Real Analysis Lecture Notes (long version)

David M. McClendon

Department of Mathematics
Ferris State University

2025 edition

Contents

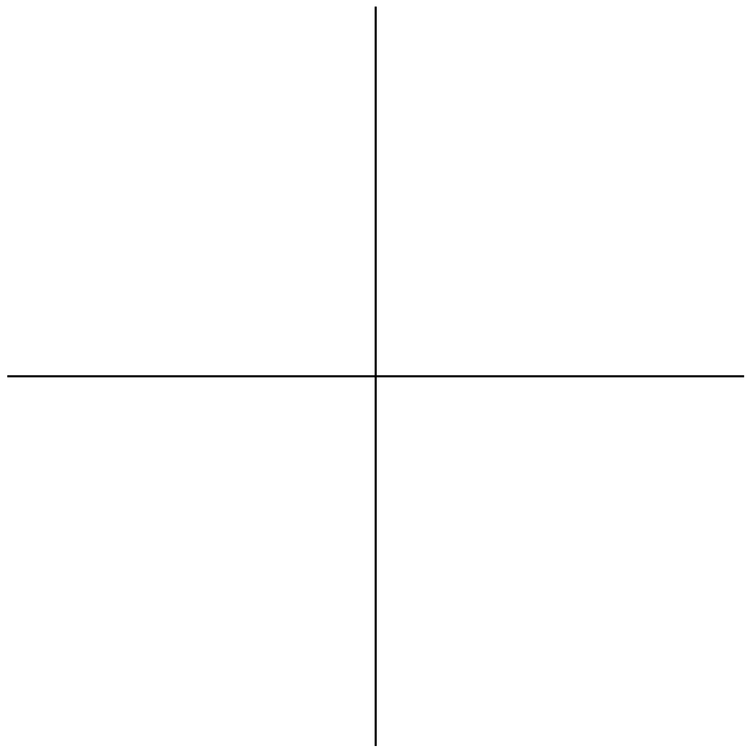
Chapter 1

Sets and functions

1.1 Why are we here?

I'd call this course **Real Analysis**, but at Ferris it's called **Advanced Calculus**.
That means this class probably has something (or a lot) to do with calculus.
So to get started, let's brainstorm what you learned (hopefully) in calculus:

Next, let's try an experiment.
Sketch the graph of any function you like on these blank axes:



Next, mark two values on the x-axis; call these values a and b (a is the smaller value, b is the bigger one).
Next, mark where $f(a)$ and $f(b)$ are on your graph.
Next, pick any number y between $f(a)$ and $f(b)$ and mark that number on your y-axis.
Question: In your example, is there a number x between a and b so that $f(x) = y$?

Once you finish, check out your classmates' graphs and see if they have the same answer to the question asked here.
Follow-up: Do you think the answer to this question is always the same, no matter what function you choose?

What we saw on the previous page is made formal in the following theorem:

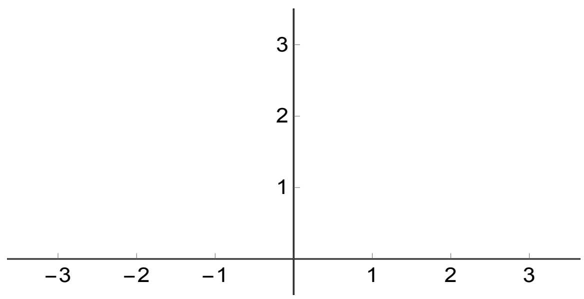
Theorem 1.1 (Intermediate Value Theorem (IVT)) *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous, and let $a < b$ in $\mathbb{R}$. Then, for any y between $f(a)$ and $f(b)$, there is $x \in (a, b)$ such that $f(x) = y$.*

REMARK

If you take this theorem and remove the hypothesis that f is continuous, the IVT is false.

To disprove the IVT in this setting, use a counterexample: let $f : \mathbb{R} \to \mathbb{R}$ be

$$f(x) = \begin{cases} 3 & x \geq 0 \\ 1 & x < 0 \end{cases}.$$



Then let $a = -1$ and $b = 1$.

2 is between $f(a) = 1$ and $f(b) = 3$, but there is no x such that $f(x) = 2$.

In light of this counterexample, the IVT must have something to do with what "continuous" means.

In Calculus 1, you are taught that a function $f : \mathbb{R} \to \mathbb{R}$ is *continuous* in an informal way and in a (slightly more) formal way:

QUESTION

What does it mean (informally) for a function $f : \mathbb{R} \to \mathbb{R}$ to be *continuous*?

Slightly more formally, you are told in Calculus 1 that $f : \mathbb{R} \to \mathbb{R}$ is *continuous* if it is *continuous at every a in its domain*, i.e.

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

for every a in the domain of f.

In particular, commonly used functions that are continuous include

(or at least you are told in Calculus 1 that these functions are continuous).

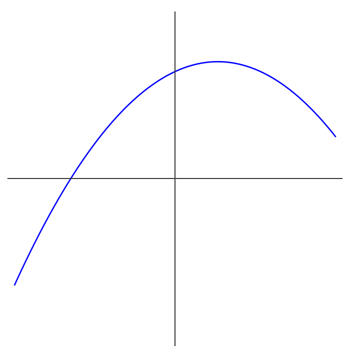
BUT... we are about to see that this approach to understanding continuity has some problems...

Here's another theorem often seen in Calculus 1:

Theorem 1.2 (Mean Value Theorem (MVT)) *Let $a < b$ be two real numbers. Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous at all $x \in [a, b]$ and differentiable at all $x \in (a, b)$. Then there is $x \in (a, b)$ such that*

$$f'(x) = \frac{f(b) - f(a)}{b - a}.$$

If you saw this theorem in Calc 1, you were probably shown a justification that relies on a picture (so it isn't a proof at all):



An alternative explanation of the MVT comes from physics (we'll talk about this later in the course when we prove the MVT).

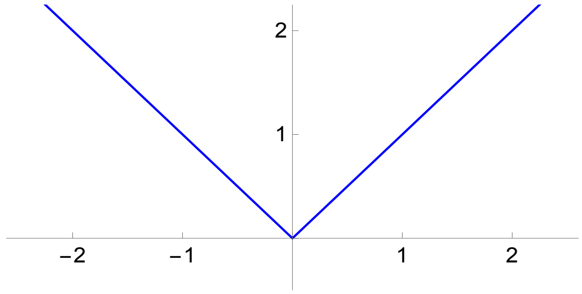
REMARK

If you remove the hypothesis that f is differentiable at all $x \in (a, b)$, the MVT is false.

Here's a counterexample: let $f(x) = |x|$, let $a = -1$ and let $b = 1$. Then

$$\frac{f(b) - f(a)}{b - a} = \frac{1 - 1}{1 - (-1)} = \frac{0}{2} = 0,$$

but at no point x between -1 and 1 is $f'(c) = 0$:



In light of this counterexample, the MVT must have a lot to do with what it means for a function $f : \mathbb{R} \to \mathbb{R}$ to be *differentiable*. In Calc 1, you are taught that a function $f : \mathbb{R} \to \mathbb{R}$ is differentiable it its derivative exists, meaning informally that its graph is...

Formally, to say that the derivative of f exists means that $f'(x)$ exists, where $f'(x)$ is defined as a limit:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Continuity and differentiability come from limits

In Calc 1, the ideas of continuity and differentiability are defined to you in terms of *limits*. But that begs the question of what "limit" means. In Calc 1, you are taught that

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

means

or some other equally meaningless garbage.

Then, you are taught how to compute limits (and subsequently derivatives and integrals) of functions with formulas including things like x^2, $\sin x$, e^x, $\ln x$, etc.

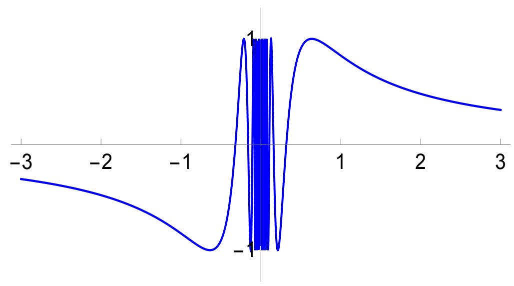
But this description of what a limit is isn't precise, can't really be applied to more sophisticated functions:

Some interesting functions

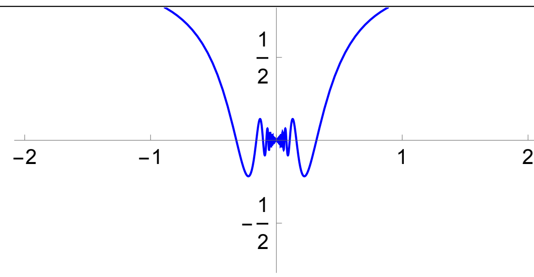
1. Functions of the form $x^m \sin \frac{1}{x^n}$, where $m \geq 0$ and $n > 0$ are integers

Are these continuous at $x = 0$? Differentiable at $x = 0$? Why or why not?

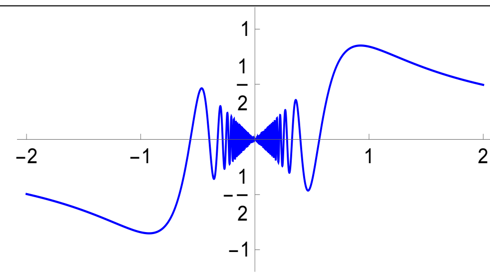
$$f(x) = \begin{cases} \sin \dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$



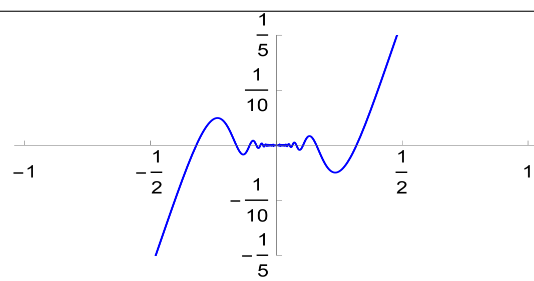
$$f(x) = \begin{cases} x \sin \dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$



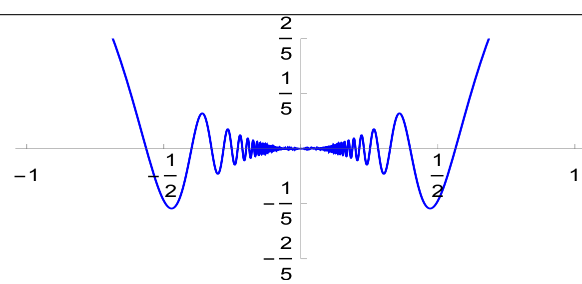
$$f(x) = \begin{cases} x \sin \dfrac{1}{x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$



$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$



$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$



2. Dirichlet's function

Let $f : \mathbb{R} \to \mathbb{R}$ (later to be called $\mathbb{1}_{\mathbb{Q}}$) be defined by $f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$.

- What would the graph of this look like?

