\lim a = x$ because for every $\epsilon_x > 0$, there exists an $N_x > \frac{1}{\epsilon_x}$ by the Archimedean Property of $\mathbb{R}$ so that

$$|a_n - x| < \frac{1}{n} \leq \frac{1}{N_x} < \epsilon_x \qquad \forall\, n \geq N_x. \tag{10.26}$$

($\Leftarrow$) This direction was essentially proved in Example 10.15. More precisely, let a be such a sequence. Then that example showed that $\lim a$ is a limit point of $a(\mathbb{N})$, but since $a(\mathbb{N}) \subseteq A$, $\lim a$ is also a limit point of A. ■

This result has many useful applications.

Example 10.27. Every point of $\mathbb{Q}$ is a limit point of $\mathbb{Q}$, but $\mathbb{Q}$ is not closed, i.e. $\mathbb{Q}$ has limit points that are not contained in $\mathbb{Q}$. The first statement follows from the Density of $\mathbb{Q}$ in $\mathbb{R}$. The second statement follows from the previous lecture and the fact that every real number is expressed as a Cauchy sequence of rational numbers. In other words, every irrational number is also a limit point of $\mathbb{Q}$.

Theorem 10.28. *A set $A \subseteq \mathbb{R}$ is closed if and only if for every Cauchy sequence $a : \mathbb{N} \to A$, $\lim a \in A$.*

Proof. This is one of your homework problems. ■

Theorem 10.29. *A set $A \subseteq \mathbb{R}$ is open if and only if $A^c \equiv \mathbb{R} \setminus A$ is closed.*

Proof.
($\Rightarrow$) Let A be open and let x be a limit point of A^c. Suppose to the contrary that $x \in A$. Then, since A is open, there exists an $\epsilon > 0$ such that $V_\epsilon(x) \subseteq A$. In particular $V_\epsilon(x) \cap A^c = \varnothing$. This contradicts that x is a limit point of A^c. Therefore, $x \in A^c$.
($\Leftarrow$) Let A^c be closed and let $y \in A$. Since A^c is closed, y is *not* a limit point of A^c. Therefore, there exists an $\epsilon > 0$ such that $V_\epsilon(y) \cap A^c = \varnothing$. By definition of complement and intersection, this means $V_\epsilon(y) \subseteq A$ showing that A is open. ■

Corollary 10.30. *A set is closed if and only if its complement is open.*

Proof. Since taking the complement is an involution (an operation that squares to the identity), this follows from Theorem 10.29. ■

Theorem 10.31.

(a) $\varnothing$ *and* $\mathbb{R}$ *are closed subsets of* $\mathbb{R}$.

(b) *The union of a finite collection of closed sets is closed.*

(c) *The intersection of an arbitrary collection of closed sets is closed.*

Proof. This follows from Theorem 10.29 and DeMorgan's laws. See exercise 3.2.9 in [2]. ■

Definition 10.32. Let $A \subseteq \mathbb{R}$. The *closure* of A, is a closed subset, $\overline{A}$, of $\mathbb{R}$ satisfying

(a) $A \subseteq \overline{A}$ and

(b) for any other closed subset $B \subseteq \mathbb{R}$ with $A \subseteq B$, $\overline{A} \subseteq B$.

The definition of closure of a set should remind you of the definition of the supremum of a bounded set. An explicit construction of $\overline{A}$ is given by the following fact.

Theorem 10.33. *Let* $A \subseteq \mathbb{R}$. *Then*

$$\overline{A} = A \cup L_A, \tag{10.34}$$

where L_A *is the set of limit points of* A. *In addition,*

$$\overline{A} = \bigcap_{B} B \tag{10.35}$$

where the intersection is over all closed sets containing A. *In particular,* A *is closed if and only if* $A = \overline{A}$.

Proof. Exercise. ■

Exercise 10.36. Prove that $\overline{\mathbb{Q}} = \mathbb{R}$.

After this lecture, it is recommended the student works through problems 2, 3, and 4 on HW #3. Additional recommended exercises include exercises 3.2.1, 3.2.2, 3.2.7, 3.2.8, 3.2.9, 3.2.11, and 3.2.14 in [2] (I'll upload the corresponding exercise in [1] shortly).

11 October 4

Definition 11.1. A subset $K \subseteq \mathbb{R}$ is <u>*compact*</u> if every sequence in K has a subsequence converging to a point in K.

Example 11.2. Let A be a finite subset of $\mathbb{R}$ and let $a : \mathbb{N} \to A$ be a sequence in A. Because A is finite, at least one of the elements, say x, of A will appear infinitely many times in the sequence a. Thus, the subsequence of a given by $\mathbb{N} \ni n \mapsto x$ converges to x, which is in A.

Example 11.3. Every closed interval $[a, b]$, with $a \leq b$, is compact.

Theorem 11.4. *Let*

$$K_1 \supseteq K_2 \supseteq K_3 \supseteq K_4 \supseteq \cdots \tag{11.5}$$

be a nested sequence of nonempty compact subsets of $\mathbb{R}$. *Then*

$$\bigcap_{n=1}^{\infty} K_n \neq \varnothing. \tag{11.6}$$

This theorem should be compared to the nested interval property. If "compact" is replaced by "closed" above, the statement is false. A counter-example is given by the collection of closed unbounded intervals $[n, \infty)$ for each $n \in \mathbb{N}$.

Proof. Because each K_n is non-empty, let $a_n \in K_n$. Let $a : \mathbb{N} \to K_1$ be the sequence defined by $\mathbb{N} \ni n \mapsto a_n$. Since K_1 is compact, there exists a one-to-one non-decreasing function $f : \mathbb{N} \to \mathbb{N}$ so that $a \circ f : \mathbb{N} \to K_1$ is a convergent subsequence of a with limit $\lim(a \circ f) \in K_1$. Because $f : \mathbb{N} \to \mathbb{N}$ is non-decreasing and one-to-one, $a_{f(n)} \in K_{f(n)} \subseteq K_n$ for all $n \in \mathbb{N}$. Therefore, for every $N \in \mathbb{N}$, let $\varphi_N : \mathbb{N} \to \mathbb{N}$ be the function defined by $\varphi_N(n) := N + n$. Then $a \circ f \circ \varphi_N$ is a subsequence of $a \circ f$ contained in K_N. Since $a \circ f$ converges, $a \circ f \circ \varphi_N$ converges to the same limit by Theorem 7.10. Since K_N is compact, $\lim(a \circ f \circ \varphi_N) \in K_N$. This shows that $\lim(a \circ f) \in K_N$ for all $N \in \mathbb{N}$ showing that $\lim(a \circ f) \in \bigcap_{i=1}^{\infty} K_i$. ■

A "picture" for this proof is

$$\begin{array}{ccccccccccccc}
a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & \cdots \\
\rotatebox{-90}{$\in$} & \rotatebox{-90}{$\in$} & \rotatebox{-90}{$\in$} & \rotatebox{-90}{$\in$} & \rotatebox{-90}{$\in$} & \rotatebox{-90}{$\in$} & \rotatebox{-90}{$\in$} & \rotatebox{-90}{$\in$} & \rotatebox{-90}{$\in$} & \rotatebox{-90}{$\in$} & \rotatebox{-90}{$\in$} & \rotatebox{-90}{$\in$} & \\
K_1 \supseteq & K_2 \supseteq & K_3 \supseteq & K_4 \supseteq & K_5 \supseteq & K_6 \supseteq & K_7 \supseteq & K_8 \supseteq & K_9 \supseteq & K_{10} \supseteq & K_{11} \supseteq & K_{12} \supseteq & \cdots \\
 & \rotatebox{90}{$\in$} & \rotatebox{90}{$\in$} & & \rotatebox{90}{$\in$} & & \rotatebox{90}{$\in$} & \rotatebox{90}{$\in$} & & & \rotatebox{90}{$\in$} & \rotatebox{90}{$\in$} & \\
 & a_{f(1)} & a_{f(2)} & & a_{f(3)} & & a_{f(4)} & a_{f(5)} & & & a_{f(6)} & a_{f(7)} &
\end{array} \tag{11.7}$$

The reason for the shift φ_N is so that the sequence is completely contained in K_N.

Lemma 11.8. *Compact subsets of* $\mathbb{R}$ *are closed.*

Proof. Let K be compact and let $x : \mathbb{N} \to K$ be a convergent sequence in K. By Theorem 7.10, every subsequence of x is convergent and converges to the same limit. Because K is compact, this limit is in K, i.e. $\lim x \in K$. Therefore, K is closed. ■

Lemma 11.9. *The intersection of an arbitrary collection of compact sets is compact.*

Proof. Let Λ be a set, $\{K_\lambda\}_{\lambda\in\Lambda}$ such a collection, and

$$a : \mathbb{N} \to \bigcap_{\lambda\in\Lambda} K_\lambda \tag{11.10}$$

a sequence in the intersection. Fix $\lambda_0 \in \Lambda$. Since K_{λ_0} is compact, there exist convergent subsequence $a \circ f : \mathbb{N} \to K_{\lambda_0}$ ($f : \mathbb{N} \to \mathbb{N}$ is one-to-one and non-decreasing) with $\lim(a \circ f) \in K_{\lambda_0}$. Because a lands in the intersection of all these compact sets, this subsequence lands in the intersection as well, namely

$$(a \circ f)(\mathbb{N}) \subseteq \bigcap_{\lambda\in\Lambda} K_\lambda. \tag{11.11}$$

Since every compact set is closed by Lemma 11.8 and since an arbitrary intersection of closed sets is closed by Theorem 10.31,

$$\lim(a \circ f) \in \bigcap_{\lambda\in\Lambda} K_\lambda. \tag{11.12}$$

Thus, $a \circ f$ is a convergent subsequence of $a : \mathbb{N} \to \bigcap_{\lambda\in\Lambda} K_\lambda$ whose limit is in $\bigcap_{\lambda\in\Lambda} K_\lambda$. ■

Definition 11.13. Let $A \subseteq \mathbb{R}$, $\tau_{\mathbb{R}}$ the set of open sets in $\mathbb{R}$, and $U : \Lambda \to \tau_{\mathbb{R}}$ an open cover of A. A *subcover* of $U : \Lambda \to \tau_{\mathbb{R}}$ is an open cover of A of the form $U \circ i : \Omega \to \tau_{\mathbb{R}}$, where $\Omega \subseteq \Lambda$ and $i : \Omega \hookrightarrow \Lambda$ is the inclusion. If Ω is a finite, then the subcover is said to be a *finite subcover*.

Theorem 11.14 (Heine-Borel Theorem)**.** *Let $K \subset \mathbb{R}$. The following statements are equivalent.*

(a) Every open cover of K has a finite subcover.

(b) K is closed and bounded.

(c) K is compact.

Proof. In this proof, we will show the following sequence of implications (a)⇒(b)⇒(c)⇒(a). In the process of proving (c)⇒(a), we actually prove (c)⇒(b) so we include that separately as well.
(a)⇒(b) Let $\epsilon := 1$. The collection

$$\mathcal{U} := \{V_\epsilon(x) \mid x \in K\} \tag{11.15}$$

of ϵ-neighborhoods around all points in K is an open cover of K. By assumption, there exists a finite subset $\{x_1, \ldots, x_n\} \subset K$ such that

$$K \subseteq \bigcup_{i=1}^{n} V_\epsilon(x_i). \tag{11.16}$$

Thus, K is bounded, for instance by

$$\max_i\{|x_i + \epsilon|, |x_i - \epsilon|\}. \tag{11.17}$$

Instead of proving that K is closed directly, we prove the contrapositive. Namely, we prove that K is not closed implies there exists an open cover of K that does not have a finite subcover. Let

x be a limit point of K that is not contained in K. Set $\epsilon_0 := 1$. Since x is a limit point of K, there exists an $x_1 \in K \cap V_{\epsilon_0}(x)$. Recursively, set

$$\epsilon_n := \frac{|x - x_n|}{2}, \tag{11.18}$$

where $x_n \in K \cap V_{\epsilon_{n-1}}(x)$. x_n always exists because $\epsilon_{n-1} > 0$ for all $n \in \mathbb{N}$ and x is a limit point of K. Then, for each $n \in \mathbb{N}$, define

$$U_n := \overline{V_{\epsilon_{n-1}}(x)}^c \equiv (-\infty, x - \epsilon_{n-1}) \cup (x + \epsilon_{n-1}, \infty). \tag{11.19}$$

This is an open set for each $n \in \mathbb{N}$ by Theorem 10.29. Furthermore, by construction, $x_n \in K \cap U_n$ for all $n \in \mathbb{N}$ and each x_n is distinct. In other words, $\{U_n\}_{n\in\mathbb{N}}$ is a countably infinite open cover of K that does not have a finite subcover containing K.
(b)⇒(c) Suppose K is closed and bounded and let $a : \mathbb{N} \to K$ be a sequence in K. Because K is bounded, the Bolzano-Weierstrass Theorem (Theorem 7.11) implies there exists a convergent subsequence $a \circ f : \mathbb{N} \to K$ with $f : \mathbb{N} \to \mathbb{N}$ one-to-one and non-decreasing. Because K is closed, Theorem 10.28 guarantees that $\lim(a \circ f) \in K$.
(c)⇒(b) Let K be compact. K is closed by Lemma 11.8. To see that K must be bounded as well, we will prove the contrapositive, namely K not bounded implies there exists a sequence $a : \mathbb{N} \to K$ with no convergent subsequence. For each $n \in \mathbb{N}$, define

$$K_0 := K \qquad \& \qquad K_n := K \cap (-n, n)^c. \tag{11.20}$$

By assumption, K_n is nonempty for all $n \in \mathbb{N}$. Thus, for each $n \in \mathbb{N}$, let $a_n \in K_n$. Then $a : \mathbb{N} \to \mathbb{R}$, defined by $\mathbb{N} \ni n \mapsto a_n$, is a divergent sequence in K containing no convergent subsequence.
(c)⇒(a) Let K be compact. By the previous paragraph, K is bounded and closed. Let $M \in \mathbb{R}$ be such a bound for K, i.e. $K \subseteq [-M, M]$. Now, let $\{U_\lambda\}_{\lambda\in\Lambda}$ be an open cover of K. Suppose, to the contrary, that $\{U_\lambda\}_{\lambda\in\Lambda}$ does not contain a finite subcover of K. Set $I_0 := [-M, M]$. By assumption, it must be that at least one of the two sets $[-M, 0] \cap K$ or $[0, M] \cap K$ cannot be covered by a finite subcover of $\{U_\lambda\}_{\lambda\in\Lambda}$ (if both of them could be, then taking the union of the two finite subcovers would give a finite subcover of K). Let I_1 be one such choice of $[-M, 0]$ or $[0, M]$. Define I_n inductively in this way. Namely, suppose $I_n := [a_n, b_n]$ with $a_n, b_n \in \mathbb{R}$ and $a_n < b_n$ and no finite subcover of $\{U_\lambda\}_{\lambda\in\Lambda}$ covers $I_n \cap K$. At least one of the two intervals

$$\left[a_n, b_n - \frac{1}{2^n}\right] \qquad \text{or} \qquad \left[a_n + \frac{1}{2^n}, b_n\right] \tag{11.21}$$

satisfies the condition that there does not exist a finite subcover of $\{U_\lambda\}_{\lambda\in\Lambda}$ covering its intersection with K. Choose one and set $I_{n+1} \equiv [a_{n+1}, b_{n+1}]$ to be this choice. This gives collection of sets

$$K_n := I_n \cap K \tag{11.22}$$

that are compact by Lemma 11.9. Furthermore, this gives a nested sequence of compact sets

$$K_1 \supseteq K_2 \supseteq K_3 \supseteq K_4 \supseteq \cdots . \tag{11.23}$$

By Theorem 11.4,

$$\bigcap_{n=1}^{\infty} K_n \neq \varnothing. \tag{11.24}$$

Because these intervals are decreasing in length to zero as n increases, $\bigcap_{n=1}^{\infty} K_n$ consists of only a single point; call this point x. Because $x \in K$, there exists a $\lambda_x \in \Lambda$ such that $x \in U_{\lambda_x}$. Because U_{λ_x} is open, there exists an $\epsilon > 0$ such that $V_\epsilon(x) \subseteq U_{\lambda_x}$. Since $\lim_{n\to\infty} a_n = x = \lim_{n\to\infty} b_n$, there exists an $N \in \mathbb{N}$ such that $a_n, b_n \in V_\epsilon(x)$ for all $n \geq N$. In particular, $I_n \subseteq V_\epsilon(x) \subseteq U_{\lambda_x}$ for all $n \geq N$. This contradicts that there is a finite subcover of I_n for $n \geq N$. ■

Remark 11.25. This theorem is of significant importance. You might run across sequences in spaces that do not have a notion of distance. Being bounded and convergence of sequences both use such a notion of distance. However, in a more abstract setting where one might not have a natural notion of distance (or in instances where one can make statements *independent* of such a notion), item (a) can be used as a *definition* of compactness.

Remark 11.26. Part of the proof of the Heine-Borel theorem illustrates the connection between the definition of a subset of $\mathbb{R}$ being compact that we gave and the equivalent concept of that subset being closed and bounded in $\mathbb{R}$. A subset $K \subseteq \mathbb{R}$ is compact if and only if either of the two columns are true (and the horizontal conditions are equivalent for any subset K of $\mathbb{R}$)

$$\begin{array}{ccc} \text{Every sequence in } K \text{ contains a Cauchy subsequence} & \xLeftrightarrow{\text{Bolzano-Weirestrass}} & K \text{ is bounded} \\ \text{The limit of every Cauchy sequence in } K \text{ is in } K & \xLeftrightarrow[\text{Definition of } K \text{ closed}]{} & K \text{ is closed} \end{array}$$

After this lecture, it is recommended the student works through problem 5 on HW #3. Additional recommended exercises include 3.3.8., 3.3.10, and 3.3.12 from [2] (I'll upload the corresponding exercise in [1] shortly).

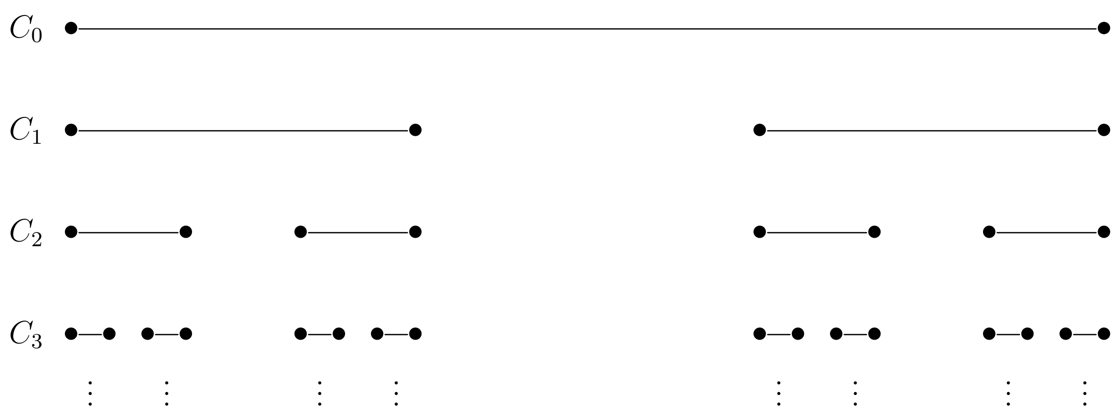
12 October 6

Unfortunately, I have to throw some definitions at you. I'm not entirely sure how much of these are necessary for what we will do in this course. I know connectedness will be crucial for functions and continuity. I also know the dense subsets show up a lot in many areas of mathematics (also see Remark 12.12). I've personally never used the fact that a set was perfect or nowhere dense to prove anything, but who knows what will happen in the future.

Definition 12.1. A subset $P \subseteq \mathbb{R}$ is *perfect* if it is closed and contains no isolated points.

Example 12.2. All closed intervals and unbounded closed intervals are perfect.

Example 12.3 (The Cantor Set)**.** Define C_n to be the subset of $[0, 1]$ depicted in the following figure.

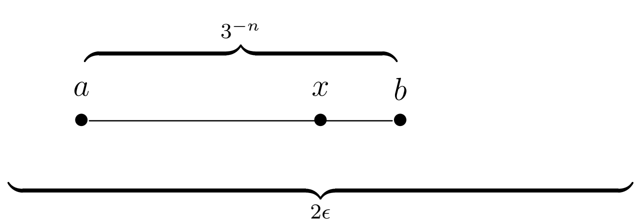


The *Cantor set* is the intersection of all these sets

$$C := \bigcap_{n=1}^{\infty} C_n. \tag{12.4}$$

Proposition 12.5. *The Cantor set is perfect.*

Proof. To see this, first notice that it is closed because it is the intersection of closed sets (see Theorem 10.31). Now let $x \in C$ and fix $\epsilon > 0$. Then, there exists an $n \in \mathbb{N}$ such that $3^{-n} < \epsilon$. Because $x \in C$, $x \in C_n$. In particular, x is contained in some interval $[a, b] \subseteq C_n$.



Because n was chosen to satisfy $3^{-n} < \epsilon$, it follows that $a, b \in V_\epsilon(x)$. Since the endpoints of any interval in C_n is in C, this proves $V_\epsilon(x) \cap C \setminus \{x\} \neq \varnothing$ showing that $x \in C$ is not isolated. ■

Corollary 12.6. *The Cantor set is compact.*

Proof. This follows immediately from the fact that perfect implies closed and since $C \subseteq [0,1]$, it is bounded and therefore compact by the Heine-Borel Theorem (Theorem 11.14). ■

Exercise 12.7. Prove that the Cantor set is compact explicitly (straight from the definitions) without referring to the Heine-Borel Theorem or Lemma 11.9.

Theorem 12.8. *A nonempty perfect set is uncountable.*

Proof. See Abbott. ■

Definition 12.9. Two nonempty subsets $A, B \subseteq \mathbb{R}$ are said to be *separated* if $\overline{A} \cap B$ and $A \cap \overline{B}$ are both empty. A subset $E \subseteq \mathbb{R}$ is *disconnected* if there exist separated subsets $A, B \subseteq E$ with $E = A \cup B$. A set E is *connected* if it is not disconnected.

Theorem 12.10. *A subset $E \subseteq \mathbb{R}$ is connected if and only if for any $a, b \in E$ with $a < b$ implies $(a, b) \subseteq E$.*

Proof. See Abbott. ■

We may take this theorem as the definition of connectedness in this course. However, please be aware that the more general definition is more relevant in higher dimensions.

Definition 12.11. A subset $G \subseteq \mathbb{R}$ is *dense* if for any open subset $U \subseteq \mathbb{R}$, $G \cap U \neq \varnothing$.

Remark 12.12. Density is a very useful concept. We will use this time and time again when we begin discussing functions and convergence of functions. Often, we will be able to prove theorems about dense subsets. We will then use density to extend the result to larger classes of functions.

Example 12.13. $\mathbb{Q}$ is a dense subset of $\mathbb{R}$.

Theorem 12.14. *A subset $G \subseteq \mathbb{R}$ is dense if and only if $\overline{G} = \mathbb{R}$.*

Proof.
($\Rightarrow$) Let G be dense and let $x \in \mathbb{R}$. Since G is dense, for every $\epsilon > 0$, $V_\epsilon(x) \cap G \setminus \{x\} \neq \varnothing$. By definition and since $V_\epsilon(x)$ is open, this means x is a limit point of G. Thus, $x \in \overline{G}$.
($\Leftarrow$) The proof will be by contrapositive. Suppose G is not dense. Then there exists an open set $U \subseteq \mathbb{R}$ such that $G \cap U = \varnothing$. Let $x \in U$. Since U is open, there exists an $\epsilon > 0$ such that $V_\epsilon(x) \subseteq U$. Then U^c is a closed set (by Theorem 10.29) containing G. Hence $\overline{G} \subseteq U^c$. Since $x \notin U^c$, $x \notin \overline{G}$ and so $\overline{G} \neq \mathbb{R}$. ■

To understand all of these different topological definitions, it helps to keep a table of several examples. The following table provides some. In this table, C is the Cantor set, $a < b$ (strict inequality), and $\varnothing$ is the empty set. It a property holds, T is written. If a property does not hold, F is written.

	open	closed	bounded	perfect	compact	dense	connected
$\mathbb{R}$	T	T	F	T	F	T	T
$\mathbb{I}$	F	F	F	F	F	T	F
$\mathbb{Q}$	F	F	F	F	F	T	F
$\mathbb{Z}$	F	T	F	F	F	F	F
(a, b)	T	F	T	F	F	F	T
$[a, b]$	F	T	T	T	T	F	T
$[a, b)$	F	F	T	F	F	F	T
$[a, \infty)$	F	T	F	T	F	F	T
C	F	T	T	T	T	F	F
$\varnothing$	T	T	T	T	T	F	T
$\{0\}$	F	T	T	F	T	F	T
$\{\frac{1}{n} \mid n \in \mathbb{N}\}$	F	F	T	F	F	F	F
$\{0\} \cup \{\frac{1}{n} \mid n \in \mathbb{N}\}$	F	T	T	F	T	F	F

Exercise 12.15. Verify the table. Try to prove each and every claim.

Definition 12.16. A subset $E \subseteq \mathbb{R}$ is <u>*nowhere dense*</u> if $\overline{E}$ contains no nonempty open sets.

Theorem 12.17 (Baire's Theorem)**.** $\mathbb{R}$ *is not a countable union of nowhere-dense sets.*

Abbott has a nice discussion at the end of chapter 3 regarding Baire's theorem and its significance. Please read it to somewhat appreciate this result a bit more.

After this lecture, it is recommended the student works through problems 6, 7, and 8 on HW #3. Additional recommended exercises include 3.4.2, 3.4.5, 3.4.6, 3.4.9, 3.5.8, 3.5.9, and 3.5.10 from [2] (I'll upload the corresponding exercise in [1] shortly).

13 October 11

Today we finish up what I didn't finish up from the rest of chapter 3.

14 October 13

Today we will review chapters 1 through 3 by going through a practice exam and any other related problems.

15 October 18

Today is the exam. Good luck!

16 October 20

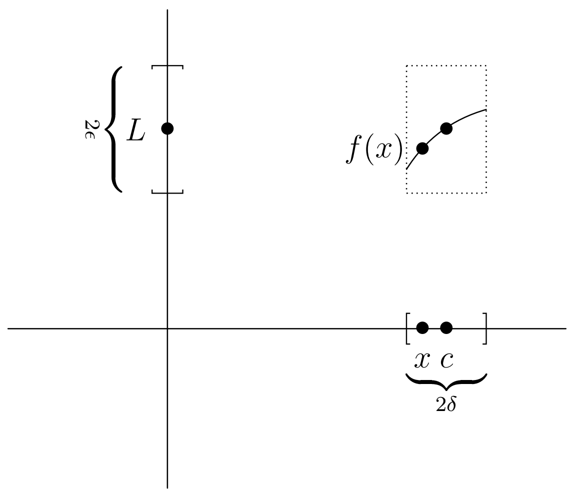
Finally, functions! Please read Abbott for motivation. One of the points about Abbott's discussion is that what we might think of as being a generic function, which is often smooth, is far from generic in the technical math sense. In fact, in the set of all functions of a real variable, *most* of the functions aren't smooth—in fact, most of them aren't even continuous! This is a result of the definition of function as it was carved over many years.

Definition 16.1. Let $A \subseteq \mathbb{R}$ and let $f : A \to \mathbb{R}$. Let c be a limit point of A. L is said to be a *limit of f as x approaches c*, written

$$\lim_{x \to c} f(x) = L, \tag{16.2}$$

if for any $\epsilon > 0$, there exists a $\delta > 0$ such that $f(x) \in V_\epsilon(L)$ for all $x \in V_\delta(c) \cap A \setminus \{c\}$. L is called a *functional limit* of f as x approaches c.