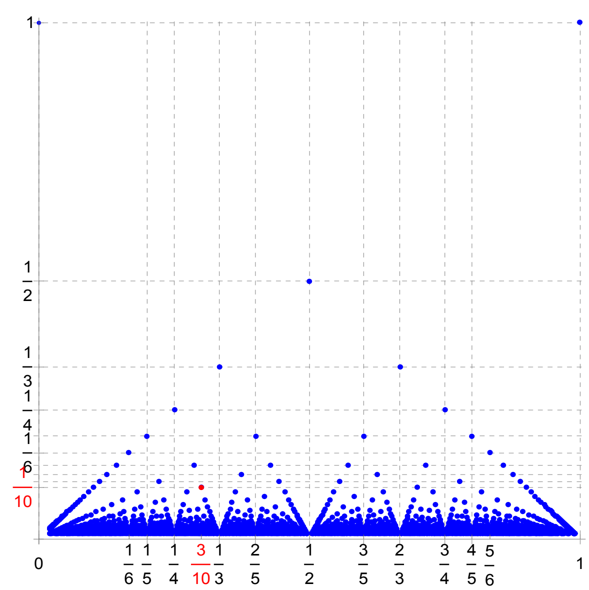
- Is this function continuous at any x? At all x? At no x?
- Is it differentiable anywhere? If so, where? What is the derivative?
- Is this function integrable? If so, what is $\int_0^1 f(x)\,dx$?

3. Thomae's function (a.k.a. raindrop function a.k.a. popcorn function)

Let $\tau : \mathbb{R} \to \mathbb{R}$ be defined by

$$\tau(x) = \begin{cases} \dfrac{1}{q} & \text{if } x \text{ is rational and } x = \dfrac{p}{q} \text{ in lowest terms, with } q > 0 \\ 0 & \text{if } x \text{ is irrational} \end{cases}.$$

- What does the graph of this τ look like?



- Is τ continuous at any x? At all x? At no x?
- Is τ differentiable anywhere? If so, where? What is its derivative?
- Is τ integrable? If so, what is $\int_0^1 \tau(x)\,dx$?

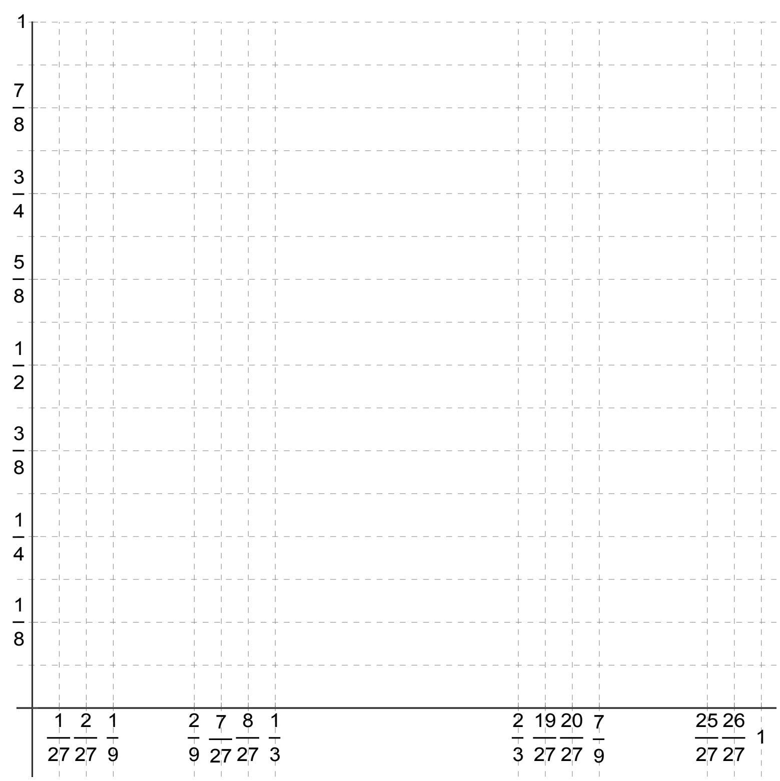
4. The Cantor function (a.k.a. the Devil's staircase)

Let $c : [0,1] \to [0,1]$ be "defined" as follows:
Set $c(0) = 0$ and $c(1) = 1$.
Now, take the interval on which f hasn't been defined yet and divide it into thirds.

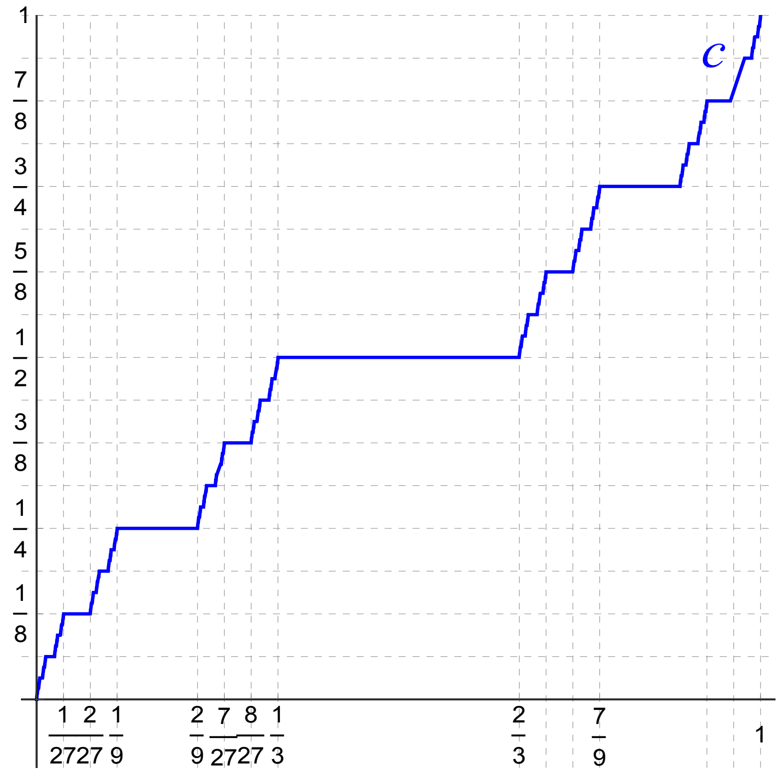
On the middle third, set $c(x) = \dfrac{f(0) + f(1)}{2} = \dfrac{1}{2}$.

Now repeat this procedure over and over: on any interval where f isn't defined yet (say (a,b)), divide that interval into thirds, and on the middle third, set

$$c(x) = \frac{f(a) + f(b)}{2}.$$



You end up with a function (or do you?) whose graph looks like this:



Questions:

- Does this actually rigorously define a function c from $[0,1]$ to $[0,1]$?
- If so, at which x is c continuous? (Is it continuous at all x?)
- If so, at which x is c differentiable? (Is it differentiable at all x?)
- Does c have an antiderivative? If so, what is it?
- Does $\int_0^1 c(x)\,dx$ exist? If so, what is it?

Why it's important to consider functions like these

If we can't adequately and precisely flesh out what's meant by the concepts of *limit*, *continuity* and *differentiability* in the context of these exotic examples, **why should we believe what these concepts are when we apply them** in calculus, differential equations, numerical analysis, probability. etc.? Put another way:

Why should we trust that the calculus computations we have been taught are actually valid?

Course goals

1. **Define precisely** what is meant by *limit*, *continuous* and *derivative* (and also what is meant by *integral*).

2. Use these precise definitions to **rigorously prove** the major theorems and techniques of MATH 220 (IVT, MVT, Fundamental Theorem of Calculus).

3. **Analyze** some of the **exotic functions** described above with regard to our precise notions of limit / continuity / differentiabilty / integrability.

It turns out that as a prerequisite to accomplishing these goals, we have to learn about some deep properties of the *real numbers* (for reasons that will be discussed later).

And to do that, we first need a refresher on some universal mathematical language, so that's where we're headed next.

1.2 Sets

The fundamental objects of mathematics are called *sets*. A set is really just a collection or list of objects (and in math, the objects are usually things like numbers, vectors, functions, perhaps other sets, etc.).

> **Definition 1.3** *A* **set** *is a definable collection of objects.*
> *The objects which comprise a set are called the set's* **elements**.
> *If x is an element of set E, we write $x \in E$.*
> *If x is not an element of set E, we write $x \notin E$.*

EXAMPLES OF SETS

Observe that sets are usually denoted by capital letters:

$$A = \{3, 5, 7, 9, 11\}$$
$$B = \{1, 2, 3, 4, 5, 6\}$$
$$C = \{3, 5, 7\}$$

The elements of set C described above are 3, 5 and 7.
For the set A above, $3 \in A$ and $5 \in A$ but $8 \notin A$.

Set-builder notation

We often define a set without listing the elements (using English language). For example, the sets A, B and C given above could be described, respectively, by saying

"let A be the set of odd numbers from 3 to 11";
"let B be the set of integers from 1 to 6";
"let C be the set of odd numbers from 3 to 7".

We also describe sets by using what is called **set-builder notation**: to describe the same sets A, B, C as above using set-builder notation, we would write (or say)

$$A = \{x : 3 \le x \le 11 \text{ and } x \text{ is odd}\}$$
$$B = \{x : 1 \le x \le 6 \text{ and } x \text{ is an integer}\}$$
$$C = \{x : 3 \le x \le 7 \text{ and } x \text{ is odd}\}.$$

The first statement above is interpreted as follows: it says that set A is equal to the set of numbers x such that (the colon means "such that" in mathematics) $3 \le x \le 11$ and x is odd. Notice that this is exactly the set $\{3, 5, 7, 9, 11\}$.

To show you a different kind of example: if you were defining some set of functions (instead of a set of numbers), then instead of x you'd write f, and then after the colon you'd describe what has to be true about f for the function f to be in the set.

For example, the set D of functions whose derivative at $x = 2$ is positive could be described by writing

$$D = \{f : f'(2) > 0\}.$$

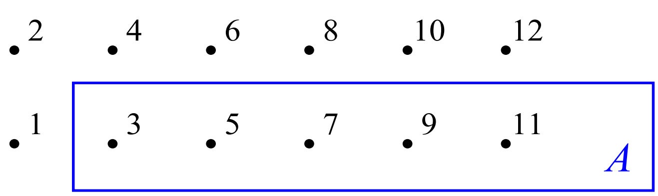
For this set D, it would be valid to say that if $g(x) = x^3$, then $g \in D$ (because $g'(2) = 3(2^2) = 12 > 0$) but if $h(x) = 3 - 4x$, then $h \notin D$ (because $h'(2) = -4 \leq 0$).

Definition 1.4 *The* **empty set***, denoted* $\varnothing$*, is the set with no elements.*

Venn diagrams

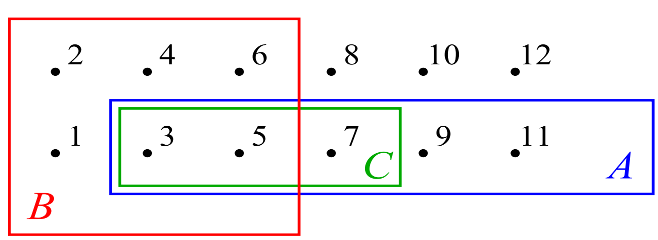
A useful way to think about sets is to draw pictures called **Venn diagrams**. To draw a Venn diagram, represent each set you're thinking about by a circle (or an oval, or a square, or a rectangle, or some other shape); think of an object as being an element of the set if and only if it is inside the shape corresponding to the set.

For example, a Venn diagram for the set A described above (recall that $A = \{3, 5, 7, 9, 11\}$) would be given by something like



because the box describing A contains exactly the elements of A (nothing more and nothing less).

Similarly, a Venn diagram representing the sets A, B and C from above would be something like



Venn diagram-style pictures can also be useful for subsets of real numbers like intervals: for example, if $S = \{x \in \mathbb{R} : x < 4\}$ and $T = \{x \in \mathbb{R} : x \geq -1\}$, we might draw S and T like this:

or

Subsets and equality of sets

Definition 1.5 *Let E and F be sets.*

We say E is a **subset** *of F, and write $E \subseteq F$, if for all $x, x \in E \Rightarrow x \in F$.*

If E is not a subset of F, we write $E \nsubseteq F$.

If $E \subseteq F$, we also write $F \supseteq E$ and say that F is a **superset** *of E.*

If $E \subseteq F$ and $F \subseteq E$. we say E and F are **equal**, *and write $E = F$,*

If E and F are not equal, we write $E \neq F$.

EXAMPLES

$\{0, 1, 2\} \subseteq \{0, 1, 2, 4, 8\}$ but $\{0, 1, 2\} \nsubseteq \{0, 2, 4\}$.

Note the difference between the symbols $\in$ and $\subseteq$: the first symbol should be preceded by an element, but the second symbol should be preceded by a set.

To say $E \subseteq F$ means "everything in E also is in F" or "E is inside F". If you draw a Venn diagram, to say $E \subseteq F$ means that the shape corresponding to set E is completely inside the shape corresponding to set F.

EXAMPLE

For the sets A and C given earlier, $C \subseteq A$ since every element of C is also in A.

To say two sets are equal means that they contain exactly the same elements.

EXAMPLE

$\{x : x^2 = x\} = \{x \in \mathbb{Z} : 0 \leq x < 2\}$ because the only elements in each set are 0 and 1.

Writing proofs about subset and set equality

The subset relationship $E \subseteq F$ can be restated as the conditional "if $x \in E$, then $x \in F$". This suggests a direct method for proving one set is a subset of another, called the **generic particular argument**:

GENERIC PARTICULAR ARGUMENT to prove $E \subseteq F$:

Suppose $x \in E$ ...... (some logical argument) .... Thus, $x \in F$.
Therefore $E \subseteq F$. □

RECALL

Sets E and F are equal iff $E \subseteq F$ and $F \subseteq E$.

This gives us a standard method of proving two sets are equal: you perform the generic particular argument twice, once to prove $E \subseteq F$ and again to prove $F \subseteq E$:

SET EQUALITY PROOF of $E = F$:

($\subseteq$) Suppose $x \in E$ ...... (some logical argument) .... Thus, $x \in F$.
($\supseteq$) Suppose $x \in F$ ...... (some logical argument) .... Thus, $x \in E$.
Since E and F are subsets of each other, $E = F$. □

Operations on sets

Definition 1.6 *Let E and F be sets.*
The **union** *of E and F, denoted $E \cup F$, is defined as*

$$E \cup F = \{x : x \in E \text{ or } x \in F\}.$$

The **intersection** *of E and F, denoted $E \cap F$, is defined as*

$$E \cap F = \{x : x \in E \text{ and } x \in F\}.$$

We say E and F are **disjoint** *if $E \cap F = \varnothing$.*
The **complement** *of E, denoted E^C, is the set $E^C = \{x : x \notin E\}$.*
The **difference** *of E and F, denoted $E - F$ and read "E minus F", is the set*

$$E - F = E \cap F^C.$$

Concepts:

- $\cup$ is set language for "or"–the union of a bunch of sets is the set consisting of elements belonging to *at least one* of the sets. For example, using the sets described earlier,
$$A \cup B = \{1, 2, 3, 4, 5, 6, 7, 9, 11\}$$
More generally, if you have a bunch of sets E_α indexed by some α, then the union of those sets, denoted $\bigcup\limits_\alpha E_\alpha$, is the set of things belonging to *at least one* of the E_α.

- $\cap$ is set language for "and"–the intersection of a bunch of sets is the set consisting of elements which belong to *all* of the given sets. For example, using the sets described A and B above,
$$A \cap B = \{3, 5\}$$
because the only numbers lying in both A and B are 3 and 5.
More generally, if you have a bunch of sets E_α indexed by some α, then the intersection of those sets, denoted $\bigcap\limits_\alpha E_\alpha$, is the set of things belonging to *all* of the E_α.

- Sets are disjoint if there are no objects which are both elements of E and elements of F.
EXAMPLE
$\{x \in \mathbb{R} : x < 0\}$ and $\{x \in \mathbb{Z} : x \geq 2\}$ are disjoint.

- Complement is set language for "not".

- The difference of E and F is the set of things in E, but not in F. For the sets A and C above, i.e.
$$A = \{3, 5, 7, 9, 11\} \quad \text{and} \quad C = \{3, 5, 7\},$$
we have $A - C = \{9, 11\}$ but $C - A = \varnothing$.

Mathematics shorthand

$\forall$ is shorthand for the phrase **for all**.

$\exists$ is shorthand for the phrase **there exists**.

"s.t." is shorthand for the words **such that**.

EXAMPLE

Suppose you see this:

$$\forall n \in \mathbb{Z}, \exists m \in \mathbb{Z} \text{ s.t. } m > n. \tag{1.1}$$

We read phrase (1.1) as

"For all n in the integers, there exists m in the integers such that $m > n$."

To internalize (1.1) better, you might re-read it as

"For all integers n, there's an integer m such that $m > n$."

If you read (1.1) again, you may understand it as

"For all integers n, there is an integer larger than n."

Last, if you think it about it a bit, you might realize (1.1) has the same intellectual content as

"There is no greatest integer."

"$\Rightarrow$" means **therefore** or **implies**. This means that whatever follows the "$\Rightarrow$" is a logical consequence of what comes before it.

EXAMPLE

$$x = 5 \;\Rightarrow\; x^2 = 25.$$

"$\Leftrightarrow$" means **if and only if (iff)**. This means that whatever precedes the "$\Leftrightarrow$" and whatever follows the "$\Leftrightarrow$" are *statements with truth values* that are true at exactly the same times.

EXAMPLE

$$x \text{ is even } \Leftrightarrow x = 2n \text{ for some integer } n.$$

What about this statement?

$$x = 5 \;\Leftrightarrow\; x^2 = 25.$$

Proofs of theorems start with PROOF and end with "$\square$". The $\square$ is a representation of a **gravestone**, which represents that the task of proving the theorem is complete, or "dead and buried".

1.3 Functions

In MATH 324 (Proofs), we learn a technical definition of a function which makes precise the idea of a "function" that you first encounter in high-school algebra or precalculus.
Generally speaking, this technical definition isn't useful, but it's worth stating:

Definition 1.7 *Let A and B be sets.*
A **function**, *a.k.a.* **map** *f from A to B is a rule that assigns to each element $x \in A$ exactly one element $f(x) \in B$.*
This $f(x)$ is called the **value of** f **at** x, *or the* **image of** x **under** f.
The notation $\boxed{f : A \to B}$ *means that f is a function from A to B.*
A is called the **domain** *of f and B is called the* **codomain** *of f.*

Definition 1.8 *Let $f : A \to B$.*
The **range**, *a.k.a.* **image** *of f, denoted $Range(f)$ or $Im(f)$, is the set of the function's values:*

$$Range(f) = Im(f) = \{y \in B : \exists x \in A \text{ s.t. } f(x) = y\}.$$

EXAMPLE

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

- the domain of f is $\mathbb{R}$;
- the codomain of f is $\mathbb{R}$;
- the range of f is $[0, \infty)$.

EXAMPLE

Let $f : \mathbb{R} - \{0\} \to \mathbb{R}$ be $f(x) = \dfrac{1}{x}$. Then:

- the domain of f is $\mathbb{R} - \{0\}$;
- the codomain of f is $\mathbb{R}$;
- the range of f is $\mathbb{R} - \{0\}$.

Definition 1.9 (Equality of functions) *To say two functions f and g are* **equal** *means that they have the same domain, and for all x in that common domain,*

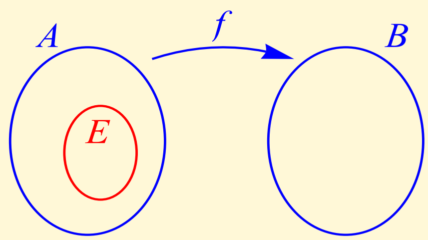
$$f(x) = g(x).$$

Images and preimages

Definition 1.10 *Let $f : A \to B$.*

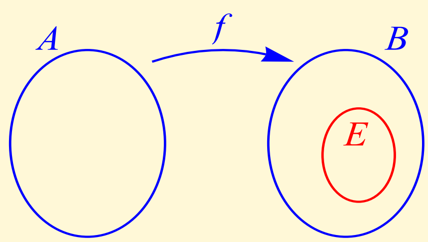
- *Given $E \subseteq A$, the* **image of E under** f*, denoted $f(E)$, is the set*

$$f(E) = \{y \in B : \exists\, x \in E \text{ s.t. } y = f(x)\}.$$



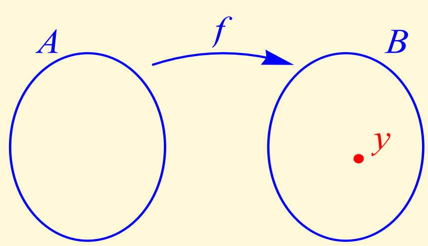
- *Given $E \subseteq B$, the* **preimage (of E under f)**, *also called the* **inverse image (of E under f)**, *denoted $f^{-1}(E)$, is the set*

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



- *Given $y \in B$, the* **preimage (of y under f)**, *also called the* **inverse image (of y under f)**, *denoted $f^{-1}(y)$, is the set defined by*

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



To emphasize, the preimage of a point is a set.

Example

Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = x^2$.
Then $f^{-1}(25) = \{-5, 5\}$ since both 5 and -5 map to 25 under f.

Theorem 1.11 *Let $f : A \to B$. Then:*

- *for any set $E \subseteq A$, $f^{-1}(f(E)) \supseteq E$;*
  WARNING: *in general, $f^{-1}(f(E)) \neq E$.*
- *for any set $E \subseteq B$, $f(f^{-1}(E)) = E \cap Im(f)$.*
  WARNING: *in general, $f(f^{-1}(E)) \neq E$.*

PROOF HW

Indicator functions

Definition 1.12 *Let A be any set and let $E \subseteq A$. The* **indicator function** *of E, a.k.a. the* **characteristic function** *of E, denoted $\mathbb{1}_E$, is the function $\mathbb{1}_E : A \to \mathbb{R}$ defined by*

$$\mathbb{1}_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{else} \end{cases} .$$

EXAMPLE

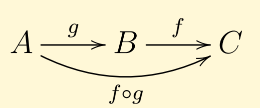
Dirichlet's function (that we encountered earlier) is the indicator function $\mathbb{1}_{\mathbb{Q}}$ of the rational numbers:

$$\mathbb{1}_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R} \qquad \mathbb{1}_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

Compositions

Definition 1.13 *Let $g : A \to B$ and let $f : B \to C$.*
Define the **composition** *of f with g, denoted $f \circ g$, to be the function from A to C defined by the rule*

$$(f \circ g)(x) = f(g(x)).$$



EXAMPLE

If $f : \mathbb{R}^2 \to \mathbb{R}$ is $f(x,y) = x^2 - y$ and $g : \mathbb{R} \to \mathbb{R}^2$ is $g(t) = (t-2, 4t+3)$, then $f \circ g : \mathbb{R} \to \mathbb{R}$ has rule

$$(f \circ g)(t) = f(g(t)) = f(t-2, 4t+3) = (t-2)^2 - (4t+3).$$

Injectivity

An *injection* (a.k.a. a $1-1$ *function*) is a function which takes different inputs to different outputs. More precisely:

> **Definition 1.14** *A function* $f : A \to B$ *is called* **injective***, a.k.a.* **one-to-one***, a.k.a.* $1-1$*, if for every* $x, y \in A$*,*
>
> $$f(x) = f(y) \text{ implies } x = y.$$

Equivalent characterizations of injectivity:

1. $f(x) = f(y)$ implies $x = y$.
2. $x \neq y$ implies $f(x) \neq f(y)$.
3. Different inputs go to different outputs.
4. For all $y \in B$, there is at most one $x \in A$ s.t. $f(x) = y$.
5. f passes the Horizontal Line Test (in the situation where $f : \mathbb{R} \to \mathbb{R}$).