Please note that if c is actually an element of A, a limit of f as x approaches c is completely independent of the value of f at c. Also note that a limit itself need not exist (the proof of the next theorem will illustrate what this negation means explicitly). Another way of phrasing this definition is: "L is a limit of f as x approaches c if for any $\epsilon > 0$, there exists a $\delta > 0$ such that $x \in V_\delta(c) \cap A \setminus \{c\}$ implies $f(x) \in V_\epsilon(L)$." A picture for this definition is



Lemma 16.3. *Let $A \subseteq \mathbb{R}$, let $f : A \to \mathbb{R}$, and let c be a limit point of A. If L and L' are both limits of f as x approaches c, then $L = L'$, i.e. functional limits are unique.*

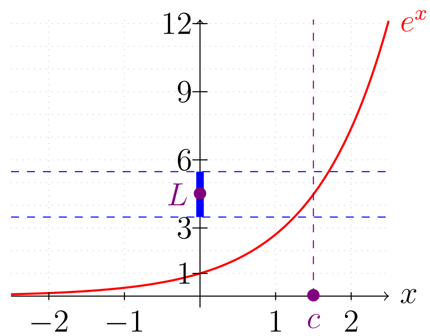
Proof. Suppose that L and L' are both functional limits of f as x approaches c and suppose to the contrary that $L \neq L'$. Without loss of generality, suppose $L < L'$. Then, by density of real numbers, there exists an $\epsilon > 0$ such that $L + 2\epsilon < L'$. By assumption, there exists a $\delta > 0$ such that $|L - f(x)| < \epsilon$ and $|L' - f(x)| < \epsilon$ for all $x \in V_\delta(c) \cap A \setminus \{c\}$. However, these two inequalities cannot hold simultaneously, which leads to a contradiction. ∎

The textbook illustrates a few examples of functions that have functional limits at certain limit points in their domain.

Example 16.4. Let $e : \mathbb{R} \to \mathbb{R}$ be the exponential function (sometimes denoted by exp to avoid confusion with the number e) defined by[36]

$$\mathbb{R} \ni x \mapsto e(x) \equiv e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}, \tag{16.5}$$

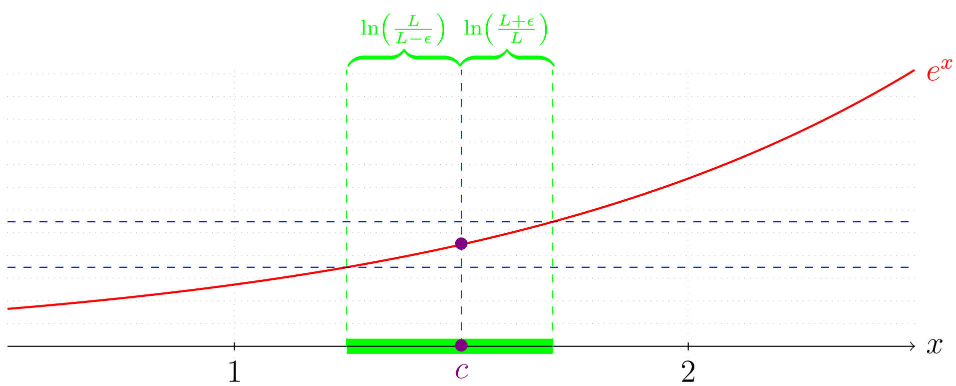
which we (essentially) proved converges several lectures ago. Let $c \in \mathbb{R}$ and let $L := e^c$. Fix $\epsilon > 0$ and consider the ϵ-neighborhood around L (depicted as a vertical blue solid interval in the following figure).



To find an appropriate δ, consider the expression $|e^x - L| < \epsilon$. When $x > c = \ln(L)$, this becomes $e^x - L < \epsilon$ and solving for x gives $\ln(L) < x < \ln(\epsilon + L)$. When $x < c = \ln(L)$, the inequality becomes $L - e^x < \epsilon$, which upon solving for x gives $\ln(L - \epsilon) < x < \ln(L)$. Thus, we should choose

$$\delta := \min\Big\{ \ln(\epsilon + L) - \ln(L), \ln(L) - \ln(L - \epsilon) \Big\} \tag{16.6}$$

motivated by a zoomed in graph of this function



The smaller of these two values is the first (as can also be seen from the figure above). Hence, set $\delta := \ln(\epsilon + L) - \ln(L)$. Then for $x \in (c - \delta, c)$,

$$|e^x - L| = L - e^x < L - e^{c-\delta} = L(1 - e^{-\delta}) = L\left(1 - \frac{L}{\epsilon + L}\right) = L\left(\frac{\epsilon}{\epsilon + L}\right) < L\left(\frac{\epsilon}{L}\right) = \epsilon. \tag{16.7}$$

[36] Actually, we defined e in an earlier lecture. We also defined what it means to raise an arbitrary real number to a real power. Hence, e^x can also be defined in this (apriori different) manner.

Similarly, for $x \in (c, c+\delta)$,

$$|e^x - L| = e^x - L < e^{c+\delta} - L = L(e^\delta - 1) = L\left(\frac{\epsilon + L}{L} - 1\right) = L\left(\frac{\epsilon}{\epsilon + L}\right) < L\left(\frac{\epsilon}{L}\right) = \epsilon. \quad (16.8)$$

Theorem 16.9. *Let $A \subseteq \mathbb{R}$, let $f : A \to \mathbb{R}$, and let c be a limit point of A. Then $\lim\limits_{x \to c} f(x) = L$ if and only if for for every sequence $a : \mathbb{N} \to A \setminus \{c\}$ with $\lim a = c$, it follows that $\lim(f \circ a) = L$.*

Keep in mind that there are two *different* usages of the symbol lim in this theorem. This theorem relates the two.

Proof.
($\Rightarrow$) Suppose $\lim\limits_{x \to c} f(x) = L$ and let $a : \mathbb{N} \to A \setminus \{c\}$ be a sequence with $\lim a = c$. Fix $\epsilon > 0$. Then, by assumption, there exists a $\delta > 0$ such that $|f(x) - L| < \epsilon$ for all $x \in A \setminus \{c\}$ satisfying $|x - c| < \delta$. But since $\lim a = c$, there exists an integer $N \in \mathbb{N}$ such that $|a_n - c| < \delta$ for all $n \geq N$. This implies $|f(a_n) - L| < \epsilon$ for all $n \geq N$ proving that $\lim(f \circ a) = L$.
($\Leftarrow$) The proof will be by contrapositive. In other words, suppose $\lim\limits_{x \to c} f(x) \neq L$. This means that there exists an $\epsilon > 0$ so that for all $\delta > 0$, there exists an $x \in V_\delta(c) \cap A \setminus \{c\}$ with $f(x) \notin V_\epsilon(L)$. In particular, for each $n \in \mathbb{N}$, set $\delta_n := \frac{1}{n}$. Then there exists an $a_n \in V_{\delta_n}(c) \cap A \setminus \{c\}$ with $f(a_n) \notin V_\epsilon(L)$. Since $\lim\limits_{n \to \infty} \delta_n = 0$, the sequence $a : \mathbb{N} \to A \setminus \{c\}$ converges to c. Furthermore, since $f(a_n) \notin V_\epsilon(L)$, the sequence $n \mapsto f(a_n)$ does not converge to L. ■

The following result is useful whenever one wants to show that a functional limit does not exist.

Corollary 16.10. *Let $A \subseteq \mathbb{R}$, let $f : A \to \mathbb{R}$, and let c be a limit point of A. If there exist two distinct sequences $a, b : \mathbb{N} \to A \setminus \{c\}$ such that $\lim a = c$ and $\lim b = c$ but $\lim(f \circ a) \neq \lim(f \circ b)$, then the limit of f as x approaches c does not exist.*

Proof. Exercise. ■

Example 16.11. A simple example illustrating this fact is the function

$$\mathbb{R} \ni x \mapsto h(x) := \begin{cases} -1 & \text{for } x < 0 \\ 0 & \text{for } x = 0 \\ 1 & \text{for } x > 0 \end{cases} \quad (16.12)$$

and the sequences $n \mapsto a_n := \frac{-1}{n}$ and $n \mapsto b_n := \frac{1}{n}$. The limit of both of these sequences is 0 but $f(a_n) = -1$ and $f(b_n) = 1$ for all $n \in \mathbb{N}$. In terms of the functional limit definition, set $\epsilon := 1$, say. Then $|f(x) + 1| < 1$ for all $x \in (-\infty, 0)$ while $|f(x) - 1| < 1$ for all $x \in (0, \infty)$. Therefore, there does not exist an $L \in \mathbb{R}$ and a $\delta > 0$ such that $|f(x) - L| < \delta$ for all $x \in (-\delta, 0) \cup (0, \delta)$.

Example 16.13. Another example is the function

$$\mathbb{R} \ni x \mapsto f(x) := \begin{cases} 1 & \text{for } x \in \mathbb{I} \\ 0 & \text{for } x \in \mathbb{Q} \end{cases} \quad (16.14)$$

known as the <u>*Dirichlet function*</u>. This function has *no* limit points for any real number. To see this let $c \in \mathbb{R}$. If $c \in \mathbb{Q}$, then the sequences defined by

$$a_n := c + \frac{1}{n} \qquad \& \qquad b_n := c + \frac{\sqrt{2}}{n} \tag{16.15}$$

for all $n \in \mathbb{N}$ both converge to c but $a_n \in \mathbb{Q}$ while $b_n \in \mathbb{I}$. As a result, the associated sequences after applying f are $f(a_n) = 0$ and $f(b_n) = 1$ for all $n \in \mathbb{N}$. These are constant sequences converging to two different values. If $c \in \mathbb{I}$ instead, then the sequences defined by

$$\alpha_n := c + \frac{1}{n} \qquad \& \qquad \beta_n := \frac{\lfloor 10^{n-1} c \rfloor}{10^{n-1}} \tag{16.16}$$

for all $n \in \mathbb{N}$ both converge to c but $\alpha_n \in \mathbb{I}$ while $\beta_n \in \mathbb{Q}$. Note that the sequence β_n is the decimal expansion of c, which is why each β_n is rational.

Exercise 16.17. Let $g : \mathbb{R} \to \mathbb{R}$ be the function defined by

$$g(x) := \begin{cases} x & \text{if } x \in \mathbb{I} \\ 0 & \text{if } x \in \mathbb{Q} \end{cases}$$

Find the values of $c \in \mathbb{R}$ for which $\lim\limits_{x \to c} g(x)$ exists and then find the limit when it exists.

After this lecture, it is recommended the student works through problems 1 and 2 on HW #4.

17 October 25

Theorem 16.9 allows us to use several facts about sequences that we have already proved.

Corollary 17.1. *[Algebraic Limit Theorem for Functional Limits] Let $A \subseteq \mathbb{R}$, let $f, g : A \to \mathbb{R}$, and let c be a limit point of A. Furthermore, suppose that $\lim_{x\to c} f(x) = L$ and $\lim_{x\to c} g(x) = M$. Then the following facts follow.*

(a) $\lim_{x\to c} \big[(kf)(x)\big] = kL$ *for all* $k \in \mathbb{R}$.

(b) $\lim_{x\to c} \big[(f+g)(x)\big] = L + M$.

(c) $\lim_{x\to c} \big[(fg)(x)\big] = LM$.

(d) Let $B \subseteq A$ be the domain over which g is nonzero and so that c is also a limit point of B. Then $\lim_{x\to c} \left[\left(\frac{f}{g}\right)(x)\right] = \frac{L}{M}$, *provided that $M \neq 0$ (here $\frac{f}{g}$ is defined on B).*

Proof. Exercise. ■

Definition 17.2. Let $A \subseteq \mathbb{R}$ and let $f : A \to \mathbb{R}$ be a function. f is continuous at $c \in A$ if for every $\epsilon > 0$, there exists a $\delta > 0$ such that $f(x) \in V_\epsilon\big(f(c)\big)$ for all $x \in V_\delta(c) \cap A$. If f is not continuous at $c \in A$, then f is said to be discontinuous at $c \in A$. f is continuous on A if f is continuous at c for all $c \in A$.

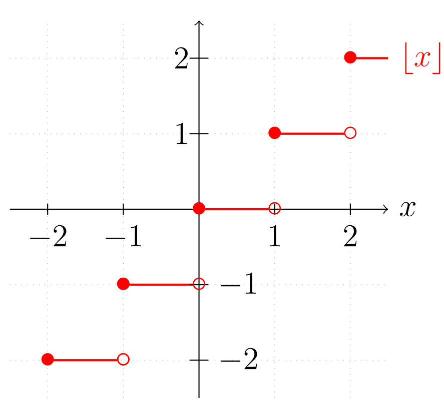
The first definition can be rephrased as: "f is continuous at $c \in A$ if for every $\epsilon > 0$, there exists a $\delta > 0$ such that $x \in V_\delta(c) \cap A$ implies $f(x) \in V_\epsilon\big(f(c)\big)$." Yet another way of rephrasing this even more concisely is: "f is continuous at $c \in A$ if for every $\epsilon > 0$, there exists a $\delta > 0$ such that $f\big(V_\delta(c)\big) \subseteq V_\epsilon\big(f(c)\big)$."

Problem 17.3. Let $\lfloor \cdot \rfloor : \mathbb{R} \to \mathbb{R}$ be the floor function.

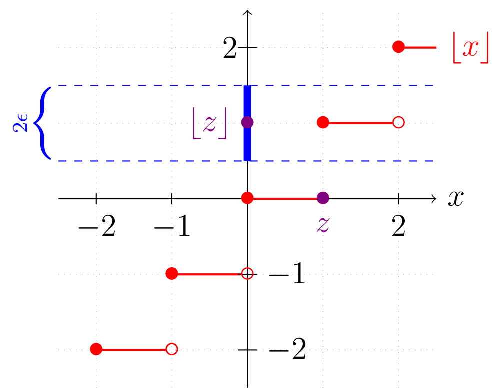
(a) Show that for $z \in \mathbb{Z}$ any integer, the function $\lfloor \cdot \rfloor$ is discontinuous at n.

(b) Show that for $r \in \mathbb{Z}^c$, i.e. for any real number that is *not* an integer, the function $\lfloor \cdot \rfloor$ is continuous at r.

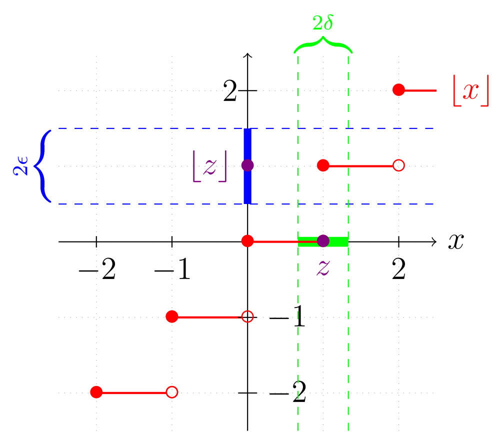
Answer. The graph of the floor function looks like the following



(a) Fix some integer $z \in \mathbb{Z}$. Set $\epsilon := \frac{1}{2}$ and consider the ϵ-neighborhood around $\lfloor z \rfloor = z$ depicted as follows (in the figure, $z = 1$ and is represented by a purple bullet and $\lfloor z \rfloor = z$ is projected onto the vertical axis in purple as well)



The region between the two dashed blue lines intercepts the vertical axis along the blue interval shown. Consider now an arbitrary $\delta > 0$, with $\delta < 1$, and the associated interval $V_\delta(z) = (z - \delta, z + \delta)$ in the domain (in the figure, vertical dashed green lines depict this region after intersecting with the horizontal axis)



Every such δ-neighborhood $(z - \delta, z + \delta)$ around z contains real numbers z_-, z_+ with

$$z - \delta < z_- < z < z_+ < z + \delta. \tag{17.4}$$

Then, by definition of the floor function,

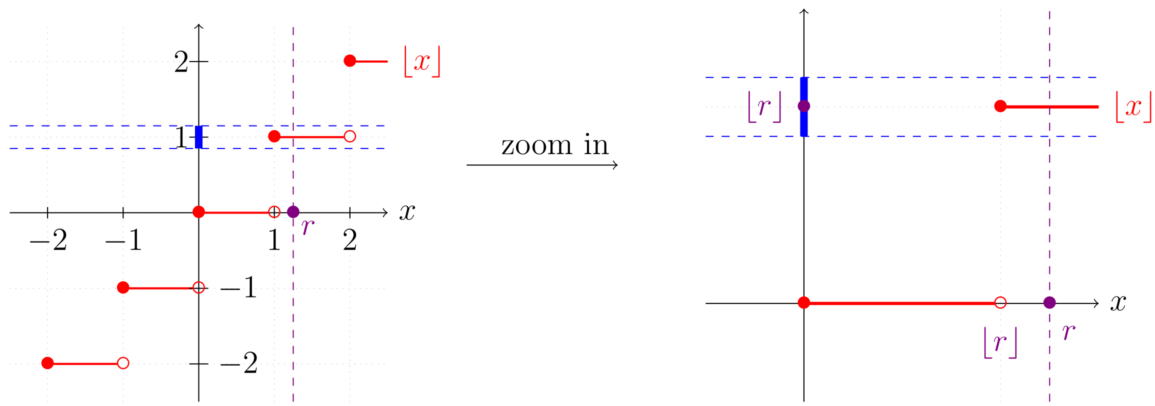
$$\lfloor z_- \rfloor = z - 1 \qquad \& \qquad \lfloor z_+ \rfloor = z. \tag{17.5}$$

Notice that $\lfloor z_- \rfloor$ is *not* in the ϵ-neighborhood around $\lfloor z \rfloor = z$. Hence, $\lfloor \cdot \rfloor$ is discontinuous at $z \in \mathbb{Z}$.

(b) Fix some real number $r \in \mathbb{Z}^c$. Then

$$\lfloor r \rfloor < z < \lfloor r \rfloor + 1 \tag{17.6}$$

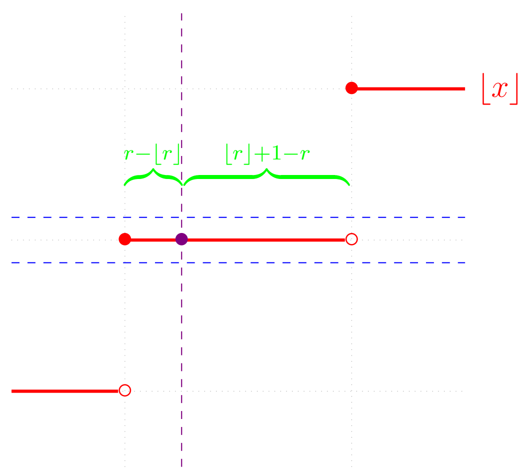
with $\lfloor r \rfloor \in \mathbb{Z}$. Thus, r is between two consecutive integers. To prove continuity of $\lfloor \cdot \rfloor$ at r, fix some $\epsilon > 0$.



Because $\lfloor r \rfloor < r < \lfloor r \rfloor + 1$, set

$$\delta := \min\big\{ r - \lfloor r \rfloor, \lfloor r \rfloor + 1 - r \big\}. \tag{17.7}$$

This choice is motivated by the following picture (the horizontal direction has been scaled differently for better viewing)



With this choice of δ, it follows that $\lfloor x \rfloor \in V_\epsilon\big(\lfloor r \rfloor\big)$ for all $x \in V_\delta(r)$. Therefore, $\lfloor \cdot \rfloor$ is continuous at $r \in \mathbb{Z}^c$.

The following two exercises were given in Professor Ben-Ari's advanced calculus course during a quiz.

Exercise 17.8. Let $f : A \to \mathbb{R}$ be a function on a domain A satisfying the condition

$$|f(x) - f(y)| \le |x^2 - y^2| \tag{17.9}$$

for all $x, y \in A$. Show that f is continuous.

Exercise 17.10. Let $f : A \to \mathbb{R}$ be a function on a domain A satisfying the condition

$$|f(x) - f(y)| \le |(x - y)^2| \tag{17.11}$$

for all $x, y \in A$. Show that f is continuous. In fact, what kind of a function is f?

Proposition 17.12. *Let $A \subseteq \mathbb{R}$ and let $f : A \to \mathbb{R}$ be a function. If $c \in A$ is an isolated point of A, then f is continuous at a.*

Proof. By definition of c being isolated, there exists a $\delta > 0$ such that $V_\delta(c) \cap A \setminus \{c\} = \varnothing$. Hence, for any $\epsilon > 0$, it is vacuously true that $x \in V_\delta(c) \cap A \setminus \{c\}$ implies $f(x) \in V_\epsilon\big(f(c)\big)$ since the former set is empty. ∎

Theorem 17.13. *Let $A \subseteq \mathbb{R}$, let $f : A \to \mathbb{R}$ be a function, and let $c \in A$. Then f is continuous at c if and only if for every sequence $a : \mathbb{N} \to A$ with $\lim a = c$, it follows that $\lim(f \circ a) = f(c)$. Furthermore, if c is a limit point of A, these conditions are equivalent to $\lim\limits_{x \to c} f(x) = f(c)$.*

Proof.
($\Rightarrow$) Suppose f is continuous at c and let $a : \mathbb{N} \to A$ be a sequence with $\lim a = c$. Fix $\epsilon > 0$. By assumption, there exists a $\delta > 0$ such that $x \in V_\delta(c) \cap A \setminus \{c\}$ implies $f(x) \in V_\epsilon\big(f(c)\big)$. Since a converges to c, there exists an $N \in \mathbb{N}$ such that $a_n \in V_\delta(c)$ for all $n \geq N$. Hence

$$\big|f(c) - f(a_n)\big| < \epsilon \qquad \forall\, n \geq N. \tag{17.14}$$

Note that if $a_n = c$, the left-hand-side of this inequality is zero, which is less than ϵ.
($\Leftarrow$) The proof will be via contrapositive. Namely, suppose that f is not continuous at c. Then there exists a $\epsilon > 0$ such that for all $\delta > 0$, there exists an $x \in V_\delta(c) \cap A$ with $f(x) \notin V_\epsilon\big(f(c)\big)$. In particular, for each $n \in \mathbb{N}$, setting $\delta_n := \frac{1}{n}$ provides the existence of elements $a_n \in V_{\delta_n}(c) \cap A$ satisfying $f(a_n) \notin V_\epsilon\big(f(c)\big)$. The sequence $a : \mathbb{N} \to \mathbb{R}$ satisfies $\lim a = c$ by construction. However, since $f(a_n) \notin V_\epsilon\big(f(c)\big)$ for all $n \in \mathbb{N}$, $\lim(f \circ a) \neq f(c)$. ∎

Remark 17.15. If f satisfies the condition "if for every sequence $a : \mathbb{N} \to A$ with $\lim a = c$, it follows that $\lim(f \circ a) = f(c)$," then f is said to be <u>*sequentially continuous at c*</u>. The distinction between this type of continuity and the definition we have presented is relevant on more general topological spaces, where these two definitions need not be equivalent.

Corollary 17.16. *Let $A \subseteq \mathbb{R}$, let $f : A \to \mathbb{R}$ be a function, and let c be a limit point of A with $c \in A$. f is discontinuous at c if and only if there exists a sequence $a : \mathbb{N} \to A$ with $\lim a = c$ but $f \circ a$ does not converge to $f(c)$.*

Proof. Exercise. ∎

Warning. Let $A \subseteq \mathbb{R}$ and let $f : A \to \mathbb{R}$ be a continuous function. If $a : \mathbb{N} \to A$ is a Cauchy sequence, it is *not* true that $f \circ a$ is always a Cauchy sequence. The homework asks you to give an example of such a domain, function, and sequence.

Theorem 17.17 (Algebraic Continuity Theorem)**.** *Let $A \subseteq \mathbb{R}$, let $f, g : A \to \mathbb{R}$, and let $c \in A$. Furthermore, suppose that f and g are continuous at c. Then the following facts follow.*

(a) The function kf is continuous at c for all $k \in \mathbb{R}$.

(b) The function $f + g$ is continuous at c.

(c) The function fg is continuous at c.

(d) *Let $B \subseteq A$ be the domain over which g is nonzero and so that $c \in B$. Then the function $\frac{f}{g}$ is continuous at c.*

Proof. Exercise. ■

Theorem 17.18. *Let $f : A \to B \subseteq \mathbb{R}$ and $g : B \to \mathbb{R}$ be functions continuous at $c \in A$ and $f(c) \in B$, respectively. Then the composition $g \circ f : A \to \mathbb{R}$ is continuous at c.*

It is more robust to prove this theorem using the definition of continuity rather than the sequential characterization.

Proof. Fix $\epsilon > 0$. By continuity of g at $f(c)$, there exists a $\delta > 0$ such that

$$y \in V_\delta\big(f(c)\big) \cap B \qquad \Rightarrow \qquad g(y) \in V_\epsilon\Big(g\big(f(c)\big)\Big). \tag{17.19}$$

By continuity of f at c, there exists a $\gamma > 0$ such that

$$x \in V_\gamma(c) \cap A \qquad \Rightarrow \qquad f(x) \in V_\delta\big(f(c)\big). \tag{17.20}$$

Therefore, if $x \in V_\gamma(c) \cap A$, then $f(x) \in V_\delta\big(f(c)\big) \cap B$, which implies $g\big(f(x)\big) \in V_\epsilon\Big(g\big(f(c)\big)\Big)$. ■

Exercise 17.21. Give a proof of Theorem 17.18 using the definition of sequential continuity for f and g.

Exercise 17.22. For each of the following conditions, give an example of, or state that the request is impossible, two functions $f : A \to B \subseteq \mathbb{R}$ and $g : B \to \mathbb{R}$ such that $f \circ g$ is continuous at $c \in A$ but

(a) f is not continuous at c.

(b) g is not continuous at $f(c)$.

(c) f is not continuous at c *and* g is not continuous at $f(c)$.

Definition 17.23. Let $A \subseteq \mathbb{R}$. A subset $U \subseteq A$ is said to be <u>*open in A*</u> if there exists an open set $V \subseteq \mathbb{R}$ with $V \cap A = U$.

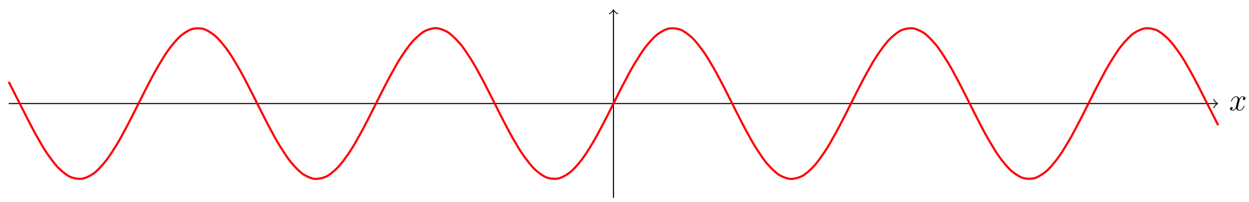
Theorem 17.24. *Let $A \subseteq \mathbb{R}$ and $f : A \to \mathbb{R}$ a function. Then f is continuous on A if and only if for every open set $U \subseteq \mathbb{R}$, the set $f^{-1}(U) \subseteq A$ is open in A.*

Proof.
($\Rightarrow$) Suppose $f : A \to \mathbb{R}$ is continuous. Let $U \subseteq \mathbb{R}$ be open and let $c \in f^{-1}(U)$. Since U is open, there exists an $\epsilon > 0$ such that $V_\epsilon\big(f(c)\big) \subseteq U$. By continuity of f at c, there exists a $\delta > 0$ such that $f(x) \in V_\epsilon\big(f(c)\big) \subseteq U$ for all $x \in V_\delta(c) \cap A$. But this implies $V_\delta(c) \cap A \subseteq f^{-1}(U)$. Since c was arbitrary, $f^{-1}(U)$ is open.
($\Leftarrow$) Suppose that $f^{-1}(U)$ is open for all open $U \subseteq \mathbb{R}$. Let $c \in A$. Fix $\epsilon > 0$. Then $V_\epsilon\big(f(c)\big)$ is an open set. By assumption $f^{-1}\Big(V_\epsilon\big(f(c)\big)\Big)$ is open and $c \in f^{-1}\Big(V_\epsilon\big(f(c)\big)\Big)$. Therefore, by definition of open sets and open sets in A, there exists a $\delta > 0$ such that $V_\delta(c) \cap A \subseteq f^{-1}\Big(V_\epsilon\big(f(c)\big)\Big)$. Therefore, $f\big(V_\delta(c)\big) \subseteq V_\epsilon\big(f(c)\big)$ proving continuity of f at c. Since c was arbitrary, this proves that f is continuous. ■

It is very important that the *preimage* of an open set is open for a function to be continuous. It is *not* true that a function is continuous if the *image* of an open set is open.

Definition 17.25. Let $A \subseteq \mathbb{R}$. A function $f : A \to \mathbb{R}$ is <u>*open*</u> if $f(U)$ is open in $\mathbb{R}$ for every open set U in A. f is <u>*closed*</u> if $f(C)$ is closed in $\mathbb{R}$ for every closed set C in A.[37]

Example 17.26. An example of a continuous function $f : \mathbb{R} \to \mathbb{R}$ that is neither open nor closed is $\mathbb{R} \ni x \mapsto f(x) := \sin(2\pi x)$.[38]



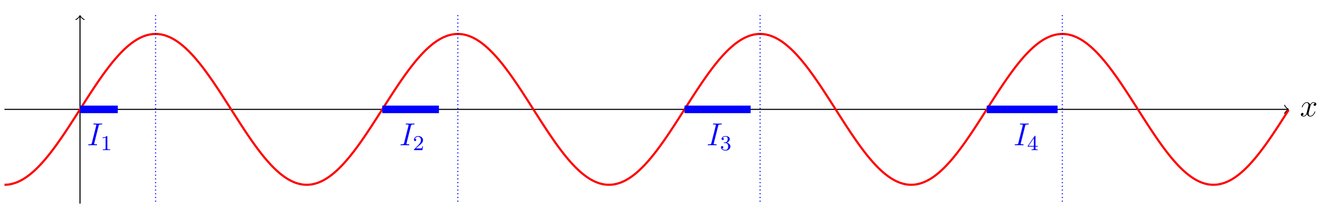
This is because

$$f\Big((0,1)\Big) = [-1,1] \qquad \& \qquad f\left(\bigcup_{n=1}^{\infty}\left[n-1, n-1+\frac{1}{4}\left(1-\frac{1}{2^{n+2}}\right)\right]\right) = [0,1). \tag{17.27}$$

The set

$$\bigcup_{n=1}^{\infty}\left[n-1, n-1+\frac{1}{4}-\frac{1}{2^{n+2}}\right] \equiv \bigcup_{n=1}^{\infty} I_n \tag{17.28}$$

is closed as can be seen by the following depiction of the individual components



The converse is not true either. Namely, an open function need not be continuous. To construct such a function, we state a definition and some facts that are surprising in their own right.

Definition 17.29. Let $A, B \subseteq \mathbb{R}$. A function $f : A \to B$ is a <u>*homeomorphism*</u> if f is a continuous bijection and $f^{-1} : B \to A$ is continuous.

Theorem 17.30. *Let $a, b \in \mathbb{R}$ with $a < b$. There exists a homeomorphism $f : (a,b) \to \mathbb{R}$.*

---

[37] Analogous to the definition of an open set in A is that of a closed subset of A. A subset C of A is said to be <u>*closed*</u> if there exists a closed set $B \subseteq \mathbb{R}$ such that $C = B \cap A$.

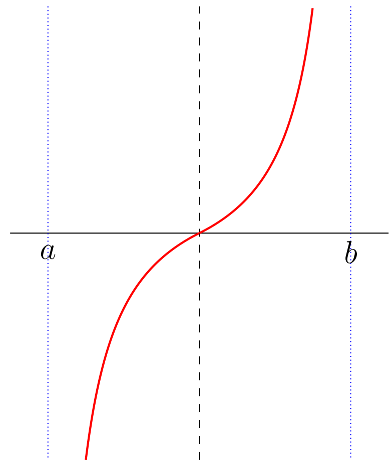
[38] Technically, we have not defined this function nor shown that it is continuous. We will be able to define it later geometrically in terms of integrals. If you are uncomfortable with this, instead, defined the function f to be

$$f(x) := \begin{cases} 1-x & \text{if } x \in \bigcup_{z\in\mathbb{Z}}\left[z+\frac{1}{4}, z+\frac{3}{4}\right] \\ x-1 & \text{if } x \in \bigcup_{z\in\mathbb{Z}}\left[z-\frac{1}{4}, z+\frac{1}{4}\right] \end{cases},$$

which achieves the same desired goal.

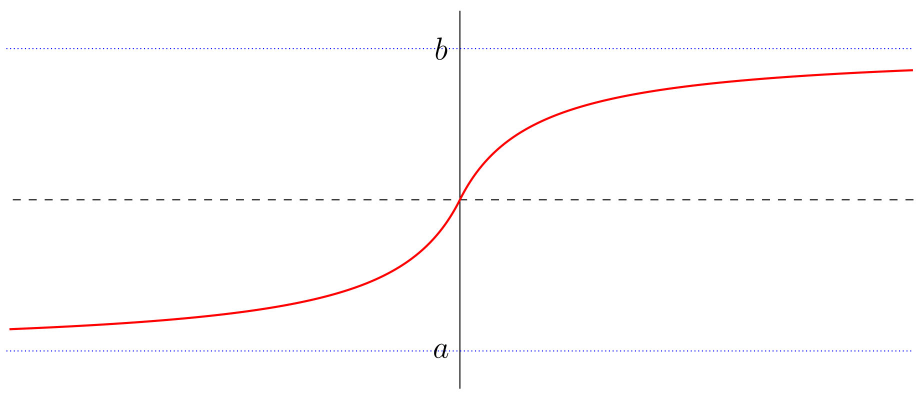
Proof. Let $f : (a,b) \to \mathbb{R}$ be the function

$$f(x) := \begin{cases} \frac{2x-(a+b)}{x-a} & \text{for } a < x \le \frac{a+b}{2} \\ \frac{2x-(a+b)}{b-x} & \text{for } \frac{a+b}{2} \le x < b \end{cases} \tag{17.31}$$



The continuity of this function is a consequence of problem 2 on HW #4 and the Algebraic Continuity Theorem (exercise). The inverse of this function is given by

$$f^{-1}(x) := \begin{cases} \frac{(a+b)-ax}{2-x} & \text{for } -\infty < x \le 0 \\ \frac{(a+b)+bx}{2+x} & \text{for } 0 \le x < \infty \end{cases} \tag{17.32}$$



Since this function is also a ratio of two continuous functions where the denominator never vanishes, it is continuous by the Algebraic Continuity Theorem. ■

Incidentally, this theorem shows that the image of a bounded set under a continuous function need not be bounded, not even under a homeomorphism.

Theorem 17.33. *Let $a \in \mathbb{R}$. There exist homeomorphisms $(-\infty, a) \to \mathbb{R}$ and $(a, \infty) \to \mathbb{R}$.*

Proof. Exercise. ■

Theorem 17.34. *Let $a, b, c, d \in \mathbb{R}$ with $a < b$ and $c < d$. There exists a homeomorphism $(a,b) \to (c,d)$.*

Proof. Exercise. ∎

Example 17.35. The present example came from StackExchange[39] Let C be the Cantor set described as the intersection $C = \bigcap_{n=1}^{\infty} C_n$ of a finite union of intervals. The complement of the Cantor set is an infinite union of disjoint open intervals (two of which are unbounded). Namely, using the notation from my solutions to HW #3, the Cantor set is given by

$$C = \bigcap_{n=1}^{\infty} \left(\bigcup_{\alpha_1,\dots,\alpha_n \in \mathbb{Z}^2} I_{\alpha_1 \dots \alpha_n} \right), \tag{17.36}$$

where the $I_{\alpha_1 \dots \alpha_n}$ are the closed intervals making up C_n, the n-th iterate for the construction of the Cantor set. By De Morgan's laws, the complement of this is given by

$$C^c = \bigcup_{n=1}^{\infty} \left(\bigcap_{\alpha_1,\dots,\alpha_n \in \mathbb{Z}^2} I^c_{\alpha_1 \dots \alpha_n} \right) = (-\infty, 0) \cup \left(\bigcup_{\lambda \in \Lambda} J_\lambda \right) \cup (1, \infty). \tag{17.37}$$

Here Λ is just some indexed set used to enumerate the open intervals J_λ in $[0,1] \cap C^c$. For each such open interval, let $f_\lambda : J_\lambda \to (-1,1)$ be a homeomorphism and let $f_- : (-\infty, 0) \to (-1,1)$ and $f_+ : (0,\infty) \to (-1,1)$ be homeomorphisms as well. Set $J_- := (-\infty, 0)$ and $J_+ := (0, \infty)$ and $\Lambda' := \Lambda \cup \{-\} \cup \{+\}$. Then the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) := \begin{cases} f_{\lambda'}(x) & \text{if } x \in J_{\lambda'} \text{ for some } \lambda' \in \Lambda' \\ 0 & \text{if } x \in C \end{cases} \tag{17.38}$$

is open but is discontinuous at every point of C.

After this lecture, it is recommended the student works through problems 3 and 4 on HW #4. Additional recommended exercises include exercises 4.3.6 and 4.3.8 from [2].

[39] http://math.stackexchange.com/questions/75589/open-maps-which-are-not-continuous.

18 October 27

Today we will state and prove both the Extreme Value Theorem (EVT) and the Intermediate Value Theorem (IVT). This is a slight departure in the presentation from the book—we will save different kinds of continuity for the next lecture.

Theorem 18.1. *Let $K \subseteq \mathbb{R}$ be a compact subset of $\mathbb{R}$ and let $f : K \to \mathbb{R}$ be continuous. Then $f(K) \subseteq \mathbb{R}$ is a compact subset of $\mathbb{R}$.*

Remark 18.2. This result should surprise you. This is because continuous functions preserve *neither* closed subsets *nor* bounded sets (remember that by the Heine-Borel theorem, compact is equivalent to closed *and* bounded). Indeed, Example 17.26 illustrates that continuous functions need not send closed subsets to closed subsets and Theorem 17.30 shows that continuous functions need not send bounded sets to bounded sets.

Abbott offers a proof of this theorem using the definition of compactness. Instead, we will use our characterization for what it means for a function to be continuous in terms of open sets.

Proof. Let $\mathcal{U} := \{U_\lambda\}_{\lambda\in\Lambda}$ be an open cover of $f(K)$. Since f is continuous, $\mathcal{V} := \{V_\lambda := f^{-1}(U_\lambda)\}_{\lambda\in\Lambda}$ is an open cover of K. Since K is compact, there exists a finite subcover $\{V_{\lambda_1}, \ldots, V_{\lambda_n}\}$ of $\mathcal{V}$ for K. Then $\{U_{\lambda_1}, \ldots, U_{\lambda_n}\}$ is a finite subcover of $\mathcal{U}$ for $f(K)$. ∎

Theorem 18.3 (Extreme Value Theorem)**.** *Let $K \subseteq \mathbb{R}$ be compact and let $f : K \to \mathbb{R}$ be a continuous function. Then there exist $x, z \in K$ such that $f(x) \le f(y) \le f(z)$ for all $y \in K$.*

Proof. Since $f(K)$ is compact by Theorem 18.1, it is closed and bounded. Hence, $M := \sup f(K)$ exists. Furthermore, since $f(K)$ is closed, $M \in f(K)$ so that there exists an $x \in K$ with $f(x) = M$. Similarly, $L := \inf f(K)$ exists and is in $f(K)$ so that there exists a $z \in K$ with $f(z) = L$. ∎

Lemma 18.4. *A subset $E \subseteq \mathbb{R}$ is disconnected if and only if there exist nonempty open sets $U \subseteq E$ and $V \subseteq E$ such that $U \cap V = \varnothing$ and $U \cup V = E$.*

Proof.
($\Rightarrow$) Suppose that (A, B) is a separation for E. Set

$$U := E \setminus (\overline{A} \cap E) \qquad \& \qquad V := E \setminus (\overline{B} \cap E). \tag{18.5}$$

Then U and V are open subsets of E, and they are nonempty and disjoint because

$$\begin{aligned} U \cap V &= \Big(E \setminus (\overline{A} \cap E)\Big) \cap \Big(E \setminus (\overline{B} \cap E)\Big) \\ &= E \setminus \Big((\overline{A} \cap E) \cup (\overline{B} \cap E)\Big) \\ &= E \setminus \Big(E \cap (\overline{A} \cup \overline{B})\Big) \\ &= E \setminus E \\ &= \varnothing \end{aligned} \tag{18.6}$$

by De Morgan's laws (as subsets of E) and a question from the first quiz in our class. Furthermore, their union is

$$U \cup V = \Big(E \setminus (\overline{A} \cap E)\Big) \cup \Big(E \setminus (\overline{B} \cap E)\Big) = (\overline{A} \cap E) \cap (\overline{B} \cap E). \tag{18.7}$$

again by De Morgan's laws. Also, note that $\overline{A} \subseteq A$ because $\overline{A} \cap B = \varnothing$ and $A = E \setminus B$. Hence,

$$U \cup V = (\overline{A} \cap E) \cap (\overline{B} \cap E) \subseteq A \cap \overline{B} \cap E = \varnothing \tag{18.8}$$

because (A, B) is a separation for E.
($\Leftarrow$) Suppose that there exist nonempty disjoint open subsets U and V of E such that $U \cup V = E$. Set

$$A := E \setminus U \qquad \& \qquad B := E \setminus V. \tag{18.9}$$

Then A and B are disjoint and closed subsets of E. Because they're closed, $(\overline{A} \cap E) \cap B = \varnothing$ and $(\overline{B} \cap E) \cap A = \varnothing$. Finally, $A \cup B = U \cup V = E$ by De Morgan's laws. Thus (A, B) is a separation for E. ∎

Theorem 18.10. *Let $A \subseteq \mathbb{R}$ be a connected subset of $\mathbb{R}$ and let $f : A \to \mathbb{R}$ be continuous. Then $f(A) \subseteq \mathbb{R}$ is a connected subset of $\mathbb{R}$.*

Proof. Suppose to the contrary that $f(A)$ is disconnected. Then by Lemma 18.4, there exist disjoint nonempty open sets $U, V \subseteq \mathbb{R}$ such that

$$\big(U \cap f(A)\big) \cup \big(V \cap f(A)\big) = f(A). \tag{18.11}$$

Because f is continuous $f^{-1}(U)$ and $f^{-1}(V)$ are open. Furthermore, they are disjoint by definition of the inverse image and because U and V are disjoint. Finally,

$$f^{-1}(U) \cup f^{-1}(V) = f^{-1}(U \cup V) = f^{-1}\big(f(A)\big) = A \tag{18.12}$$

contradicting the assumption that A is connected. ∎

Theorem 18.13 (Intermediate Value Theorem)**.** *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then for any L satisfying either $f(a) < L < f(b)$ or $f(a) > L > f(b)$, there exists a $c \in (a, b)$ with $f(c) = L$.*

Proof. Let $L \in \big(f(a), f(b)\big)$.[40] By Theorem 18.10, $f\big([a, b]\big)$ is connected and contains $f(a)$ and $f(b)$. By Theorem 12.10, $\big[f(a), f(b)\big] \subseteq f\big([a, b]\big)$. Hence, for any $L \in \big(f(a), f(b)\big)$, it follows that $L \in f\big([a, b]\big)$. Therefore, there exists a $c \in (a, b)$ such that $f(c) = L$. ∎

Definition 18.14. Let $A \subseteq \mathbb{R}$. A function $f : A \to \mathbb{R}$ is *non-decreasing* if $f(x) \leq f(y)$ for all $x, y \in A$ with $x \leq y$. f is *non-increasing* if $f(x) \geq f(y)$ for all $x, y \in A$ with $x \geq y$. f is *monotone* if it is either non-increasing or non-decreasing.

Before giving an example of such a function, we discuss some properties of monotone functions.

[40] Note that there is nothing to prove if $f(a) = f(b)$. The following proof is modified from what it was before and is based on Kevin Pratt's idea.

Definition 18.15. Let $a, b \in \mathbb{R}$ with $a \leq b$. A function $f : [a, b] \to \mathbb{R}$ has the *intermediate value property* if for all $x, y \in [a, b]$ with $x < y$ and all $L \in \big(f(x), f(y)\big)$, there exists a $c \in (x, y)$ such that $f(c) = L$.

Theorem 18.16. *Let $a, b \in \mathbb{R}$ with $a \leq b$. Let $f : [a, b] \to \mathbb{R}$ be a monotone function with the intermediate value property. Then f is continuous.*

Proof. Exercise. ∎

Definition 18.17. Let $f : A \to \mathbb{R}$ be a function. Let

$$D_f := \{x \in A \ : \ f \text{ is discontinuous at } x\} \tag{18.18}$$

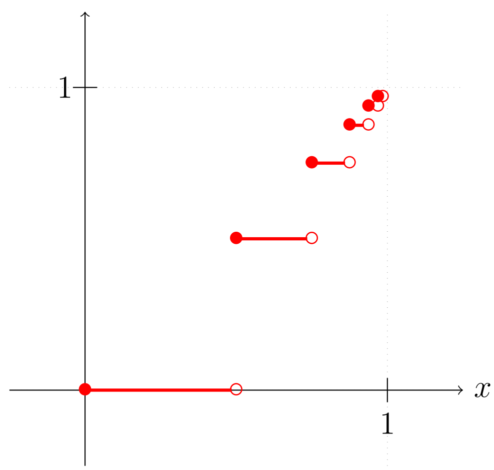
be the set of points at which f is discontinuous.

Example 18.19. For each $n \in \mathbb{N}$, define a function

$$\left[1 - \frac{1}{2^{n-1}}, 1 - \frac{1}{2^n}\right) \ni x \mapsto f_n(x) := \begin{cases} 0 & \text{if } n = 1 \\ \frac{1}{2^{n-1}} & \text{otherwise} \end{cases} \tag{18.20}$$

The function $f : [0, 1) \to \mathbb{R}$ defined by

$$[0, 1) \ni x \mapsto f(x) := f_n(x) \text{ if } x \in \left[1 - \frac{1}{2^{n-1}}, 1 - \frac{1}{2^n}\right). \tag{18.21}$$



Then f is non-decreasing and has a countable set of discontinuities.

The set of discontinuities on $\mathbb{R}$ can be separated into three types of interest. To define these types, we define left and right limits.

Definition 18.22. Let $A \subseteq \mathbb{R}$, let $f : A \to \mathbb{R}$, and let c be a limit point of A. The *left-hand limit of f as x approaches c* (a.k.a. *limit of f as x approaches c from the left*) is a real number $L_- := \lim_{x \to c^-} f(x)$ satisfying the condition that for every $\epsilon > 0$, there exists a $\delta > 0$ such that $f(x) \in V_\epsilon(L_-)$ for all $x \in (c - \delta, c) \cap A$. Similarly, the *right-hand limit of f as x approaches c* (a.k.a. *limit of f as x approaches c from the right*) is a real number $L_+ := \lim_{x \to c^+} f(x)$ satisfying the condition that for every $\epsilon > 0$, there exists a $\delta > 0$ such that $f(x) \in V_\epsilon(L_+)$ for all $x \in (c, c+\delta) \cap A$.

Theorem 18.23. *Let $A \subseteq \mathbb{R}$, let $f : A \to \mathbb{R}$, and let c be a limit point of A. Then*

$$\lim_{x\to c} f(x) = L \tag{18.24}$$

if and only if

$$\lim_{x\to c^-} f(x) = L \qquad \& \qquad \lim_{x\to c^+} f(x) = L. \tag{18.25}$$

Proof. Exercise. ■

Definition 18.26. Let $A \subseteq \mathbb{R}$, let $f : A \to \mathbb{R}$, and let c be a limit point of A with $c \in A$.

(a) If $\lim\limits_{x\to c} f(x)$ exists but does not equal $f(c)$, then f is said to have a *removable* discontinuity at c.

(b) If $\lim\limits_{x\to c^-} f(x)$ and $\lim\limits_{x\to c^+} f(x)$ both exist but are not equal, then f is said to have a *jump* discontinuity at c.

(c) If $\lim\limits_{x\to c} f(x)$ does not exist for some other reason, then f is said to have a *essential* discontinuity at c.

Theorem 18.27. *Let $C \subseteq \mathbb{R}$ and let $f : C \to \mathbb{R}$ be a non-decreasing function and let $c \in C$ be a limit point of C such that*

$$(-\infty, c) \cap C \neq \varnothing \qquad \& \qquad (c, \infty) \cap C \neq \varnothing. \tag{18.28}$$

Then $\lim\limits_{x\to c^-} f(x)$ and $\lim\limits_{x\to c^+} f(x)$ both exist and satisfy

$$\lim_{x\to c^-} f(x) \le \lim_{x\to c^+} f(x). \tag{18.29}$$

In particular, the only allowed discontinuities for a non-decreasing function are jump discontinuities.

Proof. Since f is non-decreasing and (18.28) holds, the sets

$$A := f\big((-\infty, c) \cap C\big) \qquad \& \qquad B := f\big(C \cap (c, \infty)\big) \tag{18.30}$$

are both nonempty and bounded from above and below, respectively. Then the theorem will follow if it is shown that

$$\lim_{x\to c^-} f(x) = \sup A \qquad \& \qquad \lim_{x\to c^+} f(x) = \inf B \tag{18.31}$$

since the latter quantities exist by the Axiom of Completeness. Set

$$s := \sup A \qquad \& \qquad i := \inf B. \tag{18.32}$$

Fix $\epsilon_- > 0$. By Lemma 3.31, there exists an element $a \in A$ such that $s - \epsilon_- < a$. But by definition of A, this means there exists an element $y \in A$ such that $f(y) = a$. Thus, $s - \epsilon_- < f(y)$. Set $\delta_- := c - y$. Note that $\delta_- > 0$ because $y \in (-\infty, c) \cap C$. Hence, for all $x \in (c - \delta_-, c) \cap C$, it follows that

$$s - \epsilon_- < f(y) = f(c - \delta_-) \le f(x) \le s < s + \epsilon_- \tag{18.33}$$

because f is non-decreasing. Thus $f(x) \in V_{\epsilon_-}(s)$ showing that $\lim_{x \to c^-} f(x) = s$. A similar argument shows that $\lim_{x \to c^+} f(x) = \inf B$.

The last claim follows from the fact that the elements of A are all lower bounds for B because f is non-decreasing. Thus, A is contained in the set $\mathcal{L}_B$ of *all* lower bounds of B. By a problem from our first homework set, $\sup \mathcal{L}_B = \inf B$ (this was problem 4). By another exercise earlier in the semester, since $A \subseteq \mathcal{L}_B$, it follows that $\sup A \leq \sup \mathcal{L}_B$. Combining these two inequalities gives $\sup A \leq \inf B$. ■

Theorem 18.34. *Let $f : A \to \mathbb{R}$ be a non-decreasing function. Then D_f is at most countable.*

Proof. Let $c \in D_f$. Then

$$\lim_{x \to c^-} f(x) < \lim_{x \to c^+} f(x) \tag{18.35}$$

by the previous theorem and the fact that c is a discontinuity. By the Density of $\mathbb{Q}$ in $\mathbb{R}$, there exists a rational number r_c satisfying

$$\lim_{x \to c^-} f(x) < r_c < \lim_{x \to c^+} f(x). \tag{18.36}$$

Because f is non-decreasing, the function

$$\begin{aligned} D_f &\to \mathbb{Q} \\ c &\mapsto r_c \end{aligned} \tag{18.37}$$

is one-to-one. This identifies D_f with a subset of $\mathbb{Q}$ and is therefore at most countable. ■

Remark 18.38. Given a countable subset C of (a, b) with $a < b$ and $a, b \in \mathbb{R}$, Rudin [6] (see Remark 4.31) provides a surprising construction of a *monotone* function $f : (a, b) \to \mathbb{R}$ for which f is discontinuous at all points in C.

Theorem 18.39. *Let $a, b \in \mathbb{R}$ with $a < b$. Let $f : (a, b) \to \mathbb{R}$ be a continuous 1-1 function. Then f is monotone. Let $g : [a, b] \to \mathbb{R}$ be a continuous 1-1 function. Then g is monotone and attains its maximum and minimum at a and b.*

Proof. Exercise. ■

After this lecture, it is recommended the student works through problems 5 and 7 on HW #4. It goes without saying that the reader should also attempt the exercises listed in these notes.

19 November 1

Today we will discuss different kinds of continuity for functions, their properties, and their relations to each other. To be pedantic, we will review the definition of continuity itself but this time be very careful about the notation to illustrate how every quantity depends on every other quantity.

Definition 19.1. Let $A \subseteq \mathbb{R}$ and let $f : A \to \mathbb{R}$ be a function. f is continuous at $c \in A$ if for every $\epsilon > 0$, there exists a $\delta(c, \epsilon) > 0$ such that $f(x) \in V_\epsilon\big(f(c)\big)$ for all $x \in V_{\delta(c,\epsilon)}(c) \cap A$. f is continuous on A if f is continuous at c for all $c \in A$.

Remark 19.2. Note that the δ in the definition of continuity depends on *both* the point $c \in A$ and the number ϵ. To be honest, I should have been careful from the start about this. In other words, if I wanted to show that a function f is continuous on A, if you provide me with a family of ϵ's parametrized by the elements $c \in A$, I would have to come up with a δ that may depend on both c and ϵ.

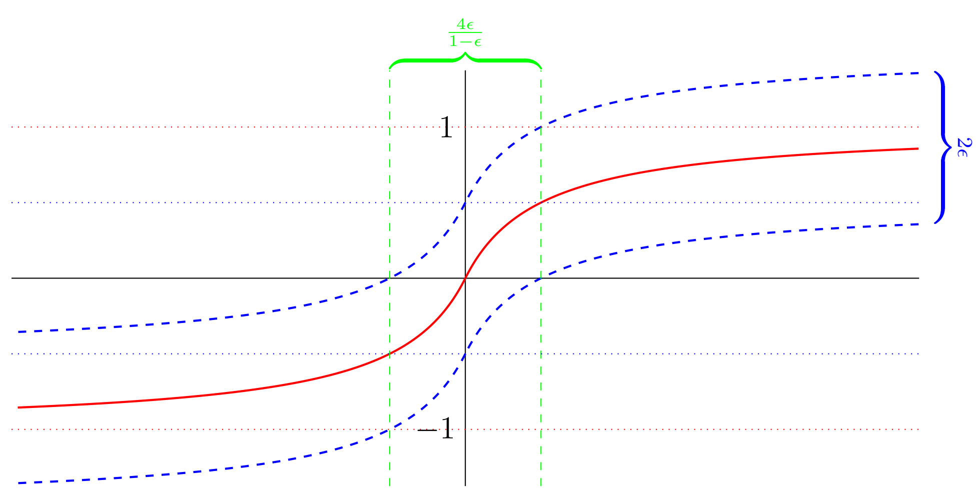
Definition 19.3. Let $A \subseteq \mathbb{R}$ and let $f : A \to \mathbb{R}$ be a function. f is uniformly continuous on A if for every $\epsilon > 0$ there exists a $\delta(\epsilon) > 0$ such that $|f(x) - f(y)| < \epsilon$ for all pairs of elements $x, y \in A$ with $|x - y| < \delta(\epsilon)$.

Remark 19.4. Using this more precise notation, we see that the definition of uniform continuity no longer depends on the point $c \in A$ but depends only on the number ϵ. In other words, if I wanted to show that a function f is uniformly continuous on A, if you provide me with a *single* ϵ, I would have to come up with a δ that depends only on ϵ.

Example 19.5. The function f^{-1} from the proof of Theorem 17.30 is uniformly continuous. As a simplification, we set $a = -1$ and $b = 1$ so that

$$\mathbb{R} \ni x \mapsto f^{-1}(x) := \begin{cases} \frac{x}{2-x} & \text{for } -\infty < x \leq 0 \\ \frac{x}{2+x} & \text{for } 0 \leq x < \infty \end{cases} \tag{19.6}$$

To see that this function is uniformly continuous, fix $\epsilon > 0$.



It is apparent from a visualization of this function that the steepest curvature occurs at $x = 0$. Hence, if we find an appropriate $\delta(\epsilon)$ at $x = 0$, this $\delta(\epsilon)$ should work for all other points. Fortunately, this $\delta(\epsilon)$ is easy to solve for because the function is increasing.[41] Furthermore, we know the inverse, so we just plug in $x = \epsilon$ for the right-hand-side of the piece-wise defined function:

$$\delta(\epsilon) := \frac{2\epsilon}{1-\epsilon} \qquad \text{for } \epsilon < 1. \tag{19.7}$$

If $\epsilon \geq 1$, any number $\delta(\epsilon) > 0$ will work. To check that this $\delta(\epsilon)$ works at every point, there are four cases to check since any pair of numbers $x, y \in \mathbb{R}$ could be in either $(-\infty, 0]$ or $[0, \infty)$. However, by symmetry, it suffices to check only two cases. Thus, first suppose $x, y \in [0, \infty)$ with $x \leq y$ and $|y - x| < \delta(\epsilon)$. Set

$$\alpha := y - x. \tag{19.8}$$

By these assumptions, $0 < \alpha < \delta(\epsilon)$. Then,

$$\begin{aligned}
\left|f^{-1}(y) - f^{-1}(x)\right| &= \frac{y}{2+y} - \frac{x}{2+x} \qquad \text{since } f^{-1} \text{ is increasing and } x < y \\
&= \frac{y(2+x) - x(2+y)}{(2+y)(2+x)} \\
&= \frac{(x+\alpha)(2+x) - x(2+x+\alpha)}{(2+x+\alpha)(2+x)} \\
&= \frac{2\alpha}{(2+x)^2 + (2+x)\alpha} \\
&\leq \frac{\alpha}{2+\alpha} \qquad \text{since } x \geq 0 \\
&< \frac{\delta}{2+\delta} \qquad \text{since } f^{-1} \text{ is strictly increasing and } \alpha < \delta \\
&= \frac{2\epsilon}{2(1-\epsilon) + 2\epsilon} \\
&= \epsilon
\end{aligned} \tag{19.9}$$

The other cases are treated similarly and are left to the reader. Thus, f^{-1} is uniformly continuous.