EXAMPLES

$f : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^2$ is not injective because $f(1) = f(-1) = 1$.

$f : [0, \infty) \to \mathbb{R}$ given by $f(x) = x^2$ is injective:
$\forall\, y \in \mathbb{R}$, there is at most one x in $[0, \infty)$ s.t. $f(x) = x^2 = y$.

> **PROVING** that $f : A \to B$ is injective:
>
> Suppose $x, y \in A$ are such that $f(x) = f(y)$.
>
> ......
>
> Therefore, $x = y$.
>
> Therefore f is $1-1$. □

EXAMPLE

Let's prove that $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 3(x-1)^3 - 2$ is injective:

DISPROVING that $f : A \to B$ is injective:

Let $x =$ and $y =$ (choose specific $x, y \in A$). Note $x \neq y$.

......

Therefore, $f(x) = f(y)$.

Therefore f is not $1-1$. $\square$

EXAMPLE

Let's prove $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \cos x$ is not injective.

Theorem 1.15 *Let* $f : B \to C$ *and* $g : A \to B$ *be functions.*

1. *If* f *and* g *are both injective, then* $f \circ g$ *is injective.*
2. *If* $f \circ g$ *is injective, then* g *is injective.*

PROOF

When proving a statement of the form "if P, then Q", start by supposing that P is true.

To show a function is injective, use the recipe on the preceding page.

Again, this is a statement of the form "if P, then Q".

As before, to show g is injective, use the recipe.

Tell the reader the proof is finished.

Surjectivity

An *surjection* (a.k.a. an *onto function*) is a function which "hits" every point in its codomain. More precisely:

> **Definition 1.16** *A function $f : A \to B$ is called* **surjective**, *a.k.a.* **onto**, *if $f(A) = B$.*

Equivalent characterizations of surjectivity:

1. $Im(f) = f(A) = B$.
2. $B \subseteq Im(f)$.
3. Every potential output of f is an actual output.
4. For every $y \in B$, there is an $x \in A$ such that $f(x) = y$.

EXAMPLE

$f : \mathbb{R} \to \mathbb{R}$ where $f(x) = x^2$ is not onto, since $-1 \notin f(\mathbb{R})$.

EXAMPLE

$f : \mathbb{R} \to [0, \infty)$ where $f(x) = x^2$ is onto:
$\forall\, y \in [0, \infty)$, we can let $x = \sqrt{y}$. Then $f(x) = y$.

PROVING that $f : A \to B$ is surjective:

Let $y \in B$.
Write a formula for some $x \in A$ (that comes from some scratch work).
Show that for the x you wrote down, $f(x) = y$.
Conclude that f is onto. □

DISPROVING that $f : A \to B$ is injective:

Find a specific $y \in B$.
Prove that there is no such $x \in A$ s.t. $f(x) = y$.
(Usually you do this by assuming there is such an x, and deriving a contradiction.)
Conclude that f is not onto. □

Theorem 1.17 *Let $f : B \to C$ and $g : A \to B$ be functions.*

- *If f and g are both surjective, then $f \circ g$ is surjective.*
- *If $f \circ g$ is surjective, then f is surjective.*

PROOF HW

Bijectivity and inverse functions

A function which is both $1-1$ and onto is called a *bijection*:

Definition 1.18 *A function $f : A \to B$ is called* **bijective** *if f is both injective and surjective.*

Equivalent characterizations of bijectivity:

1. f is both surjective and injective.
2. For every $y \in B$, there is one and only one $x \in A$ such that $f(x) = y$.
3. Every point in the codomain has a unique preimage.

Theorem 1.19 *If $f : B \to C$ and $g : A \to B$ are bijections, then $f \circ g$ is a bijection.*

The main reason we care about bijections is that bijections are exactly the functions that have inverses which are also functions:

Definition 1.20 *Let $f : A \to B$ be a function (with $Dom(f) = A$).*
If there is another function $f^{-1} : B \to A$ (with $Dom(f^{-1}) = B$) such that

$$\forall x \in A, f^{-1}(f(x)) = x \quad \text{and} \quad \forall y \in B, f(f^{-1}(y)) = y$$

then we say f is **invertible** *and that f^{-1} is an* **inverse (function)** *of f.*

EXAMPLE

Let $f : \mathbb{R} \to (0, \infty)$ be $f(x) = e^x$.
Then $f^{-1} : (0, \infty) \to \mathbb{R}$ is $f^{-1}(x) = \ln x$.
These are inverses because

$$f^{-1}(f(x)) = \ln e^x = x \quad \text{and } f(f^{-1}(x)) = e^{\ln x} = x.$$

Theorem 1.21 (Properties of inverse functions) *Let $f : A \to B$ and $g : B \to C$.*

1. *f is invertible if and only if f is bijective.*
2. *If f is invertible, then f has only one inverse function.*
3. *If f is invertible, then f^{-1} is invertible, and $(f^{-1})^{-1} = f$.*
4. *If f and g are invertible, then $f \circ g$ is invertible, and $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.*

WARNINGS on the notation "f^{-1}": the symbol f^{-1} is used for preimage and inverse function.
Unless you know (or have proved) that the function f is invertible, f^{-1} means preimage, and is not actually referring to a function named "f^{-1}".

PROVING that $f : A \to B$ is a bijection:

1. Prove f is surjective.
2. Prove f is injective.

3 Conclude that f is a bijection. □

PROVING that $f : A \to B$ is a bijection
(by constructing an inverse function of f):

Write down a formula for $f^{-1} : B \to A$.
Show that for any $x \in A$, $f^{-1}(f(x)) = x$.
Show that for any $y \in B$, $f(f^{-1}(y)) = y$.
Conclude that f is invertible, hence f is a bijection. □

DISPROVING that $f : A \to B$ is a bijection:

Either prove f is not surjective, or prove that f is not injective.
Therefore, f is not a bijection. □

1.4 Cardinality

Definition 1.22 *A set E is called* **finite** *if there is an injective function $f : E \to \{1, 2, 3..., n\}$ for some $n \in \mathbb{N}$.*
A set E is called **infinite** *if it is not finite.*
A set E is called **countable** *if there is an injective function $f : E \to \mathbb{N}$.*
A set E is called **uncountable** *if it is not countable.*

The purpose of the function f in these definitions is that f "counts" the elements in E. For instance, if

$$E = \{0, 5, 11, -\pi, \sqrt{19}\},$$

To precisely describe what we did above, define the function $f : E \to \{1, 2, 3, 4, 5\}$ by

$$f(0) = 1 \qquad f(5) = 2 \qquad f(11) = 3 \qquad f(-\pi) = 4 \qquad f(\sqrt{19}) = 5.$$

Since f takes different inputs to different outputs, f is injective, so E is finite.

It turns out that if E is finite, then there is one and only one natural number n such that there is a bijection between E and $\{1, 2, 3, ..., n\}$. This n is called the **cardinality** of E and is denoted $\#(E)$.
For the set $E = \{0, 5, 11, -\pi, \sqrt{19}\}$, we have $\boxed{\#(E) = 5}$.

Cardinality properties that are immediate from these definitions:

1. Every finite set is countable, because if $f : E \to \{1, ..., n\}$ is injective, the same f is an injection from E to $\mathbb{N}$. Put another way, every uncountable set is infinite.

2. A subset of a finite set is finite, because if $f : E \to \{1, ..., n\}$ is injective, then for any subset $F \subseteq E$, restricting f to the subset F gives an injection from F to $\{1, ..., n\}$.

3. A subset of a countable set is countable, because if $f : E \to \mathbb{N}$ is injective, then for any subset $F \subseteq E$, restricting f to E gives an injection from F to $\mathbb{N}$.

4. The empty set is finite, because technically the empty set is an injective function from $\varnothing$ to any other set (like $\mathbb{N}$).

5. The set $\mathbb{N}$ of natural numbers is countable, because the function $f(n) = n$ is an injection from $\mathbb{N}$ to $\mathbb{N}$.

What's not so easy to show is:

Theorem 1.23 *The set $\mathbb{N}$ of natural numbers is infinite.*

PROOF Suppose $\mathbb{N}$ is finite.
Then $\exists\, f : \mathbb{N} \to \{1, ..., n\}$ which is injective.
Now, consider the set $\{f(1), f(2), ..., f(n+1)\}$.
This set has $n+1$ different elements, since f is injective.
But it is a subset of $\{1, ..., n\}$ which only has n elements. Contradiction![1]
Therefore $\mathbb{N}$ is infinite. □

As an immediate consequence, any set which has $\mathbb{N}$ as a subset (like $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$) is infinite.

Theorem 1.24 *The set $\mathbb{Z}$ of integers is countable.*

Idea of proof: To show $\mathbb{Z}$ is countable, we have to "count" the integers.

$$\cdots \quad -5 \quad -4 \quad -3 \quad -2 \quad -1 \quad 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad \cdots$$

PROOF Let $f : \mathbb{Z} \to \mathbb{N}$ be defined by

We claim this f is injective. To show this,

Since $f : \mathbb{Z} \to \mathbb{N}$ is injective, then by definition, $\mathbb{Z}$ is countable. □

[1]Actually, the fact that no set of n elements has a subset with $n+1$ elements, while seemingly obvious, needs proof and is actually tricky to prove. Google "pigeonhole principle" for more on this.

Theorem 1.25 *The set $\mathbb{Z}\times\mathbb{Z}$, which is the set of ordered pairs of integers, is countable.*

PROOF To show $\mathbb{Z} \times \mathbb{Z}$ is countable, we have to "count" the points (x, y) where $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$. Here's how we do this:

$$
\begin{array}{ccccccccc}
\ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \iddots \\
\cdots & \bullet(-3,3) & \bullet(-2,3) & \bullet(-1,3) & \bullet(0,3) & \bullet(1,3) & \bullet(2,3) & \bullet(3,3) & \cdots \\
\cdots & \bullet(-3,2) & \bullet(-2,2) & \bullet(-1,2) & \bullet(0,2) & \bullet(1,2) & \bullet(2,2) & \bullet(3,2) & \cdots \\
\cdots & \bullet(-3,1) & \bullet(-2,1) & \bullet(-1,1) & \bullet(0,1) & \bullet(1,1) & \bullet(2,1) & \bullet(3,1) & \cdots \\
\cdots & \bullet(-3,0) & \bullet(-2,0) & \bullet(-1,0) & \bullet(0,0) & \bullet(1,0) & \bullet(2,0) & \bullet(3,0) & \cdots \\
\cdots & \bullet(-3,-1) & \bullet(-2,-1) & \bullet(-1,-1) & \bullet(0,-1) & \bullet(1,-1) & \bullet(2,-1) & \bullet(3,1) & \cdots \\
\cdots & \bullet(-3,-2) & \bullet(-2,-2) & \bullet(-1,-2) & \bullet(0,-2) & \bullet(1,-2) & \bullet(2,-2) & \bullet(3,2) & \cdots \\
\cdots & \bullet(-3,-3) & \bullet(-2,-3) & \bullet(-1,-3) & \bullet(0,-3) & \bullet(1,3-) & \bullet(2,-3) & \bullet(3,-3) & \cdots \\
\iddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
$$

In case you don't think this argument is rigorous enough, here's a (horrible but correct) formula for the function $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{N}$ that does exactly what's described with the picture above:

$$
f(p, q) = \begin{cases}
1 & \text{if } (p, q) = (0, 0) \\
4p^2 - 4p + q + 2 & \text{if } p > 0 \text{ and } 0 \le q \le p \\
4p^2 + 4p + q + 2 & \text{if } p > 0 \text{ and } -p < q < 0 \\
4q^2 - 2q - \frac{|q|}{q}(p + 2) & \text{if } q \ne 0 \text{ and } -|q| \le p \le |q| \\
4p^2 - q + 2 & \text{if } p < 0 \text{ and } -|p| \le q < |p|
\end{cases}
$$

With some work (actually, with quite a bit of work), you can show this function is injective. □

Theorem 1.26 *The set $\mathbb{Q}$ of rational numbers is countable.*

PROOF First, assume that every rational number is written in lowest terms as $\frac{p}{q}$, where $q > 0$.