Before we give an example of a continuous function that is not uniformly continuous, let us explicitly say what it means for $f : A \to \mathbb{R}$ to not be uniformly continuous on A. f is not uniformly continuous on A if and only if there exists an $\epsilon > 0$ for which there is no $\delta(\epsilon) > 0$ satisfying the condition that $\left|f(x) - f(y)\right| < \epsilon$ for all $x, y \in A$ satisfying $|x - y| < \delta(\epsilon)$. In other words, f is not uniformly continuous on A if and only if there exists an $\epsilon > 0$ such that for all $\delta(\epsilon) > 0$, there exist $x, y \in A$ satisfying $|x - y| < \delta(\epsilon)$ but $\left|f(x) - f(y)\right| \geq \epsilon$.

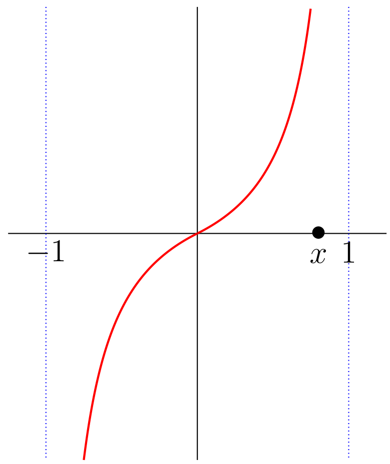
Example 19.10. The function $f : (a, b) \to \mathbb{R}$ from the proof of Theorem 17.30 is *not* uniformly continuous. Again, for simplicity, we set $a = -1$ and $b = 1$. In this case, this function is given by

$$f(x) := \begin{cases} \frac{2x}{x+1} & \text{for } -1 < x \leq 0 \\ \frac{2x}{1-x} & \text{for } 0 \leq x < 1 \end{cases} \tag{19.11}$$

[41]Exercise: prove that $f^{-1} : \mathbb{R} \to (-1, 1)$ is strictly increasing.

To see that f is not uniformly continuous, it suffices to find an $\epsilon > 0$ for which there does not exist a $\delta(\epsilon) > 0$ satisfying $|f(x) - f(y)| < \epsilon$ for all $x, y \in (-1, 1)$ satisfying $|x - y| < \delta(\epsilon)$. It turns out that we can do better. We will show that for *every* $\epsilon > 0$, there does not exist a $\delta(\epsilon) > 0$ satisfying the above property. Thus, fix $\epsilon > 0$ and fix $\delta(\epsilon) > 0$. It suffices to assume that $\delta(\epsilon) < 2$ to begin with since the domain of f is $(-1, 1)$. Then, set

$$x := 1 - \delta(\epsilon). \tag{19.12}$$



If $x \geq 0$, then set y to be any number in the range[42]

$$1 - \frac{2\delta(\epsilon)}{2 + \epsilon\delta(\epsilon)} < y < 1. \tag{19.13}$$

The following two inequalities obtained from this are useful:

$$\frac{1}{1-y} < \frac{2\delta(\epsilon)}{2 + \epsilon\delta(\epsilon)} \qquad \& \qquad y\big(2 + \epsilon\delta(\epsilon)\big) > 2 + \epsilon\delta(\epsilon) - 2\delta(\epsilon). \tag{19.14}$$

Note that such y satisfies $1 - 2\delta(\epsilon) < y < 1$ (in fact, it satisfies $1 - \delta(\epsilon) < y < 1$). With this choice of y,

$$\begin{aligned} \big|f(y) - f(x)\big| &= \frac{2y}{1-y} - \frac{2(1-\delta(\epsilon))}{\delta(\epsilon)} \\ &= \frac{2\delta(\epsilon) - 2 + 2y}{(1-y)\delta(\epsilon)} \\ &> \frac{\big(2\delta(\epsilon) - 2 + 2y\big)\big(2 + \epsilon\delta(\epsilon)\big)}{2\delta(\epsilon)^2} \\ &= \frac{\big(2\delta(\epsilon) - 2\big)\big(2 + \epsilon\delta(\epsilon)\big) + 2y\big(2 + \epsilon\delta(\epsilon)\big)}{2\delta(\epsilon)^2} \\ &> \epsilon. \end{aligned} \tag{19.15}$$

[42]This choice was motivated by considering the desired inequality

$$\frac{2y}{1-y} - \frac{2(1-\delta(\epsilon))}{\delta(\epsilon)} > \epsilon.$$

In other words, the two previous examples illustrate that even if a homeomorphism is uniformly continuous, the inverse of the homeomorphism need not be.

Proposition 19.16. *Let $f : A \to \mathbb{R}$ be a uniformly continuous function and let $B \subseteq A$. Then the restriction of f to B is uniformly continuous.*

Proof. Exercise. ∎

From the previous examples, one might conjecture that a *bounded* continuous function is uniformly continuous. This is not the case. However, it is a little difficult to prove this using ϵ's and δ's as we have done above in Example 19.10. To help us, we will provide a test for lack of uniform continuity in terms of convergent sequences, similarly to how convergent sequences can be used to prove that a particular function is not continuous.

Theorem 19.17. *Let $A \subseteq \mathbb{R}$ and let $f : A \to \mathbb{R}$ be a function. f is not uniformly continuous on A if and only if there exists an $\epsilon > 0$ and two sequences $a, b : \mathbb{N} \to A$ satisfying*

$$\lim(b - a) = 0 \qquad \& \qquad \big|f(a_n) - f(b_n)\big| \geq \epsilon \quad \forall\, n \in \mathbb{N}. \tag{19.18}$$

Proof.
($\Rightarrow$) Suppose f is not uniformly continuous. Fix $\epsilon > 0$. Then, by the discussion preceding Example 19.10, for every $n \in \mathbb{N}$, there exists a pair of elements $a_n, b_n \in A$ with $|a_n - b_n| < \frac{1}{n}$ satisfying $\big|f(a_n) - f(b_n)\big| \geq \epsilon$. By construction, these sequences satisfy the required conditions.
($\Leftarrow$) Suppose there exists an $\epsilon > 0$ and two sequences $a, b : \mathbb{N} \to A$ satisfying (19.18). Then, for any $\delta > 0$, since the sequence $a - b$ converges to zero, there exists an $N \in \mathbb{N}$ such that $|a_n - b_n| < \delta$ for all $n \geq N$. By assumption $\big|f(a_n) - f(b_n)\big| \geq \epsilon$ showing that f is not uniformly continuous. ∎

Example 19.19. An example of a continuous bounded function that is not uniformly continuous is

$$(0, 1] \ni x \mapsto f(x) := \sin\left(\frac{2\pi}{x}\right) \tag{19.20}$$

as is discussed in Abbott's book.

Remark 19.21. This function is also useful in topology where it provides several counterexamples (its graph with the origin included is connected but not path-connected). It is known as the topologist's sine curve.

Exercise 19.22. Prove or disprove the following. Let $A, B \subseteq \mathbb{R}$ with B bounded and let $f : A \to B$ be a homeomorphism. Then f is uniformly continuous.

Theorem 19.23 (Uniform continuity on compact sets)**.** *Let $K \subseteq \mathbb{R}$ be a compact set and let $f : K \to \mathbb{R}$ be a continuous function. Then f is uniformly continuous.*

Proof. See Abbott. ∎

Definition 19.24. A function $f : A \to \mathbb{R}$ is called <u>Lipschitz</u> if there exists an $M \in \mathbb{R}$ such that

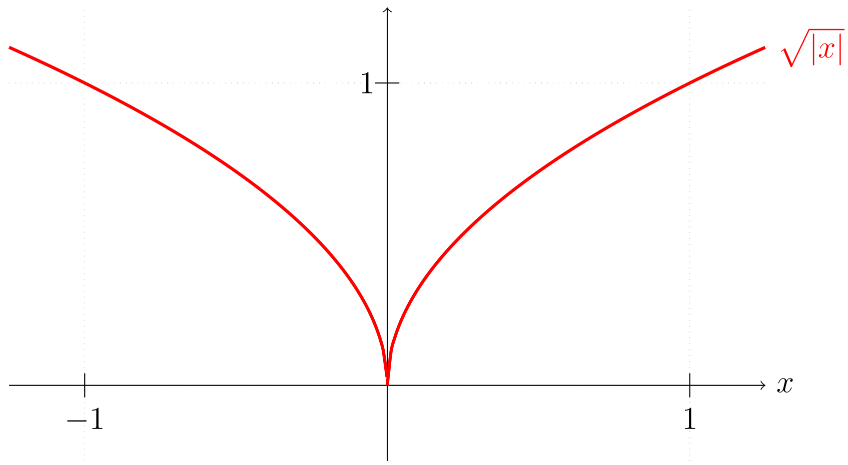
$$\left|\frac{f(x) - f(y)}{x - y}\right| \leq M \tag{19.25}$$

for all $x, y \in A$ with $x \neq y$.

Theorem 19.26. *Let $A \subseteq \mathbb{R}$ and let $f : A \to \mathbb{R}$ be a function. If f is Lipschitz continuous the f is uniformly continuous. If f is uniformly continuous then f is continuous. Neither of the converse statements are true.*

Proof. Exercise. ■

Instead of proving this theorem in full, we merely provide an example of a uniformly continuous function that is not Lipschitz.



and leave it to the reader to show that this function satisfies the mentioned properties.

After this lecture, it is recommended the student works through problem 6 on HW #4.

20 November 3

It is often useful to express physical phenomena in terms of differential equations. These equations describe the rates of change of particular quantities depending on at least one variable. For example, Schrödinger's equation describes the rate of change of a vector ψ in some (usually infinite-dimensional) vector space subject to dynamics described by a self-adjoint operator H on that vector space

$$\sqrt{-1}\frac{d\psi(t)}{dt} = (H\psi)(t). \tag{20.1}$$

Knowing the rate of change of such a quantity, one can, in principle, predict the future state of a system given some initial condition. Only the rate of change is sufficient to determine the future, which is why Schrödinger's equation is a first order differential equation. Incompressible fluid flow can be described as a second order differential equation (known as the Navier-Stokes equation) by

$$\frac{\partial \vec{u}}{\partial t} + \langle \vec{u}, \nabla \cdot \vec{u} \rangle = -\frac{\nabla P}{\rho} + \nu \nabla^2 \vec{u}, \tag{20.2}$$

where $\vec{u}(x,y,z,t)$ is the velocity of the fluid at some position (x,y,z) at the moment in time t, P is the pressure (also a function of space and time), ρ is the density of the fluid, ν is some constant, and ∇ is the gradient. As exciting as these and related differential equations are, we will first dedicate some time to understanding the notion of a rate of change of a function.

Remark 20.3. It is worth pointing out that it is not known whether there always exist reasonable solutions to the Navier-Stokes equation under some physically reasonable conditions. This question has been deemed so important that it has a bounty of $1,000,000.

Definition 20.4. Let $A \subseteq \mathbb{R}$ and $f : A \to \mathbb{R}$. Let c be a limit point of A with $c \in A$. Then the *derivative of f at c* is defined to be the functional limit

$$f'(c) := \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \tag{20.5}$$

if it exists. When the derivative of f at c exists, f is said to be *differentiable at c*. If f is differentiable at c for all $c \in A$, then f is *differentiable on A*.

Proposition 20.6. *$f : A \to \mathbb{R}$ is differentiable at $c \in A$ if and only if there exists a linear function*[43] *$D_c f : \mathbb{R} \to \mathbb{R}$ such that*

$$\lim_{x \to c} \frac{f(x) - f(c) - (D_c f)(x - c)}{x - c} = 0. \tag{20.7}$$

Proof.
($\Rightarrow$) Suppose f is differentiable at c with derivative $f'(c)$. Define the function $D_c f : \mathbb{R} \to \mathbb{R}$ to be

$$\mathbb{R} \ni y \mapsto (D_c f)(y) := f'(c) y. \tag{20.8}$$

[43] A function $g : \mathbb{R} \to \mathbb{R}$ is *linear* if $g(ax + y) = ag(x) + g(y)$ for all $a, x, y \in \mathbb{R}$.

Then, by the Algebraic Limit Theorem for functional limits

$$\begin{aligned}\lim_{x\to c}\frac{f(x)-f(c)-(D_cf)(x-c)}{x-c} &= \lim_{x\to c}\left(\frac{f(x)-f(c)}{x-c}-f'(c)\right) \quad \text{by (20.8)}\\ &= \lim_{x\to c}\left(\frac{f(x)-f(c)}{x-c}\right)-\lim_{x\to c}\big(f'(c)\big) \quad \text{by Corollary 17.1}\\ &= f'(c)-f'(c)\\ &= 0.\end{aligned} \tag{20.9}$$

($\Leftarrow$) Suppose that there exists a $D_cf:\mathbb{R}\to\mathbb{R}$ as in the description. Set

$$f'(c) := (D_cf)(1). \tag{20.10}$$

Then, by a similar calculation to the previous direction,

$$0=\lim_{x\to c}\frac{f(x)-f(c)-(D_cf)(x-c)}{x-c}=\lim_{x\to c}\left(\frac{f(x)-f(c)}{x-c}-f'(c)\right). \tag{20.11}$$

By Exercise 2.3.10 part (a)[44] in [2] and the sequential characterization for limits (see Theorem 16.9), this implies

$$\lim_{x\to c}\frac{f(x)-f(c)}{x-c}=\lim_{x\to c}f'(c)=f'(c). \tag{20.12}$$

■

Definition 20.13. Let $f:A\to\mathbb{R}$ be differentiable at c. The linear function $D_cf:\mathbb{R}\to\mathbb{R}$ from Proposition 20.6 is called the *differential of f at c*.

Remark 20.14. This equivalent definition of the derivative provides a useful perspective when generalizing to higher-dimensional derivatives.[45] We will also see in a moment that the chain rule is expressed particularly elegantly in terms of differentials.

Example 20.15. See Abbott.

Theorem 20.16. *Let $f:A\to\mathbb{R}$ be a function that is differentiable at a limit point c of A with $c\in A$. Then f is continuous at c.*

For this proof, it will be useful to recall the equivalent notions of continuity from Theorem 17.13 so that we may avoid an explicit ϵ and δ proof for continuity.

Proof. By the Algebraic Limit Theorem for functional limits

$$\begin{aligned}\lim_{x\to c}\big(f(x)-f(c)\big) &= \lim_{x\to c}\left(\left(\frac{f(x)-f(c)}{x-c}\right)(x-c)\right)\\ &= \lim_{x\to c}\left(\frac{f(x)-f(c)}{x-c}\right)\lim_{x\to c}\big((x-c)\big) \quad \text{by Corollary 17.1}\\ &= \big(f'(c)\big)(0)\\ &= 0,\end{aligned} \tag{20.17}$$

[44]This exercise says that if two sequences $a,b:\mathbb{N}\to\mathbb{R}$ satisfy $\lim(a-b)=0$, then $\lim a=\lim b$.

[45]If you recall from calculus, if you have a function $f:\mathbb{R}^3\to\mathbb{R}$, then ∇f, the gradient of f, has as its input a *vector* $\vec{v}$ at some point $x\in\mathbb{R}^3$ so that $(\nabla_{\vec{v}}f)(x)$ describes the derivative of f at x in the direction $\vec{v}$. However, to be consistent with our notation, we would instead write this as $(\nabla_x f)(\vec{v})$.

which, again by Exercise 2.3.10 part (a) in [2] and the sequential characterization for limits (see Theorem 16.9), this shows that

$$\lim_{x\to c} f(x) = f(c). \tag{20.18}$$

This proves that f is continuous at c due to the functional limit characterization of continuity from Theorem 17.13. ■

Theorem 20.19 (Chain rule)**.** *Let $A, B \subseteq \mathbb{R}$, let $f : A \to \mathbb{R}$ with $f(A) \subseteq B$ and let $g : B \to \mathbb{R}$. Let c be a limit point of A with $c \in A$ and let $f(c)$ be a limit point of B. Suppose that f is differentiable at c and g is differentiable at $f(c)$. Then $g \circ f$ is differentiable at c and its derivative is given by*

$$(g \circ f)'(c) = g'\big(f(c)\big)f'(c). \tag{20.20}$$

The chain rule takes a particularly appealing form if expressed using the equivalent definition of the derivative in terms of linear functions (see Proposition 20.6).

Theorem 20.21 (Differential form of the chain rule)**.** *Let $A, B \subseteq \mathbb{R}$, let $f : A \to \mathbb{R}$ with $f(A) \subseteq B$ and let $g : B \to \mathbb{R}$. Let c be a limit point of A with $c \in A$ and let $f(c)$ be a limit point of B. Suppose that f is differentiable at c with differential $D_c f$ and g is differentiable at $f(c)$ with differential $D_{f(c)}g$. Then $g \circ f$ is differentiable at c with differential*

$$D_c(g \circ f) = D_{f(c)}g \circ D_c f. \tag{20.22}$$

Proof. The goal is to show[46]

$$\lim_{x\to c} \left(\frac{(g \circ f)(x) - (g \circ f)(c) - (D_{f(c)}g \circ D_c f)(x - c)}{x - c} \right) = 0. \tag{20.23}$$

Set

$$h(x) := \frac{(g \circ f)(x) - (g \circ f)(c) - (D_{f(c)}g \circ D_c f)(x - c)}{x - c}. \tag{20.24}$$

Then h splits into two terms as follows

$$\begin{aligned} h(x) &= \frac{g\big(f(x)\big) - g\big(f(c)\big) - (D_{f(c)}g)\Big(f(x) - f(c) - \big(f(x) - f(c) - (D_c f)(x - c)\big)\Big)}{x - c} \\ &= \frac{g\big(f(x)\big) - g\big(f(c)\big) - (D_{f(c)}g)\big(f(x) - f(c)\big)}{x - c} + (D_{f(c)}g)\left(\frac{f(x) - f(c) - (D_c f)(x - c)}{x - c}\right). \end{aligned} \tag{20.25}$$

Focusing on the second term, because $D_{f(c)}g$ is linear,

$$\begin{aligned} \lim_{x\to c}(D_{f(c)}g)\left(\frac{f(x) - f(c) - (D_c f)(x - c)}{x - c}\right) &= D_{f(c)}g\left(\lim_{x\to c} \frac{f(x) - f(c) - (D_c f)(x - c)}{x - c}\right) \\ &= (D_{f(c)}g)(0) \\ &= 0 \end{aligned} \tag{20.26}$$

[46]The present proof follows the proof in [7] rather closely.

by part (a) of the Algebraic Limit Theorem for Functional Limits. The first term will involve a rather intricate ϵ and δ argument. Fix $\epsilon > 0$. The goal is to show that there exists a $\delta > 0$ such that

$$\Big|g\big(f(x)\big) - g\big(f(c)\big) - (D_{f(c)}g)\big(f(x) - f(c)\big)\Big| < \epsilon|x - c| \qquad \forall\, x \in V_\delta(c). \tag{20.27}$$

Because f is differentiable at c, there exists a $\delta_f > 0$ such that[47]

$$\big|f(x) - f(c) - (D_c f)(x - c)\big| < |x - c| \qquad \forall\, x \in V_{\delta_f}(c). \tag{20.28}$$

Set

$$\epsilon_g := \frac{\epsilon}{1 + |f'(c)|}. \tag{20.29}$$

Because g is differentiable at $f(c)$, there exists a $\delta_g > 0$ such that

$$\Big|g(y) - g\big(f(c)\big) - (D_{f(c)}g)\big(y - f(c)\big)\Big| < \epsilon\big|y - f(c)\big| \qquad \forall\, y \in V_{\delta_g}\big(f(c)\big). \tag{20.30}$$

Since f is differentiable at c, f is also continuous at c by Theorem 20.16. Therefore, there exists a $\delta' > 0$ such that

$$\big|f(x) - f(c)\big| < \delta_g \qquad \forall\, x \in V_{\delta'}(c). \tag{20.31}$$

Set

$$\delta := \min\{\delta', \delta_f\}. \tag{20.32}$$

Putting all these inequalities together gives

$$\begin{aligned}
&\Big|g\big(f(x)\big) - g\big(f(c)\big) - (D_{f(c)}g)\big(f(x) - f(c)\big)\Big| \\
&\quad < \epsilon_g\big|f(x) - f(c)\big| \quad \text{by (20.30) and (20.31)} \\
&\quad = \epsilon_g\big|f(x) - f(c) - (D_c f)(x - c) + (D_c f)(x - c)\big| \\
&\quad \le \epsilon_g\Big(\big|f(x) - f(c) - (D_c f)(x - c)\big| + \big|(D_c f)(x - c)\big|\Big) \quad \text{by the triangle inequality} \\
&\quad \le \epsilon_g\Big(|x - c| + |f'(c)||x - c|\Big) \quad \text{by (20.28) Proposition 20.6} \\
&\quad = \epsilon|x - c| \quad \forall\, x \in V_\delta(c) \quad \text{by (20.29)},
\end{aligned} \tag{20.33}$$

which proves that

$$\lim_{x \to c} \frac{g\big(f(x)\big) - g\big(f(c)\big) - (D_{f(c)}g)\big(f(x) - f(c)\big)}{x - c} = 0. \tag{20.34}$$

Therefore, $\lim\limits_{x \to c} h(x) = 0$ and the chain rule is proved. ■

Remark 20.35. Equation (20.22) depicts a form of the Chain Rule that can be visualized diagrammatically as follows. The functions g, f, and their compositions form a diagram

$$\begin{tikzcd}
& B \arrow[dl, "g"'] & \\
\mathbb{R} & & A \arrow[ul, "f"'] \arrow[ll, "g \circ f"]
\end{tikzcd} \tag{20.36}$$

[47] The ϵ_f here was chosen to be 1.

and this diagram gets sent to the diagram

$$\begin{tikzcd} & \mathbb{R} \arrow[dl, "D_{f(c)}g"'] & \\ \mathbb{R} & & \mathbb{R} \arrow[ul, "D_c f"'] \arrow[ll, "D_c(g\circ f)"] \end{tikzcd} \tag{20.37}$$

upon applying the differential. We say that this latter diagram *commutes* because first performing the transformation $D_c f$ and then applying $D_{f(c)}g$ gives, by definition, $D_{f(c)}g \circ D_c f$, but the diagram indicates that this latter transformation must be equal to $D_c(g \circ f)$.

Theorem 20.38. *[Algebraic Differentiability Theorem] Let $A \subseteq \mathbb{R}$, let $f, g : A \to \mathbb{R}$, and let $c \in A$. Furthermore, suppose that f and g are differentiable at c with derivatives $f'(c)$ and $g'(c)$, respectively. Then the following facts hold.*

(a) The function kf is differentiable at c for all $k \in \mathbb{R}$ with derivative $(kf)'(c) = kf'(c)$.

(b) The function $f + g$ is differentiable at c with derivative $(f + g)'(c) = f'(c) + g'(c)$.

(c) The function fg is differentiable at c with derivative $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$.

(d) Let $B \subseteq A$ be the domain over which g is nonzero and so that $c \in B$. Then the function $\frac{f}{g}$ is differentiable at c with derivative $\left(\frac{f}{g}\right)'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{[g(c)]^2}$.

Proof. Exercise. ■

After this lecture, it is recommended the student works through problems 1 and 2 on HW #5.

21 November 8

Today we will present many of the great theorems of calculus. Rather than reproving the same theorems in Abbott, we will prove only those that were left as exercises. Abbott discusses an example of a function that is differentiable everywhere, but its derivative function is not continuous. However, we have not yet established the necessary results to talk about this function. We will be better prepared to do so after a few more lectures when we discuss series of functions and their derivatives. At that point, we will define the sin and cos functions and explore their derivatives. We will also be able to discuss the derivative of the exponential function.

Theorem 21.1 (Interior Extremum Theorem)**.** *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : (a, b) \to \mathbb{R}$ be a function that is differentiable on (a, b). If there exists a point $c \in (a, b)$ for which either $f(x) \le f(c)$ for all $x \in (a, b)$ or $f(x) \ge f(c)$ for all $x \in (a, b)$ then $f'(c) = 0$.*

Proof. See Abbott. ■

Lemma 21.2. *Let $a, b \in \mathbb{R}$ with $a < b$ and let $g : [a, b] \to \mathbb{R}$ be a function that is differentiable on $[a, b]$ satisfying*

$$g'(a) < 0 < g'(b). \tag{21.3}$$

Then there exist $x, y \in (a, b)$ satisfying

$$g(a) > g(x) \qquad \& \qquad g(y) < g(b). \tag{21.4}$$

Proof. Set $\epsilon := -g'(a)$. Since g is differentiable at a, there exists a $\delta > 0$ such that

$$-\epsilon < \frac{g(x) - g(a)}{x - a} - g'(a) < \epsilon \qquad \forall\, x \in [a, a + \delta). \tag{21.5}$$

Since $x - a > 0$, multiplying throughout, using the fact that $\epsilon = -g'(a)$, and rearranging gives

$$g(x) < g(a) \qquad \forall\, x \in [a, a + \delta). \tag{21.6}$$

This proves that there exists some $x \in (a, b)$ such that $g(x) < g(a)$. A similar argument shows that there exists a $y \in (a, b)$ such that $g(y) < g(b)$. ■

The proof of this theorem actually showed more than the statement provided. In fact, under the same hypotheses, there exist neighborhoods around the points a and b for which the value of g satisfies the inequality provided for *all* such points.

Theorem 21.7 (Darboux's Theorem)**.** *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a function that is differentiable on $[a, b]$. Then f' satisfies the intermediate value property.*

Proof. Let $\alpha \in \mathbb{R}$ satisfy $f'(a) < \alpha < f'(b)$ if $f'(a) < f'(b)$ or $f'(a) > \alpha > f'(b)$ if $f'(a) > f'(b)$. Without loss of generality, consider the first case.[48] Define $g : [a, b] \to \mathbb{R}$ by

$$[a, b] \ni x \mapsto g(x) := f(x) - \alpha x. \tag{21.8}$$

[48] The goal is to prove that there exists a $c \in (a, b)$ such that $f'(c) = \alpha$.

Then g is differentiable and, by the Algebraic Differentiability Theorem, $g'(x) = f'(x) - \alpha$ for all $x \in [a, b]$. By assumption, $g'(a) < 0 < g'(b)$. By the previous Lemma, there exist points $x, y \in (a, b)$ such that $g(a) > g(x)$ and $g(y) < g(b)$. By the Extreme Value Theorem, g attains its minimum on $[a, b]$ (because g is continuous). By the preceding facts, the minimum of g is not attained at a or b since it was shown there exist x and y satisfying $g(a) > g(x)$ and $g(y) < g(b)$. Hence, there exists a $c \in (a, b)$ for which $g(c) \leq g(z)$ for all $z \in [a, b]$. By the Interior Extremum Theorem, $g'(c) = 0$. By definition of g, this means $f'(c) = \alpha$. Thus f satisfies the intermediate value property. ■

This result makes it seem even less plausible for the existence of a function that is differentiable yet its derivative function is discontinuous.

Theorem 21.9 (Rolle's Theorem)**.** *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a function satisfying*

i) f is continuous on $[a, b]$,

ii) f is differentiable on (a, b), and

iii) $f(a) = f(b)$.

Then there exists a point $c \in (a, b)$ with $f'(c) = 0$.

Proof. See Abbott. ■

A slight generalization of Rolle's theorem is the Mean Value Theorem.

Theorem 21.10. *[Mean Value Theorem] Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a function satisfying*

i) f is continuous on $[a, b]$ and

ii) f is differentiable on (a, b).

Then there exists a point $c \in (a, b)$ with

$$f'(c) = \frac{f(b) - f(a)}{b - a}. \tag{21.11}$$

Proof. See Abbott. ■

Rolle's Theorem is a special case of the Mean Value Theorem obtained by setting $f(a) = f(b)$.

Theorem 21.12 (Generalized Mean Value Theorem)**.** *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f, g : [a, b] \to \mathbb{R}$ be two functions satisfying*

i) f is continuous on $[a, b]$ and

ii) f is differentiable on (a, b).

Then there exists a point $c \in (a,b)$ with

$$\big(f(b) - f(a)\big)g'(c) = \big(g(b) - g(a)\big)f'(c). \tag{21.13}$$

Furthermore, if, in addition, g' is never zero on (a,b), then

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}. \tag{21.14}$$

Proof. Define $h : [a,b] \to \mathbb{R}$ to be the function

$$h := \big[f(b) - f(a)\big]g - \big[g(b) - g(a)\big]f. \tag{21.15}$$

Then, by the Algebraic Differentiability Theorem, h is continuous on $[a,b]$ and differentiable on (a,b). Furthermore,

$$h(a) = f(b)g(a) - g(b)f(a) \qquad \& \qquad h(b) = g(a)f(b) - f(a)g(b), \tag{21.16}$$

showing that $h(a) = h(b)$. By Rolle's Theorem, there exists a $c \in (a,b)$ such that $h'(c) = 0$, i.e. there exists a $c \in (a,b)$ such that

$$\big[f(b) - f(a)\big]g'(c) - \big[g(b) - g(a)\big]f'(c) = 0. \tag{21.17}$$

This proves the claim. ■

The Mean Value Theorem is the special case of the Generalized Mean Value Theorem with $g(x) = x$ for all $x \in [a,b]$. To see this more clearly, notice that because $a < b$, the conclusion can be equivalently written as

$$\left(\frac{f(b) - f(a)}{b - a}\right)g'(c) = \left(\frac{g(b) - g(a)}{b - a}\right)f'(c). \tag{21.18}$$

Theorem 21.19 (L'Hospital's Rule: 0/0 case)**.** *Let $a < b$ and let $c \in [a,b]$. Let $f, g : [a,b] \to \mathbb{R}$ be continuous functions and suppose that the restrictions of f and g to $[a,c) \cup (c,b]$ are differentiable. Furthermore, suppose that $f(a) = g(a) = 0$ and $g'(x) \neq 0$ for all $x \in [a,c) \cup (c,b]$. Then*

$$\lim_{x \to c} \frac{f'(x)}{g'(x)} = L \qquad \Rightarrow \qquad \lim_{x \to c} \frac{f(x)}{g(x)} = L. \tag{21.20}$$

Notice the very crucial direction of implication in this theorem.

Proof. Exercise. ■

Theorem 21.21 (L'Hospital's Rule: ∞/∞ case)**.** *Let $a < b$ and let $c \in (a,b)$. Let $f, g : (a,c) \cup (c,b) \to \mathbb{R}$ be differentiable functions. Furthermore, suppose that $g'(x) \neq 0$ for all $x \in (a,c) \cup (c,b)$. If $\lim\limits_{x \to c} g(x) = \pm\infty$, then*

$$\lim_{x \to c} \frac{f'(x)}{g'(x)} = L \qquad \Rightarrow \qquad \lim_{x \to c} \frac{f(x)}{g(x)} = L. \tag{21.22}$$

Proof. See Abbott. ■

After this lecture, it is recommended the student works through problems 3 and 4 on HW #5.

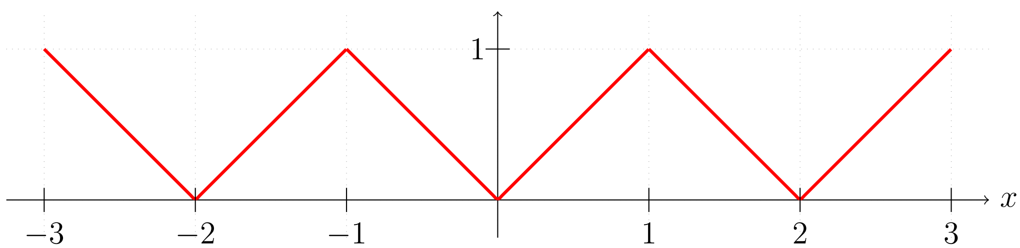
22 November 10

Earlier, we saw that a monotone function has an at most countable set of discontinuities (an arbitrary function can have an uncountable number of discontinuities, such as the Dirichlet function). We also mentioned that given a domain A and any at most countable subset $C \subseteq A$, there exists a monotone function whose set of discontinuities coincides with C. In analogy to the set of discontinuities of a function, we could ask what are the possible points at which a function is not differentiable? Such a concept is meaningless for functions that are discontinuous since it is necessary that a differentiable function is continuous. Thus, we could ask what the set of points at which a *continuous* function is not differentiable looks like. Can we construct a continuous function that is not differentiable on any arbitrary at most countable subset of its domain? Today we will construct an example of a continuous function that is not differentiable at *any* point in its domain.

Example 22.1. Let $h_0 : \mathbb{R} \to \mathbb{R}$ be the function defined as follows. For any $z \in \mathbb{Z}$, if $x \in [2z-1, 2z+1]$, set

$$h_0(x) := |x - 2z|, \tag{22.2}$$

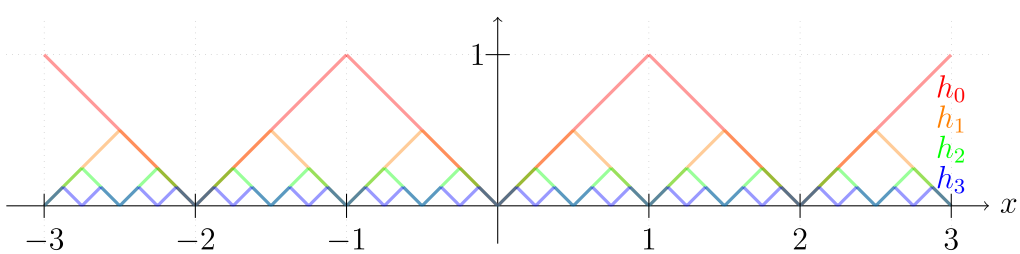
as depicted in the following graph.



For each $n \in \mathbb{N}$, set $h_n : \mathbb{R} \to \mathbb{R}$ to be the function defined by

$$\mathbb{R} \ni x \mapsto h_n(x) := \frac{1}{2^n} h_0(2^n x). \tag{22.3}$$

For $n = 1, 2, 3$, these functions are depicted in the graph below



Then, for each $x \in \mathbb{R}$, set

$$g(x) := \sum_{n=0}^{\infty} h_n(x). \tag{22.4}$$

This sum converges (absolutely) for each $x \in \mathbb{R}$ because

$$|h_n(x)| = \left| \frac{h_0(2^n x)}{2^n} \right| \leq \frac{1}{2^n} \tag{22.5}$$

since $h_0(x)$ is bounded by 1. Thus,

$$g(x) \leq \sum_{n=0}^{\infty} \frac{1}{2^n} = 2 \tag{22.6}$$

showing that the partial sums $\sum_{n=0}^{m} h_n(x)$ are bounded. Because this sequence of partial sums is monotone non-decreasing, by the monotone convergence theorem, it converges. However, notice that if we want to study the properties of the *function* $g : \mathbb{R} \to \mathbb{R}$ defined by

$$\mathbb{R} \ni x \mapsto g(x), \tag{22.7}$$

we would need to establish some results concerning infinite sums of functions. For example, is it true that $\sum_{n=0}^{\infty} h_n$ is a continuous function? Before we can answer any of these questions, we should talk about sequences of functions and what properties (such as continuity and differentiability) are preserved when taking an infinite sum of such functions. Therefore, we must put this example on hold and study these concepts.

Example 22.8. Incidentally, we will also be able to use these concepts to provide an example of a function that is differentiable everywhere but its derivative function is discontinuous. Such a function is given by the following

$$\mathbb{R} \ni x \mapsto \begin{cases} 0 & \text{if } x = 0 \\ \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+3}}{(2n+1)!} & \text{otherwise} \end{cases} \tag{22.9}$$

The derivative of this function is discontinuous at $x = 0$. However, it is not obvious that one can apply the derivative to each constituent in the infinite series and then sum said derivatives to obtain the derivative of the entire series.

Definition 22.10. Let $A \subseteq \mathbb{R}$ and let $\mathbb{R}^A$ denote the set of functions from A to $\mathbb{R}$,

$$\mathbb{R}^A := \{g : A \to \mathbb{R}\}. \tag{22.11}$$

A *sequence of functions*[49] from A to $\mathbb{R}$ is a function $f : \mathbb{N} \to \mathbb{R}^A$ whose value at $n \in \mathbb{N}$ is written as f_n. Often, a sequence f is also written as $(f_n)_{n\in\mathbb{N}}$ or just (f_n) for short. For every element $x \in A$, let $\mathrm{ev}_x : \mathbb{R}^A \to \mathbb{R}$ denote the *evaluation at x* function defined by

$$\mathbb{R}^A \ni g \mapsto \mathrm{ev}_x(g) := g(x). \tag{22.12}$$

A sequence $f : \mathbb{N} \to \mathbb{R}^A$ *converges pointwise* to a function $\lim f : A \to \mathbb{R}$, also denoted by $\lim_{n\to\infty} f_n$, iff for all $x \in A$,

$$(\lim f)(x) = \lim(\mathrm{ev}_x \circ f). \tag{22.13}$$

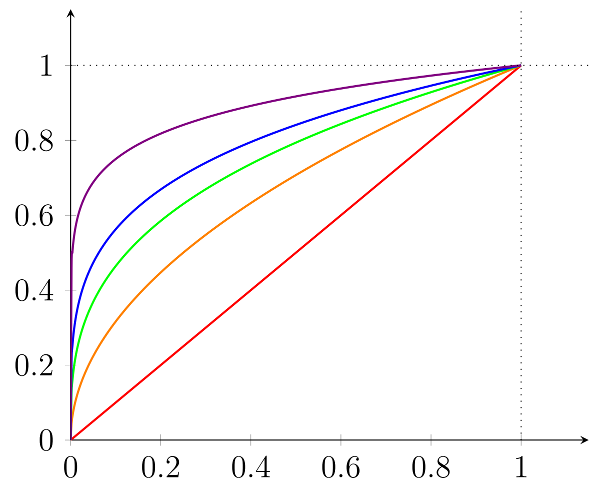
Note that the expression $\mathrm{ev}_x \circ f$ is a sequence (of real numbers) because it is the composition of functions $\mathbb{N} \xrightarrow{f} \mathbb{R}^A \xrightarrow{\mathrm{ev}_x} \mathbb{R}$. Hence, the limit on the right is the limit of a sequence of numbers and the limit on the left is a limit of a sequence of functions. The same notation is used because the two concepts are similar. In terms of the variable $n \in \mathbb{N}$, the equality (22.13) can be written as

$$\left(\lim_{n\to\infty} f_n\right)(x) = \lim_{n\to\infty} f_n(x). \tag{22.14}$$

[49]These functions need not be continuous.

Example 22.15. Let $A := [0, 1]$ and consider the sequence $f : \mathbb{N} \to \mathbb{R}^{[0,1]}$ of functions given by

$$[0, 1] \ni x \mapsto f_n(x) := \sqrt[n]{x}. \tag{22.16}$$



Then

$$\left(\lim_{n\to\infty} f_n\right)(x) = \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{otherwise} \end{cases} \tag{22.17}$$

Notice that in this example, the limit of continuous functions need *not* be continuous.

Going back to Definition 22.13, notice that when $A = \{*\}$ is a single point, then a sequence of functions is precisely a sequence of real numbers in the ordinary sense. To see this a bit more clearly, it is helpful to express the definition in terms of the definition of convergence of a sequence. Namely, a sequence f of functions on A converges pointwise to $\lim f : A \to \mathbb{R}$ iff for every $\epsilon > 0$ *and* $x \in A$, there exists an $N_{\epsilon,x} \in \mathbb{N}$ such that

$$\left|f_n(x) - (\lim f)(x)\right| < \epsilon \qquad \forall\, n \geq N_{\epsilon,x}. \tag{22.18}$$

The fact that A can have more than a single point allows one to ask for an N that depends on ϵ alone or both ϵ and x. This is analogous to how continuity differs from uniform continuity.

Example 22.19. Consider the case where A is a finite set of points $A = \{x_1, \ldots, x_k\}$. Then a function from A to $\mathbb{R}$ is defined by a k-tuple of numbers. Hence, a sequence $f : \mathbb{N} \to \mathbb{R}^A$ is a *sequence* of k-tuples of numbers. Thus, studying sequences of functions on a finite set with k elements is equivalent to studying sequences of points in $\mathbb{R}^k$, a topic for next semester. It is a little unfortunate that in our textbook, we jump from studying ordinary real numbers and sequences of real numbers all the way to a complicated extreme of sequences of functions before we study sequences of finite tuples of numbers, something that should logically be in-between the two topics.

Example 22.20. As another special case, consider $A := \mathbb{N}$. Then the set $\mathbb{R}^{\mathbb{N}}$ is precisely the set of all sequences of real numbers.

Remark 22.21. In general, for any three sets X, Y, and Z, there is a (natural) bijection

$$Z^{X\times Y} \cong (Z^Y)^X. \tag{22.22}$$

After this lecture, it is recommended the student works through problem 1 on HW #6.

23 November 15

Definition 23.1. Let $A \subseteq \mathbb{R}$. A sequence $f : \mathbb{N} \to \mathbb{R}^A$ of functions on A *converges uniformly* on A to a function $\lim f : A \to \mathbb{R}$, also denoted by $\lim_{n\to\infty} f_n$, iff for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$\big|f_n(x) - (\lim f)(x)\big| < \epsilon \qquad \forall\, n \geq N \text{ and } \forall\, x \in A. \tag{23.2}$$

It would be more precise to write N_ϵ instead of N in the definition of uniform convergence since N_ϵ depends on ϵ. The following result explains why the same notation $\lim f$ is used for both pointwise and uniform convergence.

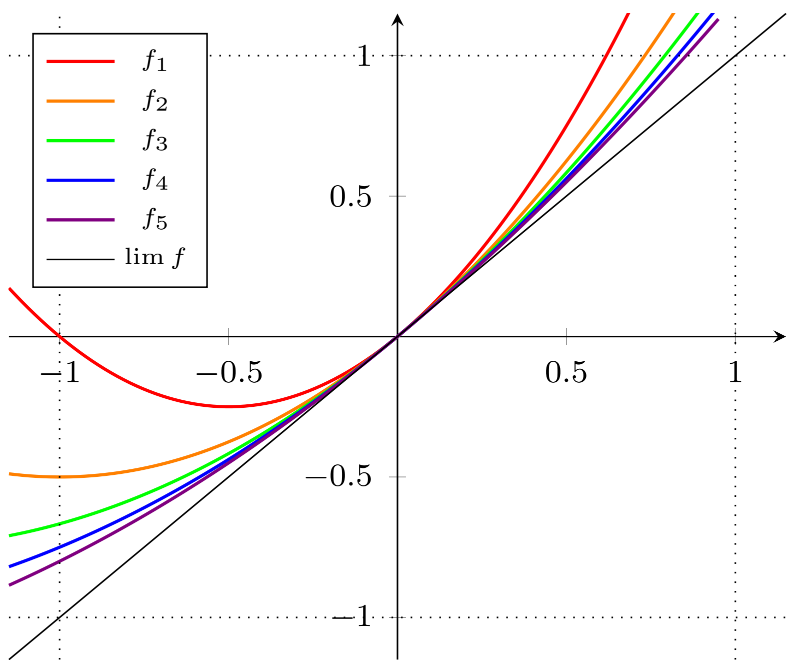
Theorem 23.3. *Let $A \subseteq \mathbb{R}$ and let $f : \mathbb{N} \to \mathbb{R}^A$ be a sequence of functions converging uniformly to a function $\lim f \in \mathbb{R}^A$. Then f converges pointwise to $\lim f$.*

Proof. Exercise. ∎

The examples of pointwise and uniform convergence from Abbott's book are quite nice, so we will repeat one of them here.

Example 23.4. Let $f : \mathbb{N} \to \mathbb{R}^{\mathbb{R}}$ be the sequence of functions defined by[50]

$$\mathbb{N} \times \mathbb{R} \ni (n, x) \mapsto f_n(x) := \frac{x^2 + nx}{n}. \tag{23.5}$$



By the Algebraic limit theorem, for each $x \in \mathbb{R}$,

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{x^2 + nx}{n} = \lim_{n\to\infty} \left(\frac{x^2}{n} + x\right) = \lim_{n\to\infty} \left(\frac{x^2}{n}\right) + \lim_{n\to\infty} (x) = x. \tag{23.6}$$

[50]Take a moment to convince yourself that a function $f : \mathbb{N} \to \mathbb{R}^A$ is equivalent to a function, denoted by the same letter, $f : \mathbb{N} \times A \to \mathbb{R}$. Also, see Remark 22.21.

This shows that f converges to $\mathrm{id}_{\mathbb{R}}$ pointwise, written either as $\lim f \underset{\text{pw}}{=} \mathrm{id}_{\mathbb{R}}$ or simply $\lim(\mathrm{ev}_x \circ f) = x$ for all $x \in \mathbb{R}$. However, the sequence f does *not* converge *uniformly* to $\mathrm{id}_{\mathbb{R}}$ on $\mathbb{R}$. To see this, fix $\epsilon > 0$ and fix $x \in \mathbb{R}$. Then

$$\big|f_n(x) - \mathrm{id}_R(x)\big| = \left|\frac{x^2 + nx}{n} - x\right| = \frac{x^2}{n}. \tag{23.7}$$

Set $N_{\epsilon,x} \in \mathbb{N}$ to be large enough so that

$$N_{\epsilon,x} > \frac{x^2}{\epsilon}. \tag{23.8}$$

Then, for this choice,

$$\big|f_n(x) - \mathrm{id}_R(x)\big| < \epsilon \qquad \forall\, n \geq N_{\epsilon,x}. \tag{23.9}$$

However, notice that as a function of x, $N_{\epsilon,x}$ is unbounded. Therefore, it will not be possible to find a single N_ϵ, independent of x for which $\big|f_n(x) - \mathrm{id}_R(x)\big| < \epsilon$ for all $n \geq N_\epsilon$ *and* for all $x \in \mathbb{R}$. However, if the domains of the functions changed to a compact domain, say K, so that f describes a sequence of functions $f : \mathbb{N} \to \mathbb{R}^K$, then by the Heine-Borel theorem, K is bounded so that there exists an M such that $K \subseteq [-M, M]$. Set $N_\epsilon \in \mathbb{N}$ to be an integer satisfying

$$N_\epsilon > \frac{M^2}{\epsilon}. \tag{23.10}$$

Then,

$$\big|f_n(x) - \mathrm{id}_R(x)\big| = \frac{x^2}{n} \leq \frac{M^2}{n} < \epsilon \qquad \forall\, n \geq N_\epsilon \text{ and } \forall\, x \in K. \tag{23.11}$$

This shows that for the sequence of functions $f : \mathbb{N} \to \mathbb{R}^K$ defined analogously to (23.5) but restricted to the domain K, $\lim f \underset{\text{un}}{=} \mathrm{id}_K$.

In your homework, you will prove a more general result about uniform convergence on compact sets (Dini's Theorem). However, it should be pointed out that it is *not* true that if $K \subseteq \mathbb{R}$ is compact and if $f : \mathbb{N} \to \mathbb{R}^K$ is a sequence that converges pointwise to some function $\lim f$, then it converges uniformly to $\lim f$.

Example 23.12. The sequence in Example 22.15 converges pointwise to (22.16). This convergence is not uniform because given $1 > \epsilon > 0$ and $x \in [0, 1]$, one needs to choose $N_{\epsilon,x} \in \mathbb{N}$ to satisfy

$$N_{\epsilon,x} > \frac{\log(x)}{\log(1-\epsilon)} \tag{23.13}$$

and $\lim\limits_{x \to 0^+} N_{\epsilon,x} = \infty$. This choice of $N_{\epsilon,x}$ illustrates that the sequence converges pointwise but not uniformly on $[0, 1]$.[51]

In fact, the assumptions can be strengthened even more. Consider the following example of a sequence of functions on a compact domain that consists of continuous functions that converge pointwise to a continuous function but not uniformly to a continuous function.

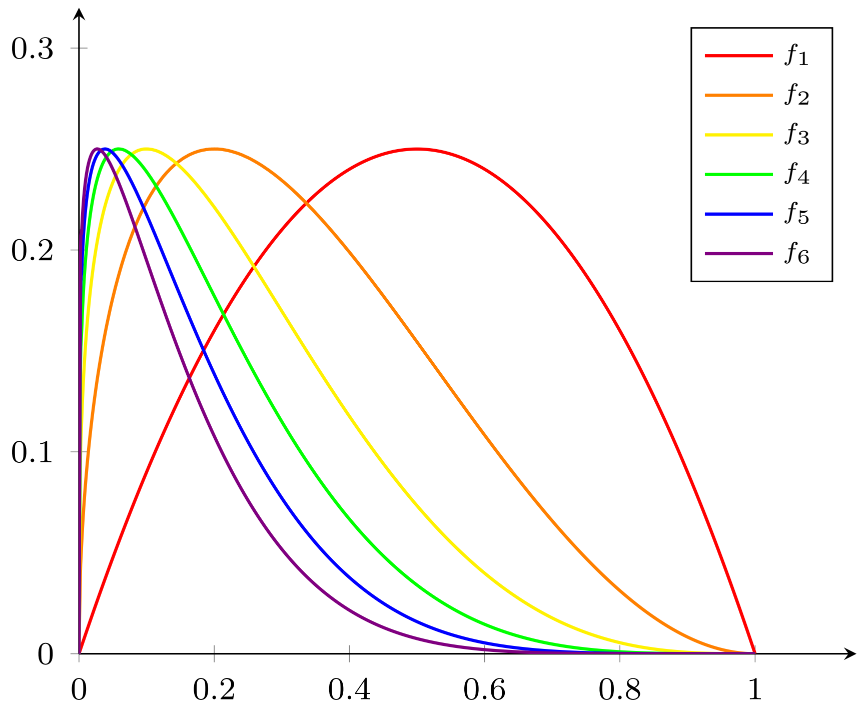
[51] Notice that this happened even though all of the f_n's were continuous on a compact domain and even increasing! Compare these assumptions to the assumptions in your homework problem.

Example 23.14. Let $c : \mathbb{N} \to \mathbb{R}$ be the sequence defined by

$$\mathbb{N} \ni n \mapsto c_n := \frac{1}{4}\left(\frac{n^2+1}{n^2}\right)^n \sqrt[n]{n^2+1} \tag{23.15}$$

and let $f : \mathbb{N} \to \mathbb{R}^{[0,1]}$ be the sequence

$$\mathbb{N} \times [0,1] \ni (n,x) \mapsto f_n(x) := c_n \sqrt[n]{x}(1-x)^n. \tag{23.16}$$



Exercise: show that f converges to the zero function pointwise but does not converge to this function uniformly. This example illustrates that the condition "the sequence f is increasing" in Dini's theorem cannot be dropped.

As for sequences of real numbers, we might want to know that a sequence converges even when we do not have a candidate limit in mind.

Definition 23.17. Let $A \subseteq \mathbb{R}$ and let $f : \mathbb{N} \to \mathbb{R}^A$ be a sequence of functions. f is a <u>*(uniform) Cauchy sequence*</u> iff for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$\big|f_n(x) - f_m(x)\big| < \epsilon \qquad \forall\, n, m \geq N \text{ and } \forall\, x \in A. \tag{23.18}$$

Theorem 23.19. *Let $A \subseteq \mathbb{R}$ and $f : \mathbb{N} \to \mathbb{R}^A$ a sequence of functions. f converges uniformly if and only if f is a uniform Cauchy sequence.*

Proof. You can try to prove this as an exercise, but we will also prove this theorem in much greater generality next semester. ■

Theorem 23.20. *[Continuous Limit Theorem] Let $A \subseteq \mathbb{R}$ and let $f : \mathbb{N} \to \mathbb{R}^A$ be a sequence of functions converging uniformly to a function $\lim f \in \mathbb{R}^A$. Let $c \in A$. If $f_n : A \to \mathbb{R}$ is continuous at c for every $n \in \mathbb{N}$, then $\lim f$ is continuous at c.*

Although this theorem is proved in Abbott, it is important and we present a proof here as well.

Proof. Fix $\epsilon > 0$. Because f uniformly converges to $\lim f$, there exists an $N \in \mathbb{N}$ such that

$$\big|(\lim f)(x) - f_n(x)\big| < \frac{\epsilon}{3} \qquad \forall\, n \geq N \text{ and } \forall\, x \in A. \tag{23.21}$$

For every $n \in \mathbb{N}$, f_n is continuous at c, so that there exists a $\delta_n > 0$ such that

$$\big|f_n(x) - f_n(c)\big| < \epsilon \qquad \forall\, x \in V_{\delta_n}(c) \cap A. \tag{23.22}$$

In particular, for $n = N$, set $\delta := \delta_N$. Using these facts,

$$\begin{aligned}
\big|(\lim f)(x) - (\lim f)(c)\big| &= \big|(\lim f)(x) - f_N(x) + f_N(x) - f_N(c) + f_N(c) - (\lim f)(c)\big| \\
&\leq \big|(\lim f)(x) - f_N(x)\big| + \big|f_N(x) - f_N(c)\big| + \big|f_N(c) - (\lim f)(c)\big| \\
&\qquad\qquad \text{by the triangle inequality} \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \qquad \text{by (23.21), (23.22), and (23.21), respectively} \\
&= \epsilon
\end{aligned} \tag{23.23}$$

for all $x \in V_\delta(c) \cap A$. Therefore, $\lim f$ is continuous at $c \in A$. ■

Example 23.24. We already provided an example where the theorem is false if the sequence f only converges pointwise to $\lim f$ (see Example 22.15).

Exercise 23.25. Prove or disprove: Let $A \subseteq \mathbb{R}$ and let $f : \mathbb{N} \to \mathbb{R}^A$ be a sequence of functions converging uniformly to a function $\lim f \in \mathbb{R}^A$. If $f_n : A \to \mathbb{R}$ is uniformly continuous for every $n \in \mathbb{N}$, then $\lim f$ uniformly continuous.

Exercise 23.26. In a similar vein to the previous exercise, see Exercise 6.2.6 in [2].

Remark 23.27. The Continuous Limit Theorem can be phrased in the following diagramatic way, but to do so, we need to introduce some notation. Let $C(A, \mathbb{R})$ be the set of continuous functions on A, let $(\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{pw}}$ denote the set of sequences of functions with a well-defined pointwise limit, and let $(\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{un}}$ denote the subset of $(\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{pw}}$ for which the limit converges to a function uniformly on A. Let

$$(C(A, \mathbb{R})^{\mathbb{N}})_{\mathrm{pw}} := C(A, \mathbb{R}) \cap (\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{pw}} \qquad \& \qquad (C(A, \mathbb{R})^{\mathbb{N}})_{\mathrm{un}} := C(A, \mathbb{R}) \cap (\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{un}} \tag{23.28}$$

be the set of sequences of continuous functions on A for which the limit is a (well-defined) function on A and for which the sequence converges uniformly, respectively. The pointwise limit of a sequence of functions on A can therefore be described as a function

$$(\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{pw}} \xrightarrow{\lim} \mathbb{R}^A. \tag{23.29}$$

This defines the function lim on all subsets of $(\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{pw}}$ including $(\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{un}}$, $(C(A, \mathbb{R})^{\mathbb{N}})_{\mathrm{pw}}$, and $(C(A, \mathbb{R})^{\mathbb{N}})_{\mathrm{un}}$. However, in each case, all we know is that the codomain of the function lim is always $\mathbb{R}^A$. The

Continuity Limit Theorem states that there exists a unique lift $\widetilde{\lim} : (C(A,\mathbb{R})^{\mathbb{N}})_{\text{un}} \to C(A,\mathbb{R})$, i.e. the diagram

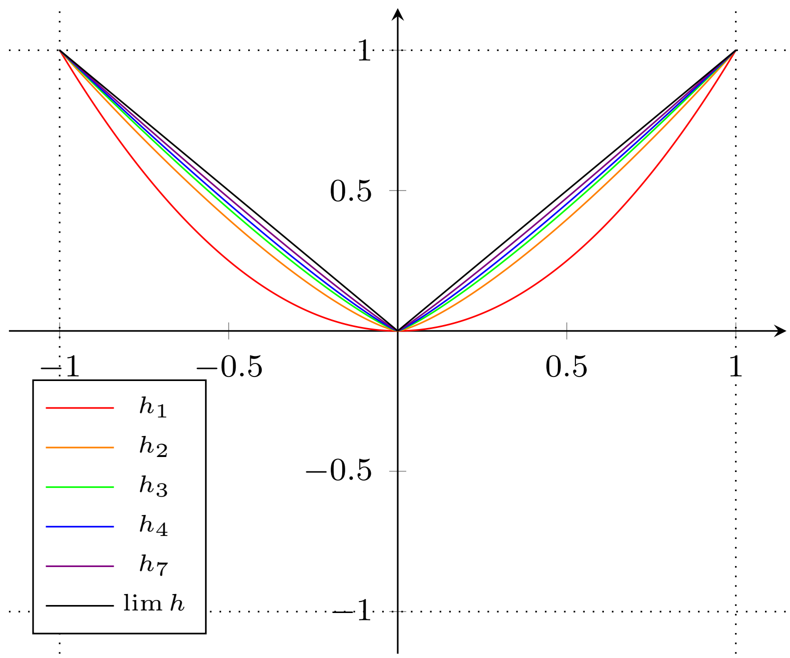
$$\begin{tikzcd} & C(A,\mathbb{R}) \arrow[dr, hook] & \\ (C(A,\mathbb{R})^{\mathbb{N}})_{\text{un}} \arrow[ur, dashed, "\widetilde{\lim}"] \arrow[rr, "\lim"'] & & \mathbb{R}^A \end{tikzcd} \tag{23.30}$$

commutes. It is common to say that the function lim <u>*factors through*</u> $C(A,\mathbb{R})$. We, however, abuse notation slightly, and denote the function $\widetilde{\lim}$ simply by $\lim$.