Then, let $g : \mathbb{Q} \to \mathbb{Z} \times \mathbb{Z}$ be defined by $g\left(\frac{p}{q}\right) = (p, q)$.

g is pretty clearly injective.

Since $\mathbb{Z} \times \mathbb{Z}$ is countable, $\exists$ injection $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{N}$ (as in the previous theorem).

The composition $f \circ g : \mathbb{Q} \to \mathbb{N}$ is the composition of two injections, so it is an injection, and therefore $\mathbb{Q}$ is countable. □

Theorem 1.27 (A countable union of countable sets is countable) *If $E_1, E_2, \ldots$ are all countable sets, then so is $\bigcup_{k=1}^{\infty} E_k$.*

PROOF Since each E_k is countable, for each k there is an injection $f_k : E \to \mathbb{N}$, which we can think of as an injection $f_k : E \to \mathbb{Z}$.

Define $f : \bigcup_{n=1}^{\infty} E_n \to \mathbb{Z} \times \mathbb{Z}$ as follows: for $x \in \bigcup_{n=1}^{\infty} E_n$, let $n(x)$ be the smallest n so that $x \in E_n$. Then set $f(x) = (n(x), f_{n(x)}(x))$.

Claim: f is injective.

Proof of Claim: Suppose $f(x) = f(y)$.

This means $(n(x), f_{n(x)}(x)) = (n(y), f_{n(y)}(y))$, so $n(x) = n(y)$.
Thus $x, y \in E_n$, where $n = n(x) = n(y)$.
Furthermore, $f_n(x) = f_n(y)$. But f_n is injective, so $x = y$.
This proves the claim.

By Theorem 1.25, $\mathbb{Z} \times \mathbb{Z}$ is countable, so $\exists$ an injection $g : \mathbb{Z} \times \mathbb{Z} \to \mathbb{N}$.

Then $g \circ f : \bigcup_{n=1}^{\infty} E_n \to \mathbb{N}$ is an injection, making $\bigcup_{n=1}^{\infty} E_n$ countable. □

QUESTION

Is there such a thing as an uncountable set?

Remember this: If a mathematician (like me) asks you **"How many ... are there?"**, almost always the mathematician is looking for one of three answers:

An application in probability theory

One of the important building blocks in probability theory is the idea of a *probability space*, which consists of:

- a set Ω,
- a σ−algebra $\mathcal{A}$ of subsets of Ω, meaning
  - $\mathcal{A}$ is non-empty, i.e. $\Omega \in \mathcal{A}$;
  - $\mathcal{A}$ is closed under complements, i.e. if $E \in \mathcal{A}$, then $E^C \in \mathcal{A}$;
  - $\mathcal{A}$ is closed under countable unions and countable intersections, i.e. for any countable set I, if $\{E_j : j \in I\}$ are in $\mathcal{A}$, then $\bigcup_{j\in I} E_j$ and $\bigcap_{j\in I} E_j$ are both in $\mathcal{A}$.
- a probability measure P, which is a function $P : \mathcal{A} \to \mathbb{R}$ such that
  - P is positive, i.e. $P(E) \geq 0$ for all $E \in \mathcal{A}$;
  - P is normalized, i.e. $P(\Omega) = 1$; and
  - P is countably additive, i.e. for any countable set I and any disjoint sets $\{E_j : j \in I\}$ in $\mathcal{A}$, we have

$$P\left(\bigcup_{j\in I} E_j\right) = \sum_{j\in I} P(E_j).$$

EXAMPLE (FOR THOSE WHO HAVE TAKEN MATH 414

Suppose X is chosen uniformly on the interval $[0,1]$. What is $P(X \in \mathbb{Q})$?

1.5 Chapter 1 Summary

DEFINITIONS AND SYMBOLS TO KNOW

Nouns

- A **set** is a definable collection of objects called the **elements** of the set.
  $x \in E$ means x is an element of set E.
  $E \subseteq F$ means E is a **subset** of F, i.e. $x \in E$ implies $x \in F$.
- The **empty set** $\varnothing$ is the set with no elements.
- The **union** $E \cup F$ is the set of things in E or F (or both).
- The **intersection** $E \cap F$ is the set of things in both E and F.
- The **complement** E^C is the set of things not in E.
- The **difference** $E - F$ is the set of things in E but not F.
- A **function** $f : A \to B$ is a rule that assigns to each $x \in A$ one element $f(x) \in B$.
  $f(x)$ is called the **value** of f at x.
  The set A of inputs to f is called the **domain** of f.
  The set B of possible outputs of f is called the **codomain** of f.
  If $E \subseteq A$, then the **image** $f(E)$ is the set of outputs obtained from inputs in E.
  $f(A)$ is called the **range** of f.
  If $E \subseteq B$, the **preimage** $f^{-1}(E)$ is the set of inputs which produce an output in E.
- The **inverse** of $f : A \to B$ is a function $f^{-1} : B \to A$ so that $f^{-1} \circ f(x) = x$ for all $x \in A$ and $f \circ f^{-1}(x) = x$ for all $x \in B$. (Not every function has an inverse.)

Adjectives that describe sets

- Two sets are called **disjoint** if they have no elements in common.
- A set is called **finite** if there is an injection from it to $\{1, 2, 3..., n\}$ for some $n \in \mathbb{N}$.
- A set is called **countable** if there is an injection from it to $\mathbb{N}$.
  Otherwise, the set is called **uncountable**.

Adjectives that describe functions

- f is **injective (**$1-1$**)** if $f(a) = f(b)$ implies $a = b$.

- $f : A \to B$ is **surjective (onto)** $f(A) = B$, i.e. for every $y \in B$ there is $x \in A$ so that $f(x) = y$.
- f is **bijective** if it is injective and surjective; this is equivalent to f being **invertible**, which means the inverse function f^{-1} exists.

Symbols

- $\forall$ means "for all".
- $\exists$ means "there exists".
- s.t. means "such that".
- $\Rightarrow$ means "therefore" or "implies".
- $\Leftrightarrow$ means "if and only if"
- $\square$ means "end of proof".
- $\mathbb{N}$ is the set of natural numbers.
  $\mathbb{Z}$ is the set of integers.
  $\mathbb{Q}$ is the set of rational numbers.
- $\#(E)$ is the **cardinality** of E (the number of elements in E).

Important examples of functions to remember

- Functions of the form $f(x) = x^m \sin \frac{1}{x^n}$ where $m, n \in \mathbb{N}$.
- The **indicator function** of set E is the function $\mathbb{1}_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}$.
- The **Dirichlet function** is $\mathbb{1}_{\mathbb{Q}}(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$.
- **Thomae's function** is the function $\tau : \mathbb{R} \to \mathbb{R}$ defined by

$$\tau(x) = \begin{cases} \frac{1}{q} & x \in \mathbb{Q}, x = \frac{p}{q} \text{ in lowest terms with } q > 0 \\ 0 & x \notin \mathbb{Q} \end{cases}$$

- The **Cantor function** c has the weird staircase-looking graph shown in §1.1.

Theorems with names

Intermediate Value Theorem (IVT) If $f : \mathbb{R} \to \mathbb{R}$ is continuous, then for any y between $f(a)$ and $f(b)$, $\exists\, x$ between a and b such that $f(x) = y$.

Mean Value Theorem (MVT) Let $a < b$. Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous and differentiable. Then $\exists\, x \in (a, b)$ such that $f'(x) = \dfrac{f(b) - f(a)}{b - a}$.

OTHER THEOREMS TO REMEMBER

- Compositions of injections are injections.
- Compositions of surjections are surjections.
- Compositions of bijections are bijections.
- Subsets of finite sets are finite.
- Subsets of countable sets are countable.
- $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Z} \times \mathbb{Z}$ and $\mathbb{Q}$ are countable infinite sets.
- Countable unions of countable sets are countable.

STANDARD PROOF TECHNIQUES

To prove $E \subseteq F$**,** assume $x \in E$ and deduce $x \in F$.

To prove $E = F$**,** prove $E \subseteq F$ and $F \subseteq E$.

To prove $f : A \to B$ **is injective,** assume $f(x) = f(y)$ and deduce $x = y$.

To prove $f : A \to B$ **is surjective,** assume $y \in B$ and write down a formula for an $x \in A$ so that $f(x) = y$.

To prove $f : A \to B$ **is bijective,** do one of these two things:

1. Prove f is injective, and prove f is surjective.
2. Prove that f has an inverse, by writing down a formula for f^{-1} and showing $f^{-1} \circ f(x) = x$ and $f \circ f^{-1}(x) = x$.

1.6 Chapter 1 Homework

Exercises from Section 1.2

1. Let A, B and C be sets. Prove $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

   Hint: This is a set equality proof, so you should prove each side is a subset of the other.

2. Let A, B and C be sets. Prove $A - (B \cap C) = (A - B) \cup (A - C)$.

Exercises from Section 1.3

3. Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = 2x^2 + 3$, and let $E = \{x \in \mathbb{R} : x \leq 5\}$. Describe the sets $f(E)$ and $f^{-1}(E)$ using inequalities.

4. Let $f : A \to B$ be a function. Prove that for any set $E \subseteq A$, $E \subseteq f^{-1}(f(E))$.

   Hint: This is a subset proof. Start by letting $x \in E$. What does that mean about $f(x)$?

5. Give a specific example of a function $f : A \to B$ and a set $E \subseteq A$ such that $f^{-1}(f(E)) \neq E$.

6. Let $f : A \to B$ be a function. Prove that for any set $E \subseteq B$, $f(f^{-1}(E)) = E \cap Im(f)$.

   Hint: This is a set equality proof. For the ($\subseteq$) direction, let $x \in f(f^{-1}(E))$. Explain why $x \in Im(f)$ and $x \in E$. For the ($\supseteq$) direction, let $x \in E \cap Im(f)$. Since $x \in Im(f)$, $\exists w \in A$ s.t. $f(w) = x$. To what subset of A must w belong?

7. Give a specific example of a function $f : A \to B$ and a set $E \subseteq B$ with $f(f^{-1}(E)) \neq E$.

   Hint: Choose $E \subseteq B$ so that it includes some elements not in the image of f.