Let $A \subseteq \mathbb{R}$. Given a sequence $f : \mathbb{N} \to \mathbb{R}^A$ of *differentiable* functions on A, it is not always true that $\lim f$ is differentiable *even if* the sequence f converges to $\lim f$ uniformly (compare this to to the Continuous Limit Theorem above). The following example illustrates this.

Example 23.31. Let $h : \mathbb{N} \to \mathbb{R}^{[-1,1]}$ be the sequence of functions defined by

$$\mathbb{N} \times [-1,1] \ni (n,x) \mapsto h_n(x) := x^{1+\frac{1}{2n-1}}. \tag{23.32}$$



Then each of the h_n are differentiable, converge to $\lim h$ uniformly, but $\lim h$ is not differentiable at 0.

However, an additional assumption guarantees that the limiting function will be differentiable.

Theorem 23.33. *[Differentiable Limit Theorem] Let $a,b \in \mathbb{R}$ with $a < b$ and let $f : \mathbb{N} \to \mathbb{R}^{[a,b]}$ be a sequence of functions satisfying the conditions*

i) f converges to $\lim f$ pointwise,

ii) f_n is differentiable for all $n \in \mathbb{N}$, and

iii) the sequence $f' : \mathbb{N} \to \mathbb{R}^{[a,b]}$ (defined by $(f')_n := f'_n$) converges to a function $\lim f'$ uniformly.

Then $\lim f$ *is differentiable and* $(\lim f)' = \lim f'$.

Proof. Fix $c \in [a,b]$ and $\epsilon > 0$.[52] First, since f' converges to $\lim f'$ uniformly, it converges pointwise by Theorem 23.3. Hence, there exists an $N_1 \in \mathbb{N}$ such that

$$\left| f_n'(c) - (\lim f')(c) \right| < \frac{\epsilon}{3} \qquad \forall \, n \geq N_1. \tag{23.34}$$

Since f' converges to to $\lim f'$ uniformly, Theorem 23.19 asserts that there exists an $N_2 \in \mathbb{N}$ such that

$$\left| f_m'(x) - f_n'(x) \right| < \frac{\epsilon}{3} \qquad \forall \, n, m \geq N_2 \text{ and } \forall \, x \in [a,b]. \tag{23.35}$$

Set $N := \max\{N_1, N_2\}$. Since each f_n is differentiable at c, f_N is differentiable as well so that there exists a $\delta > 0$ such that

$$\left| \frac{f_N(x) - f_N(c)}{x - c} - f_N'(c) \right| < \frac{\epsilon}{3} \qquad \forall \, x \in V_\delta(c) \cap [a,b] \setminus \{c\}. \tag{23.36}$$

Now, notice that because the absolute value function is continuous and because f converges to $\lim f$ pointwise,

$$\lim_{m \to \infty} \left| \frac{f_m(x) - f_m(c)}{x - c} - \frac{f_n(x) - f_n(c)}{x - c} \right| = \left| \frac{(\lim f)(x) - (\lim f)(c)}{x - c} - \frac{f_n(x) - f_n(c)}{x - c} \right| \tag{23.37}$$

for all $x \in V_\delta(c) \cap [a,b] \setminus \{c\}$ and all $n \in \mathbb{N}$. However, before taking this limit, set

$$F := f_m - f_n, \tag{23.38}$$

and rewrite the left-hand-side as

$$\left| \frac{f_m(x) - f_m(c)}{x - c} - \frac{f_n(x) - f_n(c)}{x - c} \right| = \left| \frac{(f_m - f_n)(x) - (f_m - f_n)(c)}{x - c} \right| = \left| \frac{F(x) - F(c)}{x - c} \right| \tag{23.39}$$

for all $x \in V_\delta(c) \cap [a,b] \setminus \{c\}$ and all $m, n \in \mathbb{N}$. Note that $F : [a,b] \to \mathbb{R}$ is differentiable by the Algebraic Differentiability Theorem (Theorem 20.38). In particular, restricting F to $[x,c]$ or $[c,x]$ depending on whether $x < c$ or $c < x$, respectively, the Mean Value Theorem (Theorem 21.10) holds for F so that there exists an $\alpha \in (x,c)$ or $\alpha \in (c,x)$, respectively, satisfying

$$\frac{F(x) - F(c)}{x - c} = F'(\alpha). \tag{23.40}$$

Hence, (23.39) becomes

$$\left| \frac{F(x) - F(c)}{x - c} \right| = \left| F'(\alpha) \right| = \left| f_m'(\alpha) - f_n'(\alpha) \right| < \frac{\epsilon}{3} \qquad \forall \, n, m \geq N \tag{23.41}$$

[52]The goal is to show there exists a $\delta > 0$ such that

$$\left| \frac{(\lim f)(x) - (\lim f)(c)}{x - c} - (\lim f')(c) \right| < \epsilon$$

for all $x \in V_\delta(c) \cap [a,b] \setminus \{c\}$. The idea of the proof will actually be similar to the proof of the Continuity Limit Theorem.

by the inequality (23.35). Since such an α can be found for any pair of distinct elements x and c,

$$\left|\frac{f_m(x)-f_m(c)}{x-c}-\frac{f_n(x)-f_n(c)}{x-c}\right|<\frac{\epsilon}{3}\qquad\forall\,x\in V_\delta(c)\cap[a,b]\setminus\{c\}\text{ and }n\ge N.\tag{23.42}$$

Then, by the Order Limit Theorem (Theorem 6.15) and (23.37),

$$\left|\frac{(\lim f)(x)-(\lim f)(c)}{x-c}-\frac{f_n(x)-f_n(c)}{x-c}\right|\le\frac{\epsilon}{3}\qquad\forall\,x\in V_\delta(c)\cap[a,b]\setminus\{c\}\text{ and }n\ge N.\tag{23.43}$$

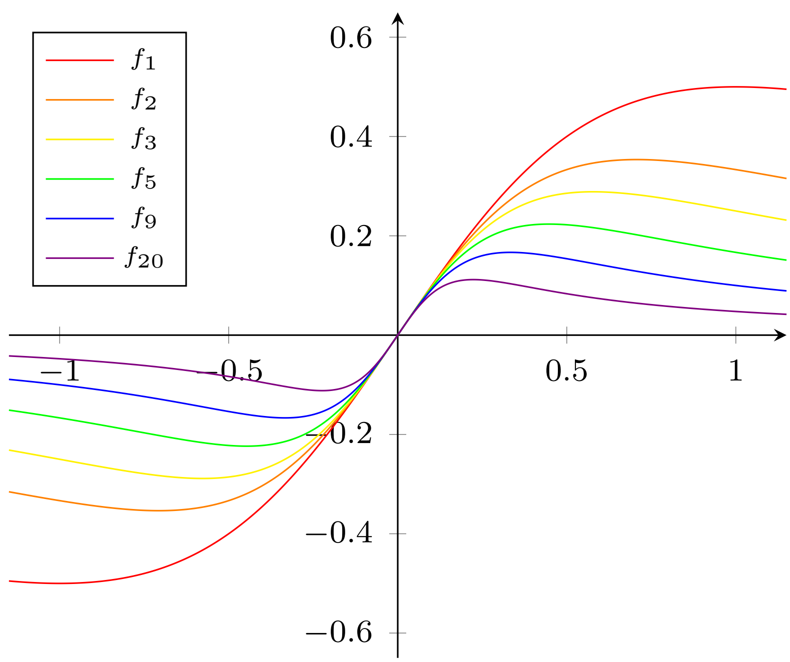
Using the triangle inequality and the inequalities (23.37), (23.36), and (23.34),

$$\begin{aligned}\left|\frac{(\lim f)(x)-(\lim f)(c)}{x-c}-(\lim f')(c)\right|&\le\left|\frac{(\lim f)(x)-(\lim f)(c)}{x-c}-\frac{f_N(x)-f_N(c)}{x-c}\right|\\&\quad+\left|\frac{f_N(x)-f_N(c)}{x-c}-f_N'(c)\right|+\left|f_N'(c)-(\lim f')(c)\right|\\&<\frac{\epsilon}{3}+\frac{\epsilon}{3}+\frac{\epsilon}{3}\\&=\epsilon\qquad\forall\,x\in V_\delta(c)\cap[a,b]\setminus\{c\}.\end{aligned}\tag{23.44}$$

Since this is true for all $c\in[a,b]$, it follows that $(\lim f)'=\lim f'$. ■

Example 23.45. Let $f:\mathbb{N}\to\mathbb{R}^{\mathbb{R}}$ be the sequence defined by

$$\mathbb{N}\times\mathbb{R}\ni(n,x)\mapsto f_n(x):=\frac{x}{1+nx^2}.\tag{23.46}$$

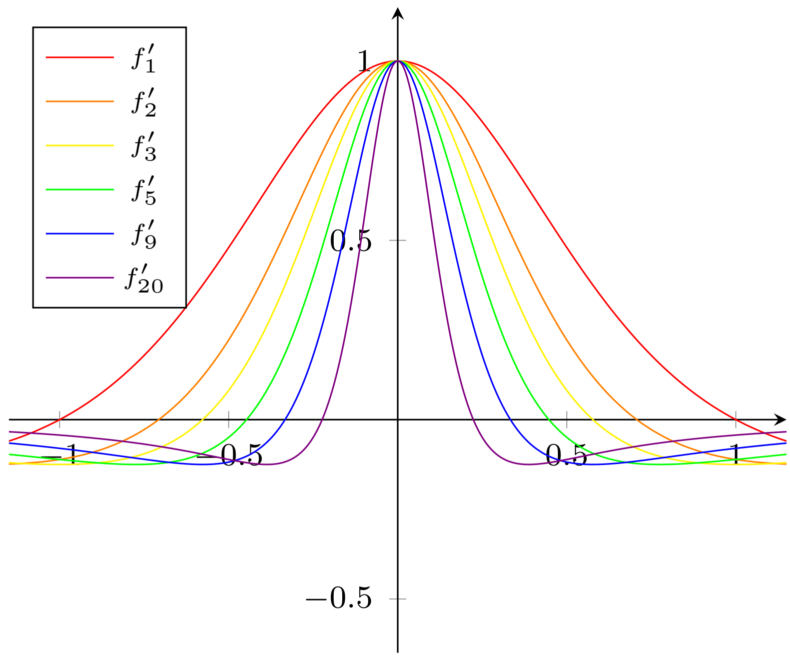


For each $x\in\mathbb{R}$,

$$\lim(\mathrm{ev}_x\circ f)=0.\tag{23.47}$$

Thus, $\lim f \underset{\text{pw}}{=} 0$. In fact, this convergence is uniform. To see this, first notice by the Algebraic Differentiability Theorem (the quotient rule in Theorem 20.38),

$$\mathbb{N}\times\mathbb{R}\ni(n,x)\mapsto f_n'(x)=\frac{1+nx^2-2nx^2}{(1+nx^2)^2}=\frac{1-nx^2}{(1+nx^2)^2}.\tag{23.48}$$



The zero's of this function provide sequences $y : \mathbb{N} \to \mathbb{R}$ and $z : \mathbb{N} \to \mathbb{R}$ that describe the minima and maxima of the sequence f. Namely, for each $n \in \mathbb{N}$, set

$$y_n := -\frac{1}{\sqrt{n}} \qquad \& \qquad z_n := \frac{1}{\sqrt{n}}. \tag{23.49}$$

Then y_n is the point at which f_n achieves its minimum and z_n is the point at which f_n achieves its maximum. Furthermore, at these points, the value of f_n is given by

$$f_n(y_n) = -\frac{1}{2\sqrt{n}} \qquad \& \qquad f_n(z_n) = \frac{1}{2\sqrt{n}} \tag{23.50}$$

so that

$$\big|f_n(x)\big| \le \frac{1}{2\sqrt{n}} \qquad \forall\, x \in \mathbb{R}. \tag{23.51}$$

Therefore, to see that $\lim f \underset{\text{un}}{=} 0$, fix $\epsilon > 0$. Set $N \in \mathbb{N}$ to be a natural number satisfying

$$N > \frac{1}{4\epsilon^2}. \tag{23.52}$$

Then, with this choice

$$\big|f_n(x)\big| \le \frac{1}{2\sqrt{n}} < \epsilon \qquad \forall\, n \ge N \text{ and } \forall\, x \in \mathbb{R}. \tag{23.53}$$

However, notice that

$$(\lim f')(x) \underset{\text{pw}}{=} \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases} \tag{23.54}$$

whereas

$$(\lim f)'(x) = 0 \qquad \forall\, x \in \mathbb{R}. \tag{23.55}$$

This does not contradict the Differentiable Limit Theorem (Theorem 23.33) because the sequence f' does *not* converge to the function in (23.54) uniformly.

Example 23.56. See Exercise 6.3.4 in [2] for another counter-example to this theorem when the assumption that f' converges to $\lim f'$ uniformly is dropped but it is assumed that f converges to $\lim f$ uniformly.

Remark 23.57. Just as the Continuity Limit Theorem can be described diagrammatically, so can the Differentiable Limit Theorem. Again, for this, we need some notation. Let $\mathrm{Diff}^1(A,\mathbb{R})$ denote the set of differentiable functions on A. Note that $\mathrm{Diff}^1(A,\mathbb{R}) \subseteq C(A,\mathbb{R})$. Let

$$(\mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}})_{\mathrm{pw}} := \mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}} \cap (\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{pw}} \quad \& \quad (\mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}})_{\mathrm{un}} := \mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}} \cap (\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{un}}. \tag{23.58}$$

In addition, let $D : \mathrm{Diff}^1(A,\mathbb{R}) \to \mathbb{R}^A$ denote the function that assigns the derivative to any differentiable function. Similarly, let $D^{\mathbb{N}} : \mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}} \to (\mathbb{R}^A)^{\mathbb{N}}$ denote the function that assigns the sequence of derivatives to any sequence of differentiable functions. Finally, let

$$\begin{aligned}
(\mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}})_{\mathrm{pw}} &:= \mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}} \cap (\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{pw}} \\
(\mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}})_{\mathrm{un}} &:= \mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}} \cap (\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{un}} \\
(\mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}})_{D\text{-pw}} &:= \left\{ f \in \mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}} \ : \ D^{\mathbb{N}} f \in (\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{pw}} \right\} \\
(\mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}})_{D\text{-un}} &:= \left\{ f \in \mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}} \ : \ D^{\mathbb{N}} f \in (\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{un}} \right\}
\end{aligned} \tag{23.59}$$

and denote intersections of these appropriately with commas. So for example,

$$(\mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}})_{\mathrm{pw},D\text{-un}} = (\mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}})_{D\text{-un}} \cap (\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{pw}} \tag{23.60}$$

denotes the set of sequences of differentiable functions that converge pointwise and whose associated derivative sequences converge uniformly. The Differentiable Limit Theorem then says that there exists a unique lift $\widetilde{\lim} : (\mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}})_{\mathrm{pw},D\text{-un}} \to \mathrm{Diff}^1(A,\mathbb{R})$ such that the diagrams

$$\begin{tikzcd}
& \mathrm{Diff}^1(A,\mathbb{R}) \arrow[dr, hook] & \\
(\mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}})_{\mathrm{pw},D\text{-un}} \arrow[ur, dashed, "\widetilde{\lim}"] \arrow[rr, "\lim"'] & & \mathbb{R}^A
\end{tikzcd} \tag{23.61}$$

and

$$\begin{tikzcd}
(\mathrm{Diff}^1(A,\mathbb{R})^{\mathbb{N}})_{\mathrm{pw},D\text{-un}} \arrow[r, "\widetilde{\lim}"] \arrow[d, "D^{\mathbb{N}}"'] & \mathrm{Diff}^1(A,\mathbb{R}) \arrow[d, "D"] \\
(\mathbb{R}^A)^{\mathbb{N}}_{\mathrm{un}} \arrow[r, "\lim"'] & \mathbb{R}^A
\end{tikzcd} \tag{23.62}$$

both commute. It is common to say that the function lim *factors through* $\mathrm{Diff}^1(A,\mathbb{R})$. We, however, abuse notation slightly, and denote the function $\widetilde{\lim}$ simply by $\lim$.

After this lecture, it is recommended the student works through problems 2, 3, and 4 on HW #6.

24 November 17

Definition 24.1. Let $A \subseteq \mathbb{R}$ and let $f : \mathbb{N} \to \mathbb{R}^A$ be a sequence of functions. The <u>*partial sums*</u> of f is the sequence of functions $S : \mathbb{N} \to \mathbb{R}^A$ defined by

$$\mathbb{N} \ni n \mapsto \sum_{m=1}^{n} f_m. \tag{24.2}$$

Lemma 24.3. *Let $A \subseteq \mathbb{R}$ and let $f : \mathbb{N} \to \mathbb{R}^A$ be a sequence of functions.*

(a) If f_n is continuous for every $n \in \mathbb{N}$, then the partial sums are all continuous.

(b) If f_n is uniformly continuous on A for every $n \in \mathbb{N}$, then the partial sums are all uniformly continuous on A.

(c) If f_n is differentiable for every $n \in \mathbb{N}$, then the partial sums are all differentiable and

$$S_n' = \sum_{m=1}^{n} f_m' \tag{24.4}$$

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

Proof. Exercise. ■

Since the partial sums themselves are sequences, the results from last lecture immediately imply the following facts.

Theorem 24.5. *[Term-by-term Continuity Theorem] Let $A \subseteq \mathbb{R}$ and let $f : \mathbb{N} \to \mathbb{R}^A$ be a sequence of functions. If f_n is continuous for every $n \in \mathbb{N}$ and the partial sums, S, of f converges uniformly to $\lim S$, then $\lim S$ is continuous on A.*

Proof. This follows from Lemma 24.3 and the Continuous Limit Theorem (Theorem 23.20). ■

Theorem 24.6. *[Term-by-term Differentiability Theorem] Let $a, b \in \mathbb{R}$ with $a < b$. Set $A := [a, b]$. Let $f : \mathbb{N} \to \mathbb{R}^A$ be a sequence of differentiable functions on A and let S denote the associated sequence of partial sums of f. Suppose that f and S satisfy*

i) S converges to $\lim S$ pointwise,

ii) f_n is differentiable for all $n \in \mathbb{N}$, and

iii) the sequence $S' : \mathbb{N} \to \mathbb{R}^{[a,b]}$ (defined by $(S')_n := S_n'$) converges uniformly to a function $\lim S'$.

Then $\lim S$ is differentiable and $(\lim S)' = \lim S'$.

Proof. This follows from Lemma 24.3, the fact that

$$S_n' = \sum_{m=1}^{n} f_m' \tag{24.7}$$

by the Algebraic Differentiability Theorem (Theorem 20.38), and the Differentiable Limit Theorem (Theorem 23.33). ■

Theorem 24.8. *[Cauchy Criterion for Uniform Convergence of Series] Let $A \subseteq \mathbb{R}$ and let $f : \mathbb{N} \to \mathbb{R}^A$ be a sequence of functions on A. The series $\sum_{n=1}^{\infty} f_n$ converges uniformly on A if and only if for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that*

$$\big|f_{m+1}(x) + f_{m+2}(x) + f_{m+3}(x) + \cdots + f_n(x)\big| < \epsilon \qquad \forall\, n > m \geq N \text{ and } \forall\, x \in A. \tag{24.9}$$

Proof. This follows from Theorem 23.19 applied to the partial sums associated to f. ■

Theorem 24.10 (Weierstrass M-test)**.** *Let $A \subseteq \mathbb{R}$ and let $f : \mathbb{N} \to \mathbb{R}^A$ be a sequence of functions on A. Suppose that for each $n \in \mathbb{N}$, there exists an $M_n \in \mathbb{R}$ such that*

$$\big|f_n(x)\big| \leq M_n \qquad \forall\, x \in A. \tag{24.11}$$

If $\sum_{n=1}^{\infty} M_n$ converges, then $\sum_{n=1}^{\infty} f_n$ converges uniformly on A.

Proof. Notice that for each $x \in A$, the partial sums of the sequence of numbers $\mathrm{ev}_x \circ f$ satisfy

$$\mathbb{N} \ni N \mapsto \sum_{n=1}^{N} \big|f_n(x)\big| \leq \sum_{n=1}^{N} M_n. \tag{24.12}$$

By the Comparison Test, this shows that

$$\sum_{n=1}^{\infty} \big|f_n(x)\big| \leq \sum_{n=1}^{\infty} M_n \tag{24.13}$$

showing that $\sum_{n=1}^{\infty} f_n(x)$ converges (in fact, absolutely) for each $x \in A$. In particular, the sequence of partial sums is Cauchy. Hence, for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that

$$\big|f_{m+1}(x) + f_{m+2}(x) + f_{m+3}(x) + \cdots + f_n(x)\big| < \epsilon \qquad \forall\, n \geq m \geq N \text{ and } \forall\, x \in A. \tag{24.14}$$

Thus, by the Cauchy Criterion for Uniform Convergence of Series (Theorem 24.8), $\sum_{n=1}^{\infty} f_n$ converges uniformly on A. ■

Example 24.15. We can use the Weierstrass M-test to prove that the sequence of partial sums $g_n := \sum_{m=0}^{n-1} h_m$ from Example 22.1 converges uniformly to the function $g : [0,2] \to \mathbb{R}$ defined in that same example. As was shown, $|h_n(x)| \leq \frac{1}{2^n}$ for all $x \in [0,2]$. Since $\sum_{m=0}^{\infty} \frac{1}{2^m} = 2$, the Weirstrass M-test guarantees that $\sum_{m=0}^{\infty} h_m$ converges uniformly on $[0,2]$. Unfortunately, nothing yet can be said about the derivatives because h_n is not differentiable for any $n \in \mathbb{N}$.

Exercise 24.16. Let $R \in (0,\infty)$ and set $A := [-R,R]$. Prove that the functions $\exp, \sin, \cos : A \to \mathbb{R}$ defined on $x \in A$ by

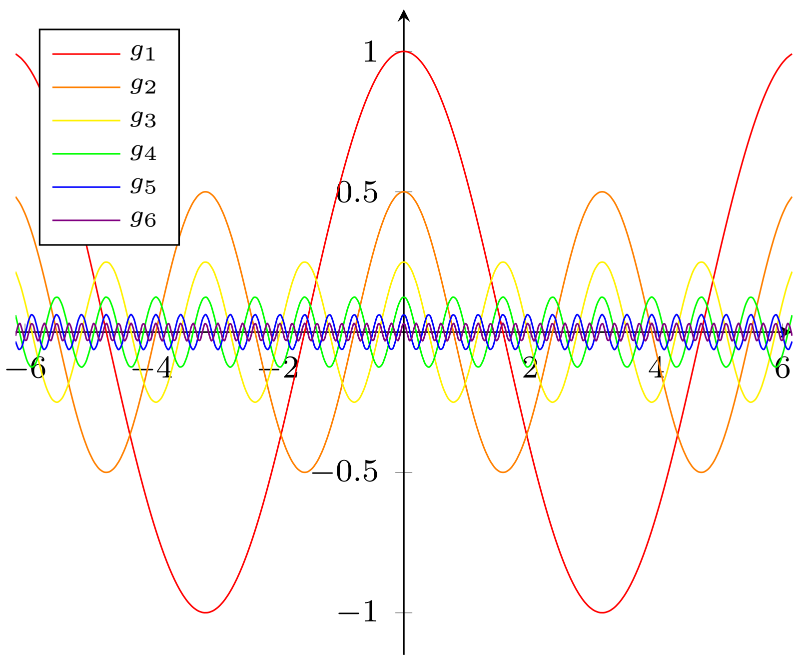
$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}, \qquad \sin(x) := \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}, \qquad \& \qquad \cos(x) := \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!} \tag{24.17}$$

are all continuous and differentiable on A. Furthermore, show that

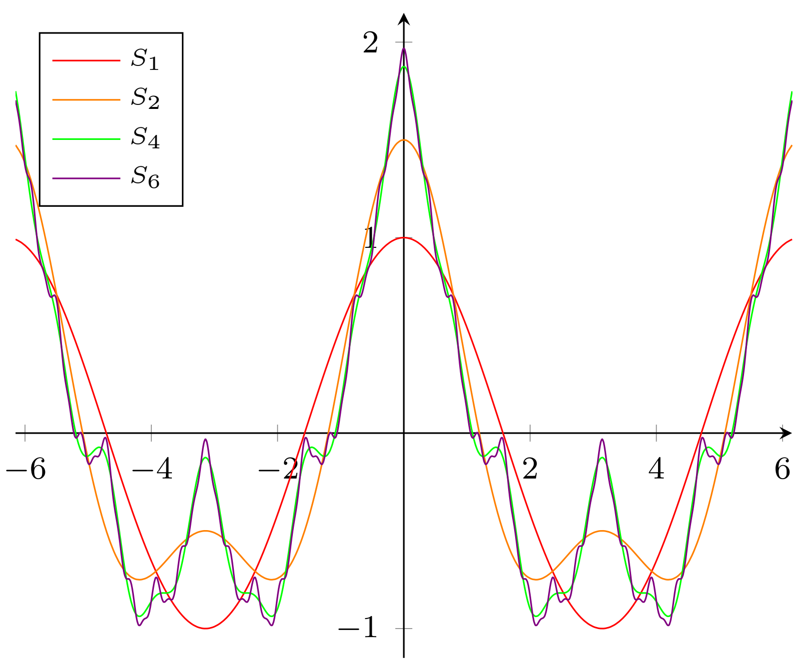
$$\exp' = \exp, \qquad \sin' = \cos, \qquad \& \qquad \cos' = -\sin. \tag{24.18}$$

Example 24.19. A better example of a sequence of functions whose associated partial sums converge to a function that is continuous but not differentiable is given by the following. Let $g : \mathbb{N} \to \mathbb{R}^{\mathbb{R}}$ be the sequence defined by

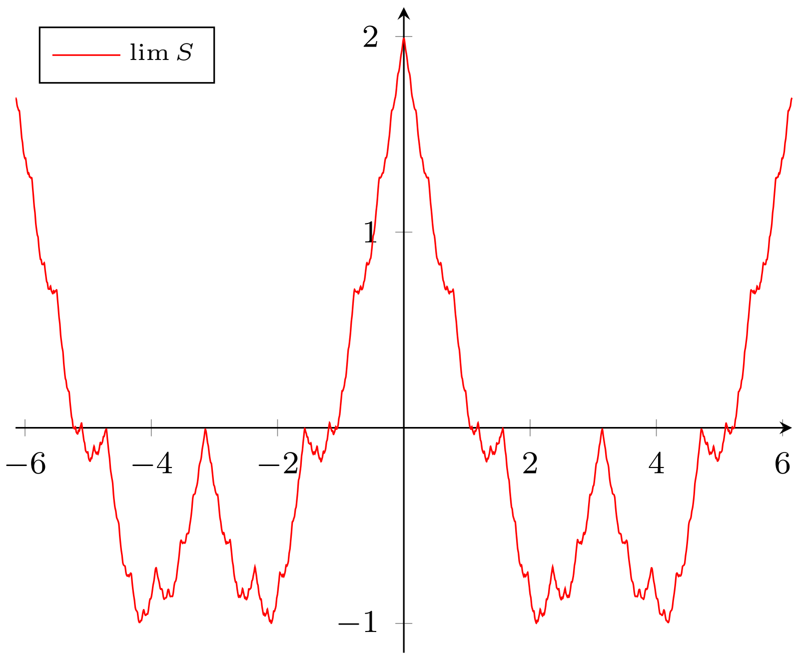
$$\mathbb{N} \times \mathbb{R} \ni (n,x) \mapsto g_n(x) := \frac{\cos(2^{n-1}x)}{2^{n-1}} \tag{24.20}$$