8. Prove or disprove: If $f : A \to B$ is a function, $E \subseteq A$ and $F \subseteq A$, then

$$f(E \cup F) = f(E) \cup f(F).$$

9. Prove or disprove: If $f : A \to B$ is a function, $E \subseteq A$ and $F \subseteq A$, then

$$f(E \cap F) = f(E) \cap f(F).$$

10. Prove or disprove: If $f : A \to B$ is a function, $E \subseteq B$ and $F \subseteq B$, then

$$f^{-1}(E \cap F) = f^{-1}(E) \cap f^{-1}(F).$$

11. Prove or disprove: If $f : A \to B$ is a function, $E \subseteq B$ and $F \subseteq B$, then

$$f^{-1}(E \cup F) = f^{-1}(E) \cup f^{-1}(F).$$

12. Let $A = \left\{x \in \mathbb{R} : x < \frac{1}{2}\right\}$; let $B = \{x \in \mathbb{R} : x \geq 0\}$ and let $C = \{x \in \mathbb{R} : x \geq 1\}$.

    a) Evaluate $\mathbb{1}_A(4)$.
    b) Evaluate $\mathbb{1}_C(5)$.
    c) Evaluate $\mathbb{1}_{A \cap B}\left(\frac{3}{4}\right)$.
    d) Evaluate $\mathbb{1}_C \circ \mathbb{1}_B(3)$.
    e) Evaluate $\mathbb{1}_C \circ \mathbb{1}_A(3)$.
    f) Describe the set $\mathbb{1}_A^{-1}(C)$.

13. Let A, B and C be as in Exercise 12.

    a) Sketch a graph of $\mathbb{1}_A$.
    b) Sketch a graph of $\mathbb{1}_{A\cup C}$.
    c) Sketch a graph of the function $f(x) = 4 \cdot \mathbb{1}_C(3x)$.
    d) Sketch a graph of the function $g(x) = 3 \cdot \mathbb{1}_A(x) + \mathbb{1}_B(x) - \mathbb{1}_C(x)$.

14. Determine, with proof, whether or not the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $f(x,y) = (2x - y, x + 2y)$ is injective.

15. Determine, with proof, whether or not the function $f :\to [-5,5] \to [0,5]$ defined by $f(x) = \sqrt{25 - x^2}$ is surjective.

16. Prove the first part of Theorem 1.17 from the notes, which says that if $f : B \to C$ and $g : A \to B$ are both surjective, then $f \circ g : A \to C$ is surjective.

17. Prove the second statement of Theorem 1.17, which says that if $f : B \to C$ and $g : A \to B$ are functions so that $f \circ g : A \to C$ is surjective, then f must be surjective.

18. Give an example of functions $f : B \to C$ and $g : A \to B$ where $f \circ g : A \to C$ is surjective, but g is <u>not</u> surjective.

19. Determine, with proof, whether or not the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 7x - 13$ is a bijection.

20. Determine, with proof, whether or not the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x|x|$ is a bijection.

21. Let $a, b, c, d \in \mathbb{R}$ with $ad \neq bc$. Prove that the function $h : \mathbb{R} - \left\{\frac{d}{c}\right\} \to \mathbb{R} - \left\{\frac{a}{c}\right\}$ defined by $h(x) = \dfrac{ax + b}{cx + d}$ is a bijection.

22. Let $f : \mathbb{R} - \{-1\} \to \mathbb{R} - \{0\}$ be $f(x) = \dfrac{2}{x^3 + 1}$. Compute a formula for $f^{-1}(x)$ and prove that your f^{-1} is indeed the inverse of f (by verifying that $f^{-1} \circ f(x) = x$ and $f \circ f^{-1}(x) = x$).

Exercises from Section 1.4

23. Prove that the union of two finite sets is finite.

    Hint: This is "obvious", but not so easy to prove. Start like this: let E and F be finite sets. Then, by definition of finite set, there exist injections $f_E : E \to \{1, ..., m\}$ and $f_F : F \to \{1, ..., n\}$, where $m, n \in \mathbb{N}$. Use these functions to construct an injection from $E \cup F$ to $\{1, ..., m + n\}$.

24. Let E be an infinite set and let x_1 be an arbitrary element of E. Prove that $E - \{x_1\}$ is infinite.

    Hint: Try a proof by contradiction that applies the result of Exercise 23.

25. Prove that every infinite set contains a countable infinite subset.

    Hint: Use the result of Problem 24 to select an $x_1 \in E$, then $x_2 \in E - \{x_1\}$, etc.

26. a) Give an example of a function $f : \mathbb{N} \to \mathbb{N}$ that is injective but not surjective.

    b) Prove that given any infinite set E, there is a function $f : E \to E$ which is injective but not surjective.

    Hint: Use the result of Problem 24 to find a countable infinite subset F of E. Using part (a) of this problem as a prototype, construct a function $f : F \to F$ which is injective but not surjective, and then extend the function f to the rest of E.

Chapter 2

The real numbers

The rational numbers $\mathbb{Q}$ and the IVT

Definition 2.1 *The set of* **rational numbers***, denoted $\mathbb{Q}$, is the set of quotients of integers:*

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

Strictly speaking, this definition isn't quite right, because it obscures the fact that some of these quotients are actually the same rational number. For example,

$$\frac{30}{60} = \frac{-8}{-16} = \frac{7}{14} = \frac{1}{2}.$$

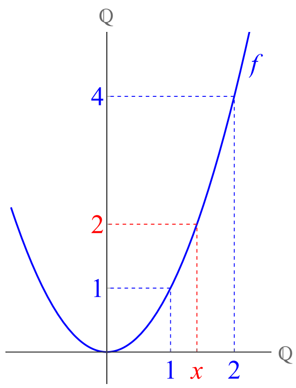
Technically, $\mathbb{Q}$ is a set of *equivalence classes* of pairs of integers. But we don't need that level of precision in our course, so we won't worry about it.

Theorem 2.2 (Intermediate Value Theorem (IVT)) *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous, and let $a < b$ in $\mathbb{R}$. Then, for any y between $f(a)$ and $f(b)$, there is $x \in \mathbb{R}$ between a and b such that $f(x) = y$.*

What would happen if we tried to formulate an Intermediate Value Theorem for $\mathbb{Q}$? It would look like this:

Conjecture 2.3 (Intermediate Value Theorem (IVT) for $\mathbb{Q}$) *Let $f : \mathbb{Q} \to \mathbb{Q}$ be continuous, and let $a < b$ in $\mathbb{Q}$. Then, for any y between $f(a)$ and $f(b)$, there is $x \in \mathbb{Q}$ between a and b such that $f(x) = y$.*

Consider the function $f : \mathbb{Q} \to \mathbb{Q}$ defined by $f(x) = x^2$. Let $a = 1$ and $b = 2$.



So if the IVT for $\mathbb{Q}$ is true, then there would have to be a rational number $x \in \mathbb{Q}$ such that $x^2 = 2$. But...

Theorem 2.4 (Hippasus' Theorem) *There is no $x \in \mathbb{Q}$ such that $x^2 = 2$.*

PROOF

What's going wrong here? There are two possibilities:

1.

2.

PUNCHLINE

$\mathbb{R}$ must be "different" from $\mathbb{Q}$ in some way. So our first goal in MATH 430 is to understand everything there is to know about $\mathbb{R}$, and especially what makes it "different" from $\mathbb{Q}$.

2.1 Algebraic and order properties

Rather than worrying about exactly what a real number is, let's focus on what properties the *set of real numbers* $\mathbb{R}$ has. First, there's a lot of nice arithmetic and algebra you can do with real numbers:

Definition 2.5 *A* **field** *is a set F together with two binary operations of F:*

Addition: $+ : F \times F \to F$ *defined by* $(x, y) \overset{+}{\longmapsto} x + y$*, and*

Multiplication: $\cdot : F \times F \to F$ *defined by* $(x, y) \overset{\cdot}{\longmapsto} xy$

which satisfy all the following eight rules:

1. Adding and multiplying produce elements of F:
$$\forall\, x, y \in F,\ x + y \in F \text{ and } xy \in F.$$
2. $+$ and $\cdot$ are commutative:
$$\forall\, x, y \in F,\ x + y = y + x \text{ and } xy = yx;$$
3. $+$ and $\cdot$ are associative:
$$\forall\, x, y, z \in F,\ x + (y + z) = (x + y) + z \text{ and } x(yz) = (xy)z;$$
4. $\cdot$ distributes over $+$:
$$\forall\, x, y, z \in F,\ x(y + z) = xy + xz;$$
5. there is an additive identity element:
$$\exists\, 0 \in F \text{ such that } 0 + x = x\ \forall\, x \in F;$$
6. there is a multiplicative identity element
$$\exists\, 1 \in F \text{ such that } 1x = x\ \forall\, x \in F;$$
7. additive inverses exist:
$$\forall\, x \in F, \exists\, -x \in F \text{ such that } x + (-x) = 0;$$
8. reciprocals of nonzero elements exist:
$$\forall\, x \neq 0 \text{ in } F, \exists\, x^{-1} \text{ such that } x(x^{-1}) = 1.$$

If you've had abstract algebra (MATH 420), there's a much shorter way to define a field that encompasses all eight of these properties:

Definition 2.6 (Shorter definition of field) *A* **field** *is a set F with two binary operations $+$ and $\cdot$ such that $(F, +)$ and $(F - \{0\}, \cdot)$ are abelian groups, and such that multiplication distributes over addition (i.e. $\forall\, x, y, z \in F, x(y + z) = xy + xz$).*

What either of these definitions capture is that a field is a set with two operations $+$ and $\cdot$ that have all the same "nice" rules as the addition and multiplication you learn about as a kid.

In particular, you can always add and multiply in a field (by definition), and you can also always subtract, because

and you can always divide by anything other than zero, because

and these operations behave lots of nice rules (you can do them in either order, regroup, there are identity and inverse elements, etc.).

In fact, there are lots of other nice rules that all fields automatically obey:

Theorem 2.7 (Arithmetic and algebraic properties of fields) *Let F be any field. Then, $\forall\, x, y, z \in F$, we have:*

1. Additive cancellation holds:
$$x + z = y + z \text{ implies } x = y.$$
2. Multiplicative cancellation holds:
$$xz = yz \text{ implies } x = y, \text{ so long as } z \neq 0.$$
3. The additive identity element of the field is unique.
4. The multiplicative identity element of the field is unique.
5. The additive inverse of each element of F is unique.
6. The reciprocal of each nonzero element of F is unique.
7. $0x = 0$.
8. $-x = (-1)x$.
9. *If $xy = 0$, then $x = 0$ or $y = 0$.*
10. $-0 = 0$.
11. $1^{-1} = 1$.
12. $-(-x) = x$.
13. $(x^{-1})^{-1} = x$.
14. $(-x)y = -(xy) = x(-y)$.

If this was an algebra class like MATH 420, it would be useful to go through proofs of all the properties in Theorem 2.7. But that's not really the subject matter of MATH 430. What's important for us is this:

Assumption # 1 about the real numbers

$\mathbb{R}$ is a field, with additive identity element 0 and multiplicative identity element 1.

OTHER EXAMPLES OF FIELDS

- $\mathbb{Q}$
- the set $\mathbb{C}$ of complex numbers (i.e. numbers of the form $a + bi$ where $a, b \in \mathbb{R}$);
- the set $\mathbb{Z}/p\mathbb{Z}$ (a.k.a. $\mathbb{Z}_p$) of integers mod p (if p is prime);
- the set $\mathbb{R}(x)$ of rational functions (a function is rational if it is the quotient of two polynomials).