Some of the first few partial sums S associated to g are depicted by



with $\lim S$ looking something like



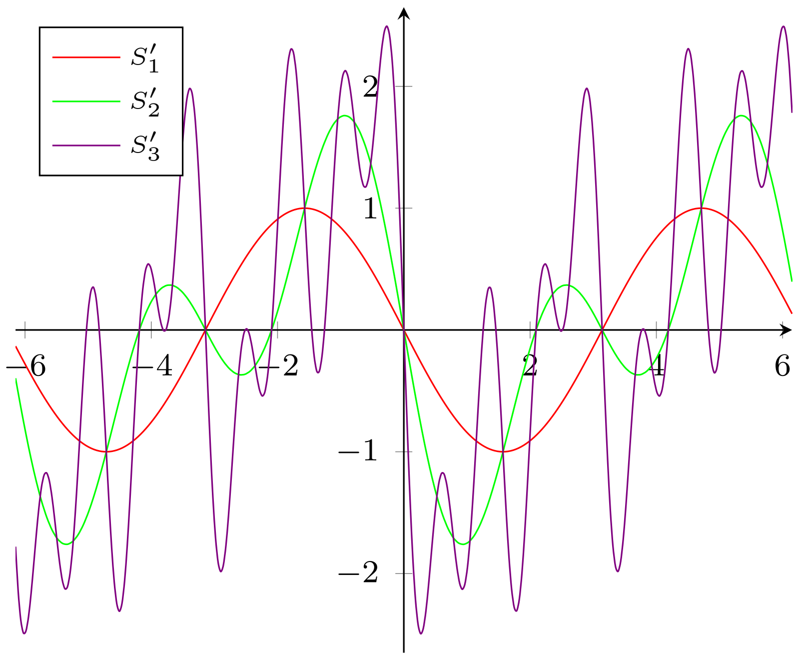
though of course this graph depicts only an approximation. Because cos is bounded by 1, the Weierstrass M-Test applies by setting $M_n := \frac{1}{2^n}$. Thus, S converges to $\lim S$ uniformly and $\lim S$ is continuous on $\mathbb{R}$. As for the differentiability (or lack thereof) of S and $\lim S$, notice that

$$g_n'(x) = -\sin(2^{n-1}x) \tag{24.21}$$

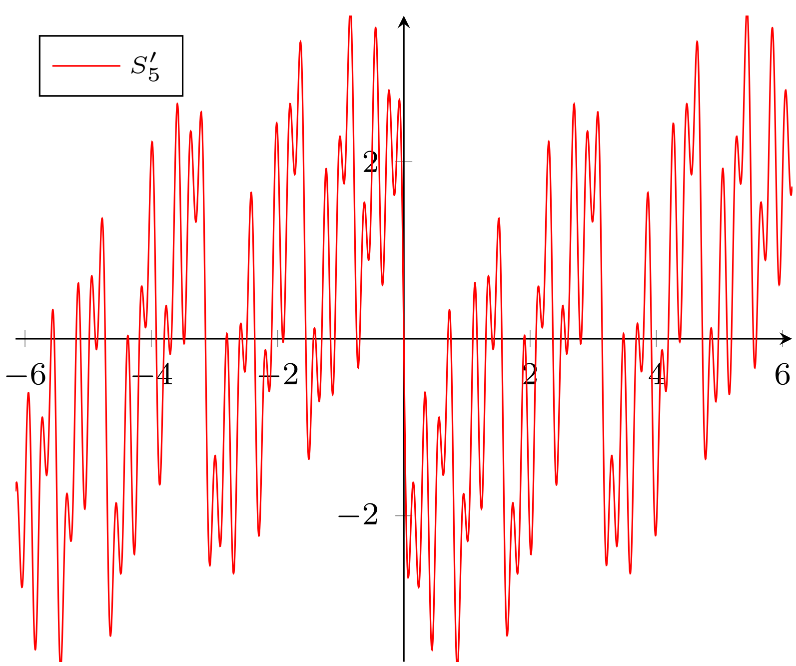
by the Chain rule and Exercise 24.16. Thus, by the Algebraic Differentiability Theorem,

$$S_n'(x) = -\sum_{m=1}^{n} \sin(2^{m-1}x). \tag{24.22}$$

There are three conditions that must be satisfied in Theorem 24.6. The one that fails for this sequence g and S is the third one. Namely, it is not true that S' converges to a function $\lim S'$ uniformly. This can be seen as follows. First, notice that S_n' is continuous for each $n \in \mathbb{N}$. If S' were to converge to $\lim S'$ uniformly, Theorem 24.5 would imply that $\lim S'$ (if it is even defined) is continuous. The first three such functions in the derivative sequence look like



while the fifth term already looks incredibly sporadic



Furthermore, Theorem 24.6 would require

$$(\lim S)'(x) = \lim S'(x) \tag{24.23}$$

for all $x \in \mathbb{R}$, which in particular, implies that $\lim S'(x)$ must be defined for all $x \in \mathbb{R}$. However, to see that the function $\lim S'$ is not differentiable at 0, consider the two sequences $y, z : \mathbb{N} \to \mathbb{R}$ given by sending $n \in \mathbb{N}$ to

$$y_n := -\frac{\pi}{2^n} \qquad \& \qquad z_n := \frac{\pi}{2^n}. \tag{24.24}$$

Both of these sequences satisfy $\lim y = 0 = \lim z$. Now, consider the associated sequences $Y := (\lim S') \circ y, Z := (\lim S') \circ z : \mathbb{N} \to \mathbb{R}$ given by

$$\mathbb{N} \ni n \mapsto Y_n := (\lim S')(y_n) = \lim_{N\to\infty} \sum_{m=1}^{N} \sin\left(2^{m-(n+1)}\pi\right) = \sum_{m=1}^{n} \sin\left(2^{m-(n+1)}\pi\right) \tag{24.25}$$

and

$$\mathbb{N} \ni n \mapsto Z_n := (\lim S')(z_n) = \lim_{N\to\infty} \left(- \sum_{m=1}^{N} \sin\left(2^{m-(n+1)}\pi\right)\right) = - \sum_{m=1}^{n} \sin\left(2^{m-(n+1)}\pi\right). \tag{24.26}$$

Notice that

$$Y_n > 0 > Z_n \qquad \forall\, n \in \mathbb{N} \tag{24.27}$$

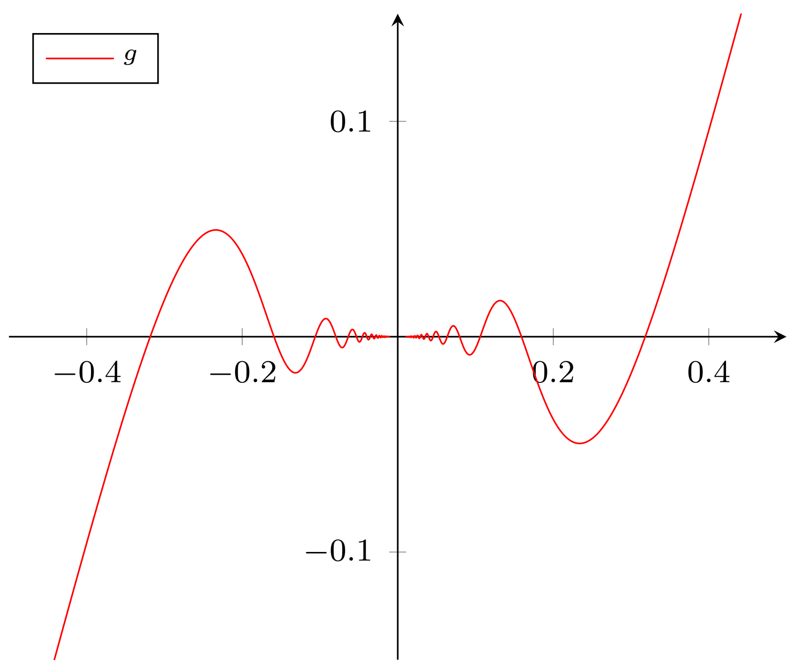
and Y is a strictly increasing sequence while Z is a strictly decreasing sequence. Therefore, we have exhibited two sequences y and z converging to 0 but satisfying

$$\lim\big((\lim S')\circ y\big) > 1 > 0 = (\lim S')(\lim y) = (\lim S')(\lim z) = 0 > -1 > \lim\big((\lim S')\circ z\big). \tag{24.28}$$

This violates the necessary condition that the function $\lim S'$ be continuous.

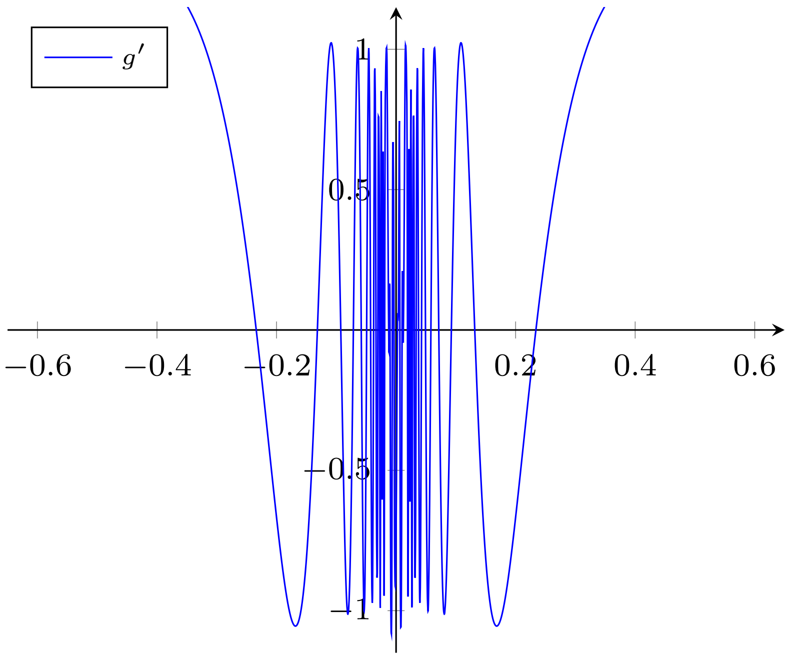
Example 24.29. The present example illustrates that not every differentiable function has a continuous derivative. We began this discussion in Example 22.8 with the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$\mathbb{R} \ni x \mapsto g(x) := \left. \begin{cases} 0 & \text{if } x = 0 \\ x^2 \sin\left(\frac{1}{x}\right) & \text{otherwise} \end{cases} \right\} = \begin{cases} 0 & \text{if } x = 0 \\ \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+3}}{(2n+1)!} & \text{otherwise} \end{cases} \tag{24.30}$$



By Theorem 24.6, the product rule, and Exercise 24.16,

$$g'(x) = \begin{cases} 0 & \text{if } x = 0 \\ 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right) & \text{otherwise} \end{cases} \tag{24.31}$$



The function g' is discontinuous at 0. One can see this by considering the two sequences $a, b : \mathbb{N} \to \mathbb{R}$ assigning $n \in \mathbb{N}$ to

$$a_n := \frac{1}{2n\pi} \qquad \& \qquad b_n := \frac{1}{(2n+1)\pi} \tag{24.32}$$

since $\lim a = 0 = \lim b$ but

$$\lim_{n\to\infty} g'(a_n) = -1 \neq 1 = \lim_{n\to\infty} g'(b_n). \tag{24.33}$$

Recall our earlier results on the different notions of continuity and differentiability for a function $f : [a, b] \to \mathbb{R}$. f differentiable $\Rightarrow$ f Lipschitz $\Rightarrow$ f uniformly continuous $\Rightarrow$ continuous. Today, we showed that there exist continuous functions that are differentiable *nowhere*. But what if we took a uniformly continuous function? Namely, does there exist a uniformly continuous function that is nowhere differentiable? What if we took a Lipschitz function? It turns out that if f is Lipschitz continuous, then f is differentiable *almost everywhere*. Here, the phrase "almost everywhere" has a precise mathematical meaning in terms of measure theory [7].

Definition 24.34. A subset $E \subseteq \mathbb{R}$ has <u>*measure zero*</u> if for any $\epsilon > 0$, there exists an at most countable cover $\{I_n = [a_n, b_n]\}_{n \in S \subseteq \mathbb{N}}$ of E consisting of closed intervals such that

$$\sum_{n \in S} (b_n - a_n) < \epsilon. \tag{24.35}$$

Example 24.36. Any at most countable subset of $\mathbb{R}$ has measure zero. In particular, $\mathbb{Q}$ has measure zero. The proof of this follows from the calculations in Example 10.9.

Theorem 24.37 (Rademacher's Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be a Lipschitz continuous. Then D_f, the set of discontinuities of f, has measure zero.*

Proof. Look elsewhere. ■

After this lecture, it is recommended the student works through problems 5, 6, and 7 on HW #6.

25 November 29

In Exercise 24.16, we introduced examples of what are known as power series. Unfortunately, I cannot find a precise enough definition of a power series in our textbook, so I have attempted to provide you with one below, first introducing some notation, and then providing a definition.

Definition 25.1. Let $p : \mathbb{N} \to \mathbb{R}^{\mathbb{R}}$ be the sequence of functions sending $n \in \mathbb{N} \cup \{0\}$ to the monomial $p_n : \mathbb{R} \to \mathbb{R}$ defined by[53]

$$\mathbb{R} \ni x \mapsto p_n(x) := x^n. \tag{25.2}$$

Let p_n also denote the restriction of p_n to A for any subset $A \subseteq \mathbb{R}$. Let $a : \mathbb{N} \cup \{0\} \to \mathbb{R}$ be a (shifted) sequence of real numbers. The <u>*power series*</u> associated to the sequence a consists of the subset $A \subseteq \mathbb{R}$ defined by

$$A := \left\{ x \in \mathbb{R} \ : \ \left| \sum_{n=0}^{\infty} a_n x^n \right| < \infty \right\} \tag{25.3}$$

together with the function $a \cdot p : A \to \mathbb{R}$ defined by

$$A \ni x \mapsto (a \cdot p)(x) := \sum_{n=0}^{\infty} a_n x^n. \tag{25.4}$$

An immediate question arises from this definition. Given two distinct (shifted) sequences a and b, is it possible for the associated power series to be the same? Before we answer this question, let us first analyze some facts about the domain A associated to the power series associated to a fixed sequence.

Theorem 25.5. *Let $a : \mathbb{N} \cup \{0\} \to \mathbb{R}$ be a (shifted) sequence and let A be the domain of the associated power series. Let $x_0 \in A$. Then*

$$\sum_{n=0}^{\infty} a_n x^n \tag{25.6}$$

converges absolutely for all $x \in \mathbb{R}$ satisfying $|x| < |x_0|$. In particular, $\big(-|x_0|, |x_0| \big) \subseteq A$.

There are two surprising statements being made here. The first statement says that if $x_0 \in A$, then *the entire open interval* $\big(-|x_0|, |x_0|\big)$ is *also* in A. The second statement is that for points x in *this* interval $\big(-|x_0|, |x_0|\big)$, the series not only converges, but it converges absolutely! Although this theorem is proved in Abbott, we'll present the proof here because I find this result quite surprising.

Proof. First note that if $x_0 = 0$, the claim is vacuously true since the interval $\big(-|x_0|, |x_0| \big)$ is empty. Hence, consider the case $x_0 \neq 0$. The sequence $\mathbb{N} \cup \{0\} \ni n \mapsto a_n x_0^n$ is bounded since the series converges at x_0. Let $M > 0$ denote such a bound, i.e. $|a_n x_0^n| \leq M$ for all $n \in \mathbb{N} \cup \{0\}$. Let $x \in \big(-|x_0|, |x_0| \big)$. Then,

$$|a_n x^n| = |a_n x_0^n| \left| \frac{x}{x_0} \right|^n \leq M \left| \frac{x}{x_0} \right|^n. \tag{25.7}$$

[53] I contemplated using the notation $\sqrt[1/n]{\ }$ instead of p_n as an offered suggestion by Ralf Schiffler. I might do this in the future, but for now, let's use p_n.

Hence,

$$\sum_{n=0}^{\infty}|a_nx^n| \le \sum_{n=0}^{\infty} M\left|\frac{x}{x_0}\right|^n = M\sum_{n=0}^{\infty}\left|\frac{x}{x_0}\right|^n = M\left(\frac{\left|\frac{x}{x_0}\right|}{1-\left|\frac{x}{x_0}\right|}\right) < \infty. \tag{25.8}$$

Therefore, $\sum_{n=0}^{\infty} a_nx^n$ converges absolutely for all $x \in \big(-|x_0|, |x_0|\big)$. ■

Definition 25.9. Let $a : \mathbb{N}\cup\{0\} \to \mathbb{R}$ be a (shifted) sequence and let A denote the domain of the associated power series $a \cdot p$. The radius of convergence of the power series associated to a is either a real number $R \in \mathbb{R}$ satisfying the two conditions

(a) $(a \cdot p)(x)$ converges absolutely for all $x \in (-R, R)$ and

(b) for all $\epsilon > 0$, the interval $(-R-\epsilon, R+\epsilon)$ is *not* contained in A

or is infinite.

Hence, the previous theorem says that if a nonzero number R is in the domain of a power series, then that domain automatically contains an interval of the form $(-R, R)$, where R is the radius of convergence of that power series.

Exercise 25.10. Let $a, b : \mathbb{N} \to \mathbb{R}$ be two sequences with associated domains with a common nonzero radius of convergence $R > 0$. Then, $a = b$ if and only if the associated power series are equal on $(-R, R)$, i.e. if and only if

$$\sum_{n=0}^{\infty} a_nx^n = \sum_{n=0}^{\infty} b_nx^n \qquad \forall\, x \in (-R, R). \tag{25.11}$$

There are several results that describe whether the domain extends beyond the radius of convergence if the radius is finite and what type of convergence holds on these domains. However, the only possibilities of such an extension are to include the limit points of these intervals.

Theorem 25.12. *Let $a : \mathbb{N}\cup\{0\} \to \mathbb{R}$ be a sequence and let $A \subseteq \mathbb{R}$ be the domain of the power series associated to a. Let $x_0 \in A$ be a point such that $(a \cdot p)(x_0)$ converges absolutely. Then the partial sums of $a \cdot p$ converges to $a \cdot p$ uniformly on $\big[-|x_0|, |x_0|\big]$. In particular, $a \cdot p$ is continuous on $\big[-|x_0|, |x_0|\big]$.*

Proof. Let $x \in \big[-|x_0|, |x_0|\big]$ and set $M_n := |a_nx_0^n|$ for each $n \in \mathbb{N}\cup\{0\}$. Then

$$|a_nx^n| \le M_n \qquad \forall\, n \in \mathbb{N}\cup\{0\} \tag{25.13}$$

and

$$\sum_{n=0}^{\infty} M_n < \infty \tag{25.14}$$

by the assumption that $\sum_{n=0}^{\infty} a_nx_0^n$ converges absolutely. Hence, by the Weierstrass M-test, the partial sums of $a \cdot p$ converges uniformly to the function $a \cdot p$ on $\big[-|x_0|, |x_0|\big]$. Continuity of $a \cdot p$ on $\big[-|x_0|, |x_0|\big]$ follows from the Term-by-term Continuity Theorem because the partial sums of $a \cdot p$ converge uniformly to $a \cdot p$ on $\big[-|x_0|, |x_0|\big]$ and the partial sums are continuous since they are all polynomials. ■

Referring to Theorem 25.12, one can remove the condition of *absolute* convergence of a power series at some point but the domain on which uniform convergence holds might not extend in both directions.

Theorem 25.15 (Abel's Theorem)**.** *Let $A \subseteq \mathbb{R}$ and let $a : \mathbb{N} \cup \{0\} \to \mathbb{R}$ and let $x_0 \in A$ be positive and such that the series $(a \cdot p)(x_0)$ converges (not necessarily absolutely). Then the partial sums of $a \cdot p$ converges uniformly to $a \cdot p$ on $\big[0, x_0\big]$. In particular, $a \cdot p$ is continuous on $\big[0, x_0\big]$. Similarly, if x_0 is negative, and $(a \cdot p)(x_0)$ converges, then the partial sums of $a \cdot p$ converges uniformly to $a \cdot p$ on $\big[x_0, 0\big]$ and $a \cdot p$ is continuous on $\big[x_0, 0\big]$.*

Proof. See Abbott. ■

Example 25.16. Consider the (shifted) sequence

$$\mathbb{N} \cup \{0\} \ni n \mapsto a_n := \begin{cases} 0 & \text{if } n = 0 \\ \frac{1}{n} & \text{otherwise} \end{cases} \tag{25.17}$$

The domain of the associated power series is $[-1, 1)$. Notice that 1 is not contained in this domain even though -1 is. This is because the series $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ does not converge absolutely.

Theorem 25.18. *Let $a : \mathbb{N} \cup \{0\} \to \mathbb{R}$ be a sequence and let A be the domain of the associated power series. For any compact set $K \subseteq A$, the partial sums of $a \cdot p$ converge uniformly to $a \cdot p$ on K. In particular, $a \cdot p$ is continuous on all compact subsets of A.*

Proof. See Abbott. ■

The above statements are regarding the nature of continuity of power series. Below is a result on differentiability of power series.

Theorem 25.19. *Let $a : \mathbb{N} \cup \{0\} \to \mathbb{R}$ be a sequence and let A be the domain of the associated power series. Then $a \cdot p$ is differentiable on any open interval $(-R, R) \subseteq A$ and the derivative is given by*

$$(a \cdot p)'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \qquad \forall\, x \in (-R, R). \tag{25.20}$$

Furthermore, $a \cdot p$ is infinitely differentiable on $(-R, R)$ and

$$(a \cdot p)^{(n)}(x) = \sum_{m=n}^{\infty} \frac{m!}{(m-n-1)!} a_m x^{m-n} \qquad \forall\, x \in (-R, R), \tag{25.21}$$

where $(a \cdot p)^{(n)}$ denotes the n-th derivative of $a \cdot p$.

Proof. See Abbott. ■

Power series have a surprising relation among the coefficients defining them.

Theorem 25.22 (Taylor's formula)**.** *Let $a : \mathbb{N} \cup \{0\} \to \mathbb{R}$ be a sequence with associated power series $a \cdot p$ defined on some interval (possibly infinite or semi-infinite) $A \subseteq \mathbb{R}$ containing 0. Then*

$$a_n = \frac{(a \cdot p)^{(n)}(0)}{n!}. \tag{25.23}$$

Proof. By Theorem 25.19,

$$\begin{aligned}
\frac{(a\cdot p)^{(n)}(0)}{n!} &= \frac{1}{n!}\left(\sum_{m=n}^{\infty}\frac{m!}{(m-n-1)!}a_m p_{m-n}\right)(0) \\
&= \frac{1}{n!}\sum_{m=n}^{\infty}\frac{m!}{(m-n-1)!}a_m p_{m-n}(0) \\
&= \frac{1}{n!}\sum_{m=n}^{\infty}\frac{m!}{(m-n-1)!}a_m \delta_{mn} \\
&= \frac{1}{n!}\frac{n!}{(n-n-1)!}a_n \\
&= a_n,
\end{aligned} \tag{25.24}$$

where δ is the Kronecker-delta function

$$(\mathbb{N}\cup\{0\})\times(\mathbb{N}\cup\{0\}) \ni (m,n) \mapsto \delta_{mn} := \begin{cases} 1 & \text{if } m=n \\ 0 & \text{otherwise} \end{cases} \tag{25.25}$$

■

This theorem motivates the following definition.

Definition 25.26. Let $A \subseteq \mathbb{R}$ be a subset (containing some interval centered at 0) and let $f : A \to \mathbb{R}$ be a function that is infinitely differentiable. The *Taylor series* of f centered at 0 is the power series associated to the sequence, called *Taylor coefficients*,

$$\mathbb{N}\cup\{0\} \ni n \mapsto \frac{f^{(n)}(0)}{n!} \tag{25.27}$$

where $f^{(n)}$ denotes the n-th derivative of f ($f^{(0)} := f$).

The previous theorem begs the question whether or not arbitrary infinitely differentiable functions $f : A \to \mathbb{R}$ can be represented by their Taylor series. If true, this would be somewhat surprising. The reason is because to specify a function requires a specification of $|A|$ many numbers (here $|A|$ is the cardinality of A). However, as we've learned, continuous functions have some constraints and so it might seem that less data captures their information. Infinitely differentiable functions are even more restricted and so one might ask if their information is completely captured by their Taylor coefficients. The answer to this suspicion is "no" as the following example illustrates.

Example 25.28. Consider the function

$$\mathbb{R} \ni x \mapsto f(x) := \begin{cases} 0 & \text{if } x=0 \\ e^{-1/x^2} & \text{otherwise} \end{cases} \tag{25.29}$$

Then the Taylor coefficients of f at 0 all vanish, $f^{(n)}(0) = 0$ for all $n \in \mathbb{N}\cup\{0\}$. However, the function f is not identically zero on any open interval around 0 showing that the Taylor series does not converge to f in any such neighborhood.

Nevertheless, the following theorem provides an error bound on how well a sufficiently differentiable function is approximated by its associated Taylor series.

Theorem 25.30 (Lagrange's Remainder Theorem)**.** *Let $R > 0$ and let $f : (-R, R) \to \mathbb{R}$ be a function that is differentiable $N + 1$ times with $N \in \mathbb{N}$. Set*

$$\{0, 1, \ldots, N\} \ni n \mapsto a_n := \frac{f^{(n)}(0)}{n!}, \tag{25.31}$$

define the function $S_N : (-R, R) \to \mathbb{R}$ by

$$(-R, R) \ni x \mapsto S_N(x) := \sum_{n=0}^{N} a_n x^n, \tag{25.32}$$

and the function $E_N : (-R, R) \to \mathbb{R}$ by

$$E_N := f - S_N. \tag{25.33}$$

Then, for any $x \in (-R, R) \setminus \{0\}$, there exists a $c \in (-|x|, |x|)$ such that

$$E_N(x) = \frac{f^{(N+1)}(c)}{(N+1)!} x^{N+1} \tag{25.34}$$

for that particular value of x.

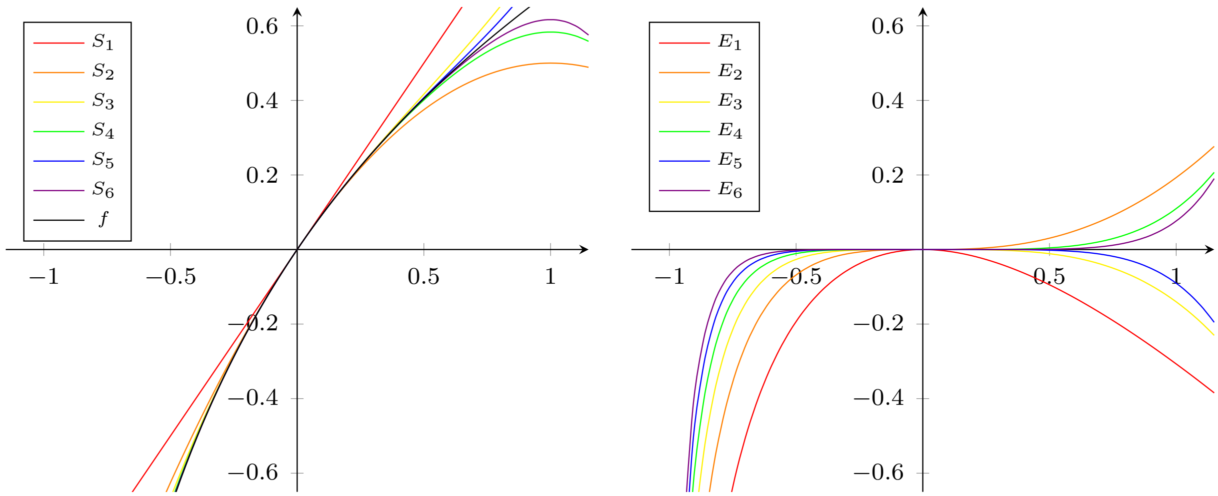
Proof. See Abbott. ■

We will not prove this theorem, and instead we will aim to understand what it is saying. Given any differentiable function f, the theorem says that the value of f at some point $x \in (-R, R) \setminus \{0\}$ can be expressed in terms of two quantities. The first is a polynomial evaluated at x all of whose data comes from the (successive) derivatives of f at a single point, namely 0. This is "infinitesimal data." The second quantity is essentially what is left over. However, it, too, can be expressed as a polynomial of the next degree. This is no longer infinitesimal since the point c could in general be larger than 0. The importance of the term E_N is not so much that it is the difference, but the actual expression for it in terms of $c \in (-|x|, |x|)$ illustrates that if the higher order derivatives of f are not unreasonably large, then the error term is very small.