SETS THAT ARE NOT FIELDS

- the set $\mathbb{N}$ of natural numbers (either $\{0, 1, 2, \dots\}$ or $\{1, 2, 3, \dots\}$ depending on the context);
- the set $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ of integers;
- the set $\mathbb{Z}/n\mathbb{Z}$ (a.k.a. $\mathbb{Z}_n$) of integers mod n (if n isn't prime);
- the set $\mathbb{R}[x]$ of polynomials with real coefficients.

Order properties of $\mathbb{R}$

First, let's talk about what a *relation* is. Given any set S, recall that $S^2 = S \times S$ is the set of ordered pairs where each entry is in S. Next, by a **relation** on a set S we technically mean any subset of S^2.

Think of a relation as a "symbol" that you put between two elements of S to produce a mathematical sentence that is either true or false. Examples of such symbols include:

The connection between the formal definition of relation and the way we usually use relations is as follows: if R is some subset of $S \times S$ (usually R is a symbol), we say $x\,R\,y$ if $(x, y) \in R$ and $x \not R\ y$ if $(x, y) \notin R$.

At this point, we care about the relation $\leq$, which has three important properties that make it something called a *total ordering*:

Definition 2.8 *Let S be any set. A* **total ordering** *on S is a relation $\leq$ on S such that $\forall\, x, y, z \in S$, the following properties hold:*

1. Connexity:
$$\textit{either } x \leq y \textit{ or } y \leq x.$$
2. Antisymmetry:
$$\textit{if } x \leq y \textit{ and } y \leq x, \textit{ then } x = y.$$
3. Transitivity:
$$\textit{if } x \leq y \textit{ and } y \leq z, \textit{ then } x \leq z.$$

In a set S that has a total ordering $\leq$, we automatically get another relation $<$. When we write $x < y$, we formally mean "$x \leq y$ and $x \neq y$".

Definition 2.9 *An* **ordered field** *is a field F, together with a total ordering $\leq$ on F such that $\forall\, x, y, z \in F$,*

1. Addition preserves inequalities:
$$\textit{if } x \leq y, \textit{ then } x + z \leq y + z.$$
2. Products of non-negative elements are non-negative:
$$\textit{if } x \geq 0 \textit{ and } y \geq 0, \textit{ then } xy \geq 0.$$

If F is an ordered field, we define these subsets of F:
1. *the* **positive cone** $F_+ = \{x \in F : x > 0\}$*; and*
2. *the* **negative cone** $F_- = \{x \in F : x < 0\}$.

Notice that F is the disjoint union of F_+, F_- and $\{0\}$.

Theorem 2.10 (Properties of ordered fields) *Let F be an ordered field. Then, for all $w, x, y, z \in F$:*
1. Additive inverses have opposite sign as the original number:
$$\textit{either } -x \leq 0 \leq x \textit{ or } x \leq 0 \leq -x.$$
2. Inequalities can be added:
$$\textit{if } w \leq x \textit{ and } y \leq z, \textit{ then } w + y \leq x + z.$$
3. Multiplying by positive constant preserves inequalities:
$$\textit{if } x \leq y \textit{ and } z \geq 0, \textit{ then } xz \leq yz.$$
4. Multiplying by negative constant reverses inequalities:
$$\textit{if } x \leq y \textit{ and } z \leq 0, \textit{ then } xz \geq yz.$$

5. Positive times negative is negative:

$$\text{if } y \geq 0 \text{ and } z \leq 0, \text{ then } yz \leq 0.$$

6. Negative times negative is positive:

$$\text{if } x \leq 0 \text{ and } z \leq 0, \text{ then } xz \geq 0.$$

7. Squares are non-negative:

$$x^2 \geq 0.$$

8. Reciprocal has same sign as original number:

$$\text{if } x > 0, \text{ then } x^{-1} > 0; \text{ if } x < 0, \text{ then } x^{-1} < 0.$$

9. Reciprocals reverse inequalities:

$$\text{if } 0 < x \leq y, \text{ then } x^{-1} \geq y^{-1}, \text{ and if } x \leq y < 0, \text{ then } x^{-1} \geq y^{-1}.$$

PROOF To prove (1), we use cases. By connexity of $\leq$, either $x \geq 0$ or $x \leq 0$.

Case 1: If $x \geq 0$, then $x + (-x) \geq 0 + (-x)$ so $0 \geq -x$. Thus $-x \leq 0 \leq x$ as wanted.

Case 2: If $x \leq 0$, then $x + (-x) \leq 0 + (-x)$ so $0 \leq -x$. Thus $x \leq 0 \leq -x$ as wanted.

To prove (2), note $w \leq x$ implies $w + y \leq x + y$ and $y \leq z$ implies $y + x \leq z + x$.
By transitivity, $w + y \leq x + z$.

To prove (3), assume $x \leq y$ and $z \geq 0$. Then $x + (-x) \leq y + (-x)$ so $0 \leq y - x$.
Since products of non-negative numbers are non-negative, $0 \leq (y-x)z = yz - xz$.
Add xz to both sides to get $xz \leq yz$.

To prove (4), assume $x \leq y$ and $z \leq 0$.
By (1), $(-z) \geq 0$.
As in the proof of (3), $0 \leq y - x$.
Since products of non-negatives are non-negative, $0 \leq (y - x)(-z) = -yz + xz$.
Add yz to both sides to get $yz \leq xz$.

Statement (5) follows from statement (4) by setting $x = 0$.

Statement (6) follows from statement (4) by setting $y = 0$.

Statement (7) follows from statement (6) and the second axiom in the definition of ordered field.

To prove (8), first observe that by (7), 1 must be positive, since $1 = 1^2$.
Now, suppose not (i.e. x and x^{-1} have opposite signs).
Then, by statement (6), $xx^{-1} = 1$ would have to be negative.
This is a contradiction.

To prove (9), suppose that $0 < x \leq y$. By (8), we know $0 < \dfrac{1}{x}$ and $0 < \dfrac{1}{y}$.

Apply (3) to $0 < x \leq y$ by multiplying everything through by $\dfrac{1}{x}$ to get $0 < 1 \leq \dfrac{y}{x}$.

Then multiply through by $\dfrac{1}{y}$ to get $0 < \dfrac{1}{y}\left(\dfrac{1}{x}\right)x \leq \dfrac{1}{y}\left(\dfrac{1}{x}\right)y$, i.e. $0 < \dfrac{1}{y} \leq \dfrac{1}{x}$. □

Assumption #2 about the real numbers

$\mathbb{R}$ is an ordered field.

Quick remark: Elements of the positive cone of the real numbers, i.e. those real numbers x which satisfy $x > 0$, are called **positive numbers**. Elements of the negative cone are called **negative numbers**.

0 is neither positive nor negative.

If we want to refer to the numbers that are ≥ 0, we call those **non-negative numbers**. "Positive" and "non-negative" are **NOT** synonyms.

OTHER EXAMPLES OF ORDERED FIELDS

- $\mathbb{Q}$
- There are others, but they are complicated.

FIELDS THAT ARE NOT ORDERED FIELDS

- $\mathbb{C}$ (reason: $i^2 = -1 < 0$, violating (9) of Theorem 2.10)
- $\mathbb{R}(x)$, the set of rational functions (reason: which is bigger, $\frac{1}{x}$ or x?)
- There are others, but they are complicated.

Intervals

In an ordered field like $\mathbb{R}$, we can define special subsets called *intervals*:

Definition 2.11 *Given* $a, b \in \mathbb{R}$, *set*

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

Any subset of $\mathbb{R}$ *which is any of these types is called an* **interval**.

As a convention, the set of all real numbers $\mathbb{R}$ *is also decreed to be an interval and can be written* $(-\infty, \infty)$.

$\mathbb{R}$

EXERCISE

What is the interval $[5, 3]$? What is the interval $[4, 4]$? What about $(4, 4)$?

Solution: $[5, 3] =$

$[4, 4] =$

$(4, 4) =$

2.2 Absolute value and distance

Definition 2.12 *The* **absolute value** *of a real number* z *is*

$$|z| = \begin{cases} z & \text{if } z \geq 0 \\ -z & \text{if } z \leq 0 \end{cases}$$

Theorem 2.13 (Elementary properties of absolute value) *Let* $z \in \mathbb{R}$. *Then:*

- $|z| = \max\{-z, z\}$.
- $-|z| \leq z \leq |z|$.
- $|z| = |-z|$.
- $|z| \geq 0$.
- $|z| = 0$ *only if* $z = 0$.

PROOF These are pretty self-evident, so I won't prove all of them. However, to be pedantic I will prove the second statement so you see how the precise arguments work. To do this, consider two cases, depending on whether or not $z \geq 0$:

Case 1: $z \geq 0$. In this situation, $|z| = z$. Therefore

$$-|z| = -z \leq 0 \leq z = |z|,$$

establishing the second statement.

Case 2: $z < 0$. In this situation, $|z| = -z$. Therefore

$$-|z| = -(-z) = z < 0 < -z = |z|,$$

proving the second statement. □

Theorem 2.14 (Multiplicativity of absolute value) *Let $x, y \in \mathbb{R}$. Then*

$$|xy| = |x|\,|y|.$$

PROOF HW

Hint: Consider some cases depending on the signs of x and/or y.

The distance between two real numbers

Definition 2.15 *Let $x, y \in \mathbb{R}$. The* **distance** *between x and y is $|x - y|$.*

If you think about the real numbers as being points on a number line, $|x-y|$ literally gives the distance between x and y:

Theorem 2.16 (Elementary properties of distance) *Let $x, y, r \in \mathbb{R}$. Then:*

1. Distances are nonnegative:
$$|x - x| \geq 0.$$
2. Distance is definite:
$$|x - y| = 0 \text{ if and only if } x = y.$$
3. Distances are symmetric:
$$|x - y| = |y - x|.$$
4. Distance is multiplicative:
$$|rx - ry| = |r|\,|x - y|.$$

PROOF Let $z = x - y$, then all these follow from the elementary properties of absolute value. For example, to establish that distances are symmetric,

$$|x - y| = |z| = 1\,|z| = |-1|\,|z| = |-z| = |-(x - y)| = |y - x|.$$

Proofs of the other properties are similar. □

The triangle inequality

A simple idea we will repeatedly use in this course is what is called the *Triangle Inequality*. It goes like this:

Theorem 2.17 (Triangle Inequality) *If $x, y \in \mathbb{R}$, then*

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

PROOF There are two cases, depending on whether or not $x + y \geq 0$.