Example 25.35. Consider the function $f : (-1, \infty) \to \mathbb{R}$ given by

$$(-1, \infty) \ni x \mapsto f(x) := \ln(x + 1). \tag{25.36}$$

The derivative of the logarithm function is obtained by the inverse function theorem from the previous problem set. The first few approximations using only the polynomial terms associated to the derivatives of f at 0 is given on the graph on the left while the difference between these Taylor approximations from the partial sums S_n and the actual function f is drawn on the graph on the right.



As the graph on the right indicates, the error term approaches zero as n increases. We can try to provide a bound for this error term using Lagrange's Remainder Theorem. Note that for $N \in \mathbb{N} \cup \{0\}$,

$$f^{(N+1)}(x) = \frac{(-1)^N N!}{(x+1)^{N+1}} \tag{25.37}$$

for all $x \in (-1, 1)$. Hence, fix $x \in (-1, 1)$. Then, for all $c \in (-|x|, |x|)$,

$$\big|f^{(N+1)}(c)\big| = \frac{N!}{(c+1)^{N+1}} \tag{25.38}$$

so that

$$\left|\frac{f^{(N+1)}(c)}{(N+1)!}x^{N+1}\right| \le \frac{N!}{(N+1)!}\left(\frac{|x|}{c+1}\right)^{N+1} = \frac{1}{N+1}\left(\frac{|x|}{c+1}\right)^{N+1}. \tag{25.39}$$

If c were arbitrary, this could tend to infinity, so a blind estimate will not help (take for example $x = -0.9$ and $c = -0.8$). Therefore, let us instead restrict attention to $x \in \left(-\frac{1}{2}, \frac{1}{2}\right)$. In this case,

$$\frac{|x|}{c+1} < 1 \tag{25.40}$$

so that

$$\lim_{N \to \infty} \big|E_N(x)\big| = 0. \tag{25.41}$$

Lagrange's Remainder Theorem is a useful pointwise result about the value of a function at a point x. However, the point c depends on x so that the theorem does not say much about the relationship between the entire *function* f and polynomials. Furthermore, as we saw in the previous example, it has its limitations. The Weierstrass approximation theorem provides a much stronger relationship.

Theorem 25.42 (Weierstrass Approximation Theorem)**.** *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be continuous. For any $\epsilon > 0$, there exists a polynomial p on $[a, b]$ satisfying*

$$\big|f(x) - p(x)\big| < \epsilon \qquad \forall\, x \in [a, b]. \tag{25.43}$$

In other words, there exists a sequence of polynomials on $[a, b]$ that converges uniformly to f.

After this lecture, it is recommended the student works through problems 1 and 2 on HW #7.

26 December 1

Read Abbott for fantastic motivation and history. One important message I would like to emphasize is the following. Historically, according to Abbott, the integral was thought of in terms of anti-derivatives. However, it was also noticed that such anti-derivatives are closely related to areas and volumes. Conceptually, a breakthrough was made when an alternative definition of the integral was sought out using the perspective of calculating areas and volumes. This perspective led to the notion of the Riemann integral and to several other kinds of integrals as well such as the Riemann-Stieltjes integral and the Lebesgue integral in increasing order of generality.[54] The Lebesgue integral is particularly important as it focused on a deeper question—what is the meaning of measure? How does one measure area, volume, and so on? What is the "length" of the rational numbers? As we saw earlier, we found that the rationals have measure zero. We also saw that the Cantor set has measure zero as well. We will not pursue the direction of measure here and will instead be content with focusing on specific kinds of measurable subsets of $\mathbb{R}$, namely, intervals. This perspective has its drawbacks in that Lebesgue integrable functions are more robust and have a characterization in terms of linear functionals on the space of all functions. However, it is visually simpler to understand and a good first step in the direction towards more general integrals. The following consists of several important definitions used for integrating functions.

Definition 26.1. Let $a, b \in \mathbb{R}$ with $a < b$. A <u>*finite partition*</u> P of $[a, b]$ consists of a a finite distinct ordered set of points $x_0, x_1, \ldots, x_{n-1}, x_n$, where $n \in \mathbb{N}$, such that

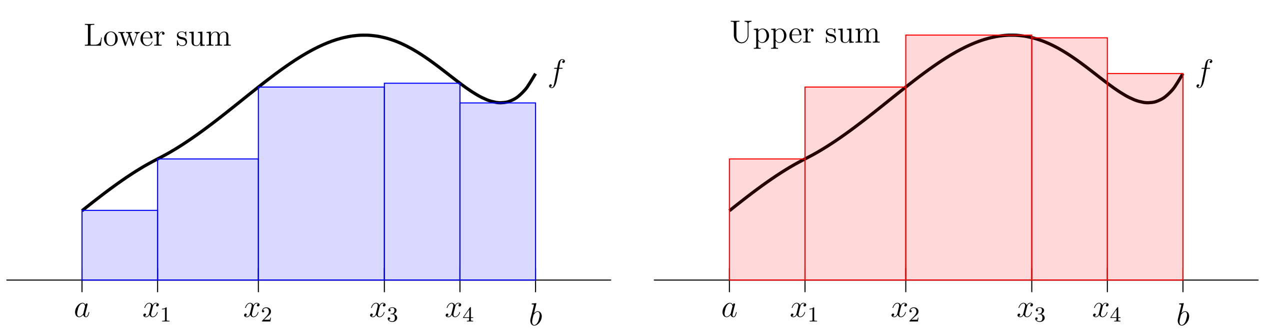
$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b. \tag{26.2}$$

Let $f : [a, b] \to \mathbb{R}$ be a bounded function. The <u>*lower sum*</u> of f with respect to P is

$$L[f; P] := \sum_{k=1}^{n} (x_k - x_{k-1}) \inf_{x \in [x_{k-1}, x_k]} \{f(x)\}. \tag{26.3}$$

The <u>*upper sum*</u> of f with respect to P is

$$U[f; P] := \sum_{k=1}^{n} (x_k - x_{k-1}) \sup_{x \in [x_{k-1}, x_k]} \{f(x)\}. \tag{26.4}$$



[54]You might think that it is difficult to do research in mathematics since there is so much to learn. However, if you ask the right question and view something that seems to be pretty well-understood from a different perspective that has never been done before, you could discover an entire new field of mathematics. Therefore, ask questions until you are completely satisfied. And do not be afraid to question traditional material, especially if you can back up your confusion with sound logic.

Note that the bounded condition on f is crucial. Indeed, there exist unbounded functions on closed intervals (remember, the functions we are considering need not be continuous).

Example 26.5. Let $f : [0,1] \to \mathbb{R}$ be the function defined by

$$[0,1] \ni x \mapsto f(x) := \begin{cases} \frac{1}{x} & \text{if } x = \frac{1}{2n} \text{ for some } n \in \mathbb{N} \\ -\frac{1}{x} & \text{if } x = \frac{1}{2n+1} \text{ for some } n \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases} \tag{26.6}$$

Then f is unbounded from above and below. In fact, since every partition of $[0,1]$ must have some interval containing 0, *all* upper and lower sums for *all* partitions are undefined.

Definition 26.7. Let P and Q be partitions of an interval $[a,b]$. Q is a *refinement* of P, written $P \le Q$, if Q contains all the points of P. Let $\mathcal{P}_a^b$ denote the set of all partitions of $[a,b]$.

Definition 26.8. Let $P,Q \in \mathcal{P}_a^b$. A *common refinement* of P and Q is a partition $P \vee Q \in \mathcal{P}_a^b$ satisfying

i) $P \le P \vee Q$ and $Q \le P \vee Q$ and

ii) for any other partition $R \in \mathcal{P}_a^b$ satisfying $P \le R$ and $Q \le R$, then $P \vee Q \le R$.

It can be shown that common refinements exist and are unique. In other words, the common refinement of P and Q is the smallest partition that contains both P and Q.

Lemma 26.9. *Let $a,b \in \mathbb{R}$ with $a < b$ and let $f : [a,b] \to \mathbb{R}$ be a bounded function.*

(a) Refinement $\le$ is a partial order on $\mathcal{P}_a^b$. This means

i) if $P,Q \in \mathcal{P}_a^b$ with $P \le Q$ and $Q \le P$, then $P = Q$, and

ii) if $P,Q,R \in \mathcal{P}_a^b$ with $P \le Q$ and $Q \le R$, then $P \le R$.

(b) If $P,Q \in \mathcal{P}_a^b$ and $P \le Q$, then $L[f;P] \le L[f;Q]$ and $U[f;P] \ge U[f;Q]$.

(c) If $P,Q \in \mathcal{P}_a^b$, then $L[f;P] \le U[f;Q]$.

(d) Let

$$U_f := (b-a) \max_{x\in[a,b]}\{f(x)\} \qquad \& \qquad L_f := (b-a) \min_{x\in[a,b]}\{f(x)\}. \tag{26.10}$$

Then

$$L_f \le L[f;P] \le U[f;P] \le U_f \qquad \forall\, P \in \mathcal{P}_a^b. \tag{26.11}$$

Proof.

(a) This follows immediately from the definition.

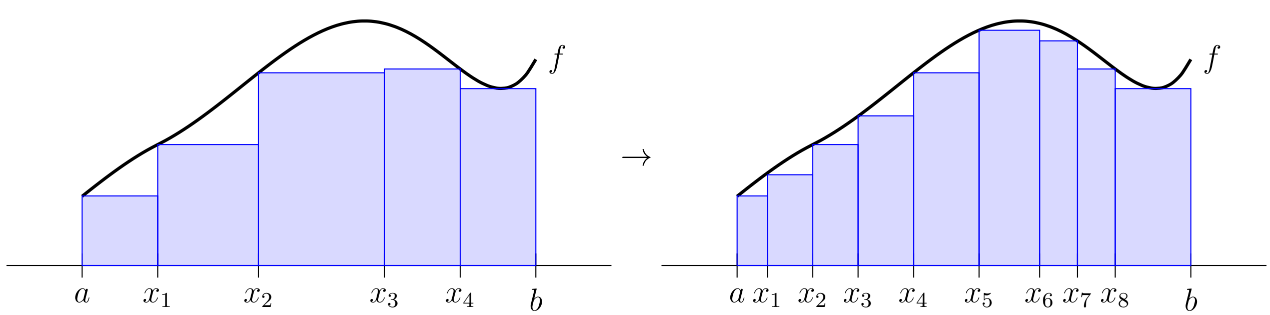
(b) Let $P = (a = x_0, x_1, \ldots, x_{n-1}, x_n = b)$ be a partition of $[a, b]$ and let Q be a refinement of P obtained by adding one extra point y, between, say x_{k-1} and x_k for some k, i.e. $Q = (a = x_0, x_1, \ldots, x_{k-1}, y, x_k, \ldots, x_{n-1}, x_n = b)$. Then,

$$\begin{aligned}(x_k - y) \inf_{x\in[y,x_k]} \{f(x)\} &+ (y - x_{k-1}) \inf_{x\in[x_{k-1},y]} \{f(x)\} \\ &\geq (x_k - y) \inf_{x\in[x_{k-1},x_k]} \{f(x)\} + (y - x_{k-1}) \inf_{x\in[x_{k-1},x_k]} \{f(x)\} \\ &= (x_k - x_{k-1}) \inf_{x\in[x_{k-1},x_k]} \{f(x)\}.\end{aligned} \tag{26.12}$$

Hence,

$$\begin{aligned} L[f; Q] &= \sum_{\substack{i=1\\ i\neq k}}^{n} (x_{i-1} - x_i) \inf_{x\in[x_{i-1},x_i]} \{f(x)\} \\ &\quad + (x_k - y) \inf_{x\in[y,x_k]} \{f(x)\} + (y - x_{k-1}) \inf_{x\in[x_{k-1},y]} \{f(x)\} \\ &\geq \sum_{\substack{i=1\\ i\neq k}}^{n} (x_{i-1} - x_i) \inf_{x\in[x_{i-1},x_i]} \{f(x)\} + (x_k - x_{k-1}) \inf_{x\in[x_{k-1},x_k]} \{f(x)\} \\ &= \sum_{i=1}^{n} (x_{i-1} - x_i) \inf_{x\in[x_{i-1},x_i]} \{f(x)\} \\ &= L[f; P]. \end{aligned} \tag{26.13}$$

Since every refinement of P is obtained by such "single step" refinements, this proves that $L[f; P] \leq L[f; Q]$ whenever $P \leq Q$.



A similar proof shows that $U[f; P] \geq U[f; Q]$.

(c) Let $P \vee Q$ be the common refinement of P and Q. Then, by definition of the upper and lower sums along with part (b) of this Lemma,

$$L[f; P] \leq L[f; P \vee Q] \leq U[f; P \vee Q] \leq U[f; Q]. \tag{26.14}$$

(d) Let I be the partition $I = \{a, b\}$ consisting of only the endpoints (the letter I is used for the word "initial"). Notice that $I \leq P$ for any partition $P \in \mathcal{P}_a^b$. Furthermore, $U_f = U[f; I]$ and $L_f = L[f; I]$. Thus, by part (b), the conclusion follows.

$\blacksquare$

With this Lemma, the following definition can be made.

Definition 26.15. Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : [a, b] \to \mathbb{R}$ be a bounded function. The <u>*upper and lower integrals*</u> of f on $[a, b]$ are

$$\overline{\int_a^b} f := \inf_{P \in \mathcal{P}_b^a} \{U[f; P]\} \qquad \& \qquad \underline{\int_a^b} f := \sup_{P \in \mathcal{P}_b^a} \{L[f; P]\}, \tag{26.16}$$

respectively. A function f as above is said to be <u>*Riemann integrable*</u> iff

$$\overline{\int_a^b} f = \underline{\int_a^b} f \tag{26.17}$$

and its <u>*Riemann integral*</u> over $[a, b]$ is given by this common value and is denoted by $\int_a^b f$ or $\int_a^b f(x) dx$.

Note that the upper integral and lower integrals always exist for bounded functions on closed intervals by the previous Lemma (L_f is a lower bound for all $U[f; P]$ and U_f is an upper bound for all $L[f; P]$).

Example 26.18. Consider the Dirichlet function f from Example 16.13 restricted to the domain $[-1, 1]$. Then $\overline{\int_{-1}^1} f = 1$ and $\underline{\int_{-1}^1} f = 0$ since every interval of nonzero length contains (infinitely many) rational and irrational numbers. Therefore, the Dirichlet function is not integrable.

We now list some necessary and/or sufficient conditions for functions to be Riemann integrable.

Theorem 26.19 (Integrability Criterion)**.** *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then the following are equivalent.*

(a) f is Riemann integrable.

(b) For every $\epsilon > 0$, there exists a partition P_ϵ of $[a, b]$ such that

$$U[f; P_\epsilon] - L[f; P_\epsilon] < \epsilon. \tag{26.20}$$

(c) There exists a sequence of partitions $(P_1, P_2, P_3, \ldots)$ of $[a, b]$ satisfying

$$\lim_{n\to\infty} \Big(U[f; P_n] - L[f; P_n] \Big) = 0. \tag{26.21}$$

Proof. See Abbott. $\blacksquare$

Theorem 26.22. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then f is Riemann integrable on $[a, b]$.*

Proof. See Abbott. $\blacksquare$

We will discuss more robust necessary and sufficient conditions for integrability in the next lecture. For homework, you will prove the following fact.

Exercise 26.23. Show that if $f : [a, b] \to \mathbb{R}$ is monotone, then f is Riemann integrable on $[a, b]$.

Lemma 26.24. *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f, g : [a, b] \to \mathbb{R}$ be two Riemann integrable functions. Let P be a partition of $[a, b]$. Then*

$$U[f + g; P] \le U[f; P] + U[g; P] \tag{26.25}$$

and

$$L[f + g; P] \ge L[f; P] + L[g; P]. \tag{26.26}$$

Proof. Let the partition P be given by the ordered set $P = (a = x_0, x_1, x_2, \ldots, x_{n-1}, x_n = b)$. For any $k \in \{1, \ldots, n\}$ and for any $y \in [x_{k-1}, x_k]$,

$$f(y) \le \sup_{x\in[x_{k-1},x_k]} \big\{f(x)\big\} \qquad \& \qquad g(y) \le \sup_{x\in[x_{k-1},x_k]} \big\{g(x)\big\}. \tag{26.27}$$

Hence,

$$f(y) + g(y) \le \sup_{x\in[x_{k-1},x_k]} \big\{f(x)\big\} + \sup_{x\in[x_{k-1},x_k]} \big\{g(x)\big\}, \tag{26.28}$$

which shows that $\sup_{x\in[x_{k-1},x_k]} \big\{f(x)\big\} + \sup_{x\in[x_{k-1},x_k]} \big\{g(x)\big\}$ is an upper bound for the set $\big\{f(y) + g(y) \ : \ y \in [x_{k-1}, x_k]\big\}$. But because the supremum of this set is the *least* upper bound, it follows that

$$\sup_{x\in[x_{k-1},x_k]} \big\{f(x) + g(x)\big\} \le \sup_{x\in[x_{k-1},x_k]} \big\{f(x)\big\} + \sup_{x\in[x_{k-1},x_k]} \big\{g(x)\big\}. \tag{26.29}$$

Hence,

$$\begin{aligned}
U[f + g; P] &= \sum_{k=1}^{n} (x_k - x_{k-1}) \sup_{x\in[x_{k-1},x_k]} \big\{f(x) + g(x)\big\} \\
&\le \sum_{k=1}^{n} (x_k - x_{k-1}) \Big(\sup_{x\in[x_{k-1},x_k]} \big\{f(x)\big\} + \sup_{x\in[x_{k-1},x_k]} \big\{g(x)\big\}\Big) \\
&= \sum_{k=1}^{n} (x_k - x_{k-1}) \sup_{x\in[x_{k-1},x_k]} \big\{f(x)\big\} + \sum_{k=1}^{n} (x_k - x_{k-1}) \sup_{x\in[x_{k-1},x_k]} \big\{g(x)\big\} \\
&= U[f; P] + U[g; P].
\end{aligned} \tag{26.30}$$

A similar calculation shows that $L[f + g; P] \ge L[f; P] + L[g; P]$. ■

Theorem 26.31. *Let $a, b, c \in \mathbb{R}$ with $a < b < c$ and let $f : [a, c] \to \mathbb{R}$. Then f is Riemann integrable on $[a, c]$ if and only if f is Riemann integrable on $[a, b]$ and $[b, c]$. In this case,*

$$\int_a^c f = \int_a^b f + \int_b^c f. \tag{26.32}$$

Proof. See Abbott. ■

Theorem 26.33. *Let $a, b \in \mathbb{R}$ with $a < b$, let $f, g : [a, b] \to \mathbb{R}$ be Riemann integrable, and let $k \in \mathbb{R}$. Then the following facts, hold.*

(a) $f + g$ is Riemann integrable on $[a, b]$ and

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g. \tag{26.34}$$

(b) kf is Riemann integrable on $[a, b]$ and

$$\int_a^b (kf) = k \int_a^b f. \tag{26.35}$$

(c) Let $m, M \in \mathbb{R}$ satisfy $m \le f(x) \le M$ for all $x \in [a, b]$. Then

$$m(b-a) \le \int_a^b f \le M(b-a). \tag{26.36}$$

(d) If $f(x) \le g(x)$ for all $x \in [a, b]$, then

$$\int_a^b f \le \int_a^b g. \tag{26.37}$$

(e) fg is Riemann integrable on $[a, b]$.

(f) $|f|$ is Riemann integrable on $[a, b]$ and

$$\left| \int_a^b f \right| \le \int_a^b |f|. \tag{26.38}$$

Proof.

(a) Fix $\epsilon > 0$. By the Integrability Criterion, there exist partitions P_f and P_g of $[a, b]$ such that

$$U[f; P_f] - L[f; P_f] < \frac{\epsilon}{2} \qquad \& \qquad U[g; P_g] - L[g; P_g] < \frac{\epsilon}{2}. \tag{26.39}$$

Set $P_\epsilon := P_f \vee P_g$ to be the common refinement of P_f and P_g. Then, by Lemma 26.9 part (b),

$$U[f; P_\epsilon] - L[f; P_\epsilon] < \frac{\epsilon}{2} \qquad \& \qquad U[g; P_\epsilon] - L[g; P_\epsilon] < \frac{\epsilon}{2}. \tag{26.40}$$

Hence, by Lemma 26.24,

$$\begin{aligned} U[f+g; P_\epsilon] - L[f+g; P_\epsilon] &\le U[f; P_\epsilon] + U[g; P_\epsilon] - L[f; P_\epsilon] - L[g; P_\epsilon] \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} \qquad \text{by (26.40)} \\ &= \epsilon. \end{aligned} \tag{26.41}$$

By the Integrability Criterion, this shows that $f + g$ is Riemann integrable. The fact that $\int_a^b (f+g) = \int_a^b f + \int_a^b g$ follows from the following list of equalities and inequalities

$$
\begin{tikzcd}[column sep=small]
 & \overline{\int_a^b} f + \overline{\int_a^b} g \arrow[r, equal] \arrow[dl, equal] & \underline{\int_a^b} f + \underline{\int_a^b} g \arrow[dr, equal] & \\
\inf\limits_{P\in\mathcal{P}_a^b} \{U[f;P]\} + \inf\limits_{P\in\mathcal{P}_a^b} \{U[g;P]\} \arrow[d, phantom, "\rotatebox{-90}{$\geq$}"] & & & \sup\limits_{P\in\mathcal{P}_a^b} \{L[f;P]\} + \sup\limits_{P\in\mathcal{P}_a^b} \{L[g;P]\} \arrow[d, phantom, "\rotatebox{-90}{$\leq$}"] \\
\inf\limits_{P\in\mathcal{P}_a^b} \{U[f+g;P]\} \arrow[dr, equal] & & & \sup\limits_{P\in\mathcal{P}_a^b} \{L[f+g;P]\} \arrow[dl, equal] \\
 & \overline{\int_a^b} (f+g) \arrow[r, equal] & \underline{\int_a^b} (f+g) &
\end{tikzcd}
\tag{26.42}
$$

where the inequalities follow from Lemma 26.24 and the equalities between the upper and lower integrals follow because f, g, and $f + g$ are Riemann integrable.

(b) See Abbott.

(c) See Abbott.

(d) See Abbott.

(e)

(f) See Abbott.

■

Remark 26.43. For homework, you will also show that for any non-negative Riemann integrable function $f : [a, b] \to \mathbb{R}$ (f is non-negative iff $f(x) \geq 0$ for all $x \in [a, b]$), the integral satisfies $\int_a^b f \geq 0$. Let I_a^b denote the set of integrable functions on $[a, b]$. Putting the previous theorem and this fact together shows that $\int_a^b$ can be viewed as a linear function $\int_a^b : I_a^b \to \mathbb{R}$ that sends non-negative elements (non-negative functions) to non-negative numbers. Similarly, given any Riemann integrable function $g : [a, b] \to \mathbb{R}$, the function

$$
I_a^b \ni f \mapsto \int_a^b fg \tag{26.44}
$$

is also linear and non-negative whenever g is non-negative (it is well-defined by the previous Theorem). One could then ask, are *all* linear and non-negative functions $I_a^b \to \mathbb{R}$ of the above form? Namely, given an arbitrary linear and non-negative function $\varphi : I_a^b \to \mathbb{R}$, does there exist a non-negative function $g \in I_a^b$ such that

$$\varphi(f) = \int_a^b fg \qquad \forall\ f \in I_a^b\ ? \tag{26.45}$$

If you've taken linear algebra, this should seem like a familiar question. First note that the above theorem says, in particular, that I_a^b is a real vector space (in fact, an algebra, since the product of two Riemann integrable functions is Riemann integrable). In the context of linear algebra, the question is phrased as follows. Let V be a vector space with an inner product $\langle\ \cdot\ ,\ \cdot\ \rangle$ and $\varphi : V \to \mathbb{R}$ a linear function. Does there exist a vector $v \in V$ such that $\varphi(w) = \langle v, w\rangle$ for all $w \in V$? I won't answer the above question for I_a^b and leave you to contemplate it—one often finds an answer in a course on measure theory or functional analysis.

After this lecture, it is recommended the student works through problem 3 on HW #7.

27 December 6

Theorem 27.1. *Let $a, b \in \mathbb{R}$ with $a < b$ and let $f : \mathbb{N} \to \mathbb{R}^{[a,b]}$ be a sequence of Riemann integrable functions that converge uniformly to* $\lim f$*. Then* $\lim f$ *is Riemann integrable and*

$$\lim_{n\to\infty} \int_a^b f_n = \int_a^b \lim f. \tag{27.2}$$

Proof. See Abbott. ∎

We will give a counter-example when the convergence in Theorem 27.1 is not uniform.

Example 27.3. Let $A := \mathbb{Q} \cap [0, 1]$. Since A is countable, there exists a bijection $\varphi : \mathbb{N} \to A$. Define the sequence of functions $f : \mathbb{N} \to \mathbb{R}^{[0,1]}$ by

$$\mathbb{N} \times [0, 1] \ni (n, x) \mapsto f_n(x) := \begin{cases} 1 & \text{if } x = \varphi(m) \text{ for some } m \in \{1, 2, \ldots , n\} \\ 0 & \text{otherwise} \end{cases} \tag{27.4}$$

Then, each f_n is Riemann integrable and its integral is (exercise!)

$$\int_0^1 f_n = 0. \tag{27.5}$$

However, $\lim f$ is the Dirichlet function

$$(\lim f)(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases} \tag{27.6}$$

which we have shown is not integrable. Therefore,

$$\lim_{n\to\infty} \int_a^b f_n \neq \int_a^b \lim f. \tag{27.7}$$

In fact, the right-hand-side is not even defined. Abbott gives an example of a sequence of integrable functions where the right-hand-side is defined but the two limits disagree.

Theorem 27.8 (Fundamental Theorem of Calculus)**.** *Let $a, b \in \mathbb{R}$ and let $f : [a, b] \to \mathbb{R}$ be Riemann integrable.*

(a) Let $F : [a, b] \to \mathbb{R}$ be a function satisfying $F' = f$ on $[a, b]$. Then

$$\int_a^b f = F(b) - F(a). \tag{27.9}$$

(b) Define $G : [a, b] \to \mathbb{R}$ by

$$[a, b] \ni x \mapsto G(x) := \int_a^x f. \tag{27.10}$$

Then G is continuous on $[a, b]$. Furthermore, if, in addition, f is continuous at $c \in [a, b]$, then G is differentiable at c and $G'(c) = f(c)$.

Proof. See Abbott. ∎

Theorem 27.11 (Lebesgue's Theorem)**.** *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Then f is Riemann integrable if and only if D_f, the set of discontinuities of f, has measure zero.*

Proof. See Abbott. ∎

28 December 8

Today we will review for the final.

References

[1] Stephen Abbott, *Understanding analysis*, 1st ed., Undergraduate Texts in Mathematics, Springer, 2001.

[2] ———, *Understanding analysis*, 2nd ed., Undergraduate Texts in Mathematics, Springer, 2015.

[3] David S. Dummit and Richard M. Foote, *Abstract algebra*, 3rd ed., Wiley, 2003.

[4] P. R. Halmos, *Naive set theory*, 1st ed., Undergraduate Texts in Mathematics, Springer, 1974.

[5] Tom Leinster, *Rethinking set theory*, American Mathematical Monthly **121** (2014), no. 5, 403–415, available at 1212.6543.

[6] Walter Rudin, *Principles of mathematical analysis*, 3rd ed., International Series in Pure & Applied Mathematics, McGraw-Hill, 1976.

[7] Michael Spivak, *Calculus on manifolds. A modern approach to classical theorems of advanced calculus*, W. A. Benjamin, Inc., New York-Amsterdam, 1965.

[8] N. Ya. Vilenkin, *Stories about sets*, Academic Press, 1968.