Case 1: If $x + y \geq 0$, we have

$$|x + y| = x + y \leq |x| + y \leq |x| + |y|.$$

Case 2: If $x + y < 0$, we have

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

This completes the proof. □

If we start with the triangle inequality

$$|x + y| \leq |x| + |y|$$

and let $x = a - b$ and let $y = b - c$, then we get

This alternate version of the triangle inequality is good to know:

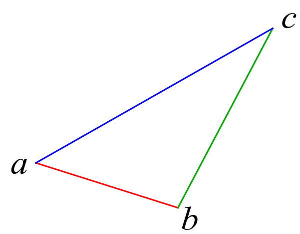
Theorem 2.18 (Triangle Inequality (version 2)) *If $a, b, c \in \mathbb{R}$, then*

$$|a - c| \le |a - b| + |b - c|.$$

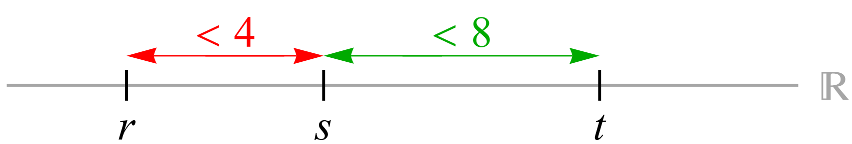
In other words,

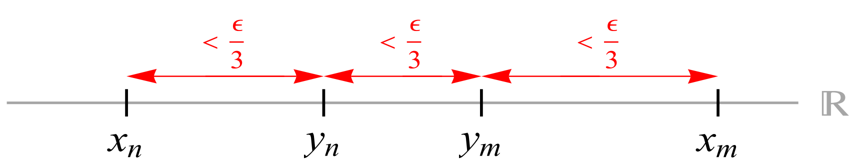
$$\text{the distance from } a \text{ to } c \le \left(\text{the distance from } a \text{ to } b\right) + \left(\text{the distance from } b \text{ to } c\right)$$

This version explains why we call these facts the "Triangle Inequality". If you think of a, b and c as points in space instead of numbers, you get the following picture:



Often, we think of points on a number line, and reason like this:





Notice that the order in which the points r, s, t (or x_n, y_n, y_m, x_m) appear on the number line doesn't affect this logic.

Bounded sets

Definition 2.19 *A subset $S \subseteq \mathbb{R}$ is called* **bounded** *if $\exists\, B > 0$ such that $\forall\, x \in S$, $|x| \le B$. In this case B is called a* **bound** *for S. An* **unbounded** *set is one that is not bounded.*

Examples of bounded sets

Examples of unbounded sets

2.3 Sequences; convergence and divergence

AN UPDATE ON THE BIG PICTURE

We are investigating what makes $\mathbb{R}$ different from $\mathbb{Q}$.

Both $\mathbb{Q}$ and $\mathbb{R}$ have the same formulas for absolute value and distance, so these concepts don't (by themselves) distinguish $\mathbb{R}$ from $\mathbb{Q}$.

BUT: we will find a big difference between $\mathbb{R}$ and $\mathbb{Q}$ by further investigating ideas related to absolute value and distance. The new ideas involve the *convergence* of *sequences* of numbers.

Definition 2.20 *Let F be either $\mathbb{Q}$ or $\mathbb{R}$, and let $m \in \mathbb{Z}$.*

A **sequence** *(in F) is a function $\{m, m+1, m+2, \ldots\} \to F$.*

If we write x_n for the image of n under this function (which might ordinarily be denoted $x(n)$), the entire sequence is denoted $\{x_n\}$ or $\{x_n\}_{n=m}^{\infty}$, so in particular,

$$\{x_n\} = \{x_m, x_{m+1}, x_{m+2}, x_{m+3}, \ldots\}.$$

The variable n is called the **index** *of the sequence $\{x_n\}$.*

Note 1: sequences have infinite index sets:

$\{1, 2, 3, 4, 5\}$ is **not** a sequence.

Note 2: sequences are ordered:

$\{1, 2, 3, 4, 5, \ldots\}$ is not the same sequence as $\{1, 3, 2, 4, 5, 7, 6, 8, \ldots\}$, even though these objects are the same sets.

Note 3: "x_n" versus "$\{x_n\}$"

Definition 2.21 *A sequence $\{x_n\}$ is called...*

... **increasing** *if $x_n \le x_{n+1}$ for every n;*

... **decreasing** *if $x_n \ge x_{n+1}$ for every n;*

... **strictly increasing** *if $x_n < x_{n+1}$ for every n;*

... **strictly decreasing** *if $x_n > x_{n+1}$ for every n;*

... **monotone** *if either $\{x_n\}$ is increasing or $\{x_n\}$ is decreasing;*

... **strictly monotone** *if $\{x_n\}$ is strictly increasing or $\{x_n\}$ is strictly decreasing;*

... **bounded** *if $\{x_n\}$ is a bounded subset of $\mathbb{R}$.*

MAIN EXAMPLES

Remark: Throughout this chapter, we will be using the four sequences in this example as "prototypes".

So you should remember the definitions of the four sequences $\{a_n\}$, $\{b_n\}$, $\{c_n\}$ and $\{d_n\}$ in this example as we go forward.

For now, let's determine which, if any, of the adjectives in Definition 2.21 apply to each sequence.

1. $a_n = \frac{1}{n}$ (i.e. $\{a_n\} = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right\}$)

2. $b_n = 1 + (-1)^n$

3. $c_n = n^2$

4. $\{d_n\} = \{1, 1.4, 1.41, 1.414, 1.4142, \ldots\}$
   (formally, d_n is the largest decimal with $\leq n$ places whose square is at most 2)

Question

What happens to the terms of a sequence, as its index n as n gets larger and larger (without bound)? In other words, what is the *limiting behavior* of the sequence?

	Q1: Do the numbers in the sequence get closer and closer to a single number?	**Q2:** Do the numbers in the sequence get closer and closer to each other?
$a_n = \frac{1}{n}$		
$b_n = 1 + (-1)^n$		
$c_n = n^2$		
$\{d_n\} = \{1.4, 1.41, 1.414, \ldots\}$		

Main Examples

The big difference between $\mathbb{R}$ and $\mathbb{Q}$ is that for sequences in $\mathbb{R}$, the answers to **Q1** and **Q2** above always coincide.

But in $\mathbb{Q}$, there are sequences like $\{d_n\}$ which get closer and closer to each other and therefore "should" have a limit, but don't (at least not in $\mathbb{Q}$).

More on this later–for now, let's talk about exactly what **Q1** means in detail.

Definition 2.22 *Let F denote either $\mathbb{Q}$ or $\mathbb{R}$.*

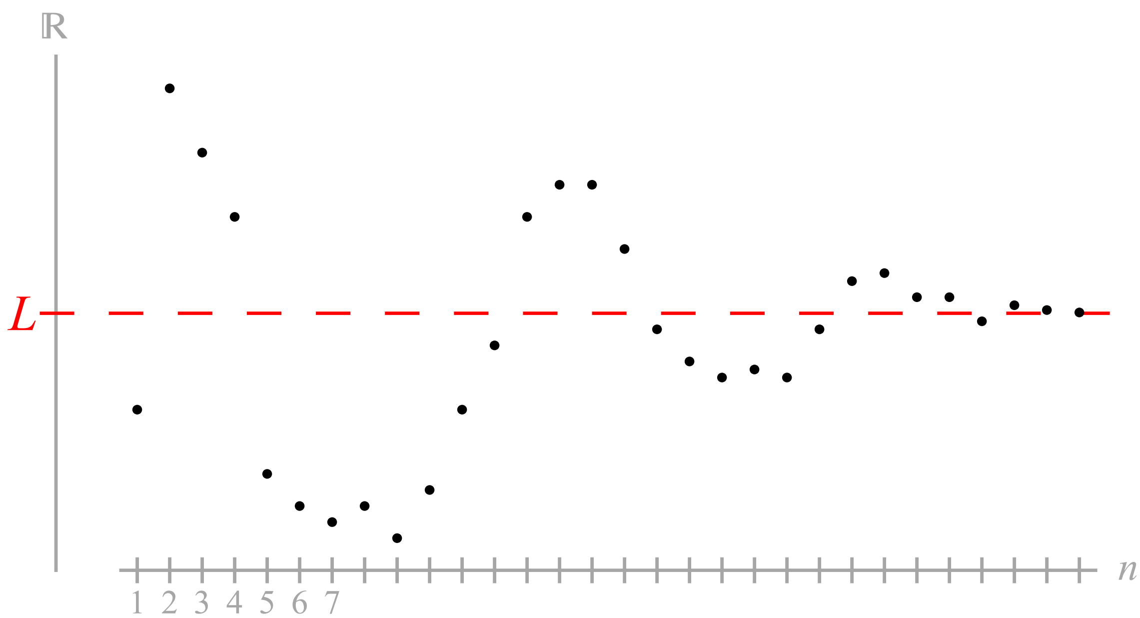
A sequence $\{x_n\}$ in F is said to **converge (in F)** *to a number $L \in F$ if $\forall\, \epsilon > 0$ ($\epsilon \in F$), $\exists\, N$ (this N usually depends on ϵ, so it might be denoted $N(\epsilon)$) such that if $n \geq N$, $|x_n - L| < \epsilon$.*

We write $\lim_{n\to\infty} x_n = L$ or $\lim x_n = L$ or $x_n \xrightarrow{n\to\infty} L$ or just $\{x_n\} \to L$ or $\boxed{x_n \to L}$ to express this.

In this situation, L is called a **limit** *of the sequence $\{x_n\}$.*

A sequence $\{x_n\}$ is said to **diverge** *if it does not converge to any limit.*

A PICTURE TO EXPLAIN THIS DEFINITION



Remarks on the definition of convergence

- If, given $\epsilon > 0$, a certain value of N works as $N(\epsilon)$ in the definition of convergence of $\{x_n\}$, then any number larger than N also works as $N(\epsilon)$.
- For any M, altering the first M terms of a sequence doesn't affect its convergence (reason: you can always choose $N \geq M$, based on the previous comment).

  So if you know $x_n \to L$, then even if you change (or delete) the values of $x_1, x_2, \ldots, x_{100}$, $\{x_n\}$ still converges to L.
- When you know a sequence converges, expressions (so long as they are > 0) can play the role of ϵ in the definition of convergence, i.e.

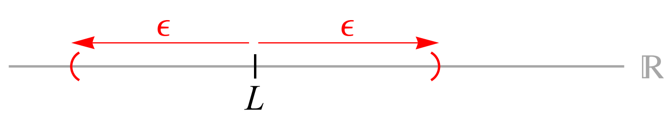
  $$\text{if } x_n \to L, \text{ then “}\forall\, \epsilon > 0\; \exists\, N \text{ so that } n \geq N \text{ implies } |x_n - L| < \frac{\epsilon}{2}.\text{”}$$

  and

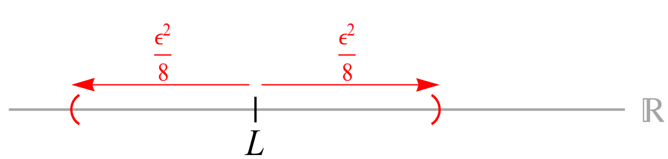
  $$\text{if } x_n \to L, \text{ then “}\forall\, \epsilon > 0\; \exists\, N \text{ so that } n \geq N \text{ implies } |a_n - L| < \frac{\epsilon^4}{100}.\text{”}$$

  but you can't say

  $$\text{if } a_n \to L, \text{ then “}\forall\, \epsilon > 0\; \exists\, N \text{ so that } n \geq N \text{ implies } |a_n - L| < \epsilon - 1.\text{”}$$
- If you know $x_n \to L$, then for any $\epsilon > 0$, you can turn the picture on the previous page “sideways” and think of a number line like this:



or more generally, something like this:



EXAMPLE

Prove that the sequence $\{a_n\}$ converges, where $a_n = \dfrac{1}{n}$.

Unfortunately, there is a problem with this proof (that is very hard to find).

Definition 2.23 *An ordered field F containing $\mathbb{N}$ is called* **Archimedean** *if, for every $x \in F$, there is $N \in \mathbb{N}$ so that $N > x$.*

Assumption #3 about the real numbers

$\mathbb{R}$ is Archimedean.

The example we did on the previous page (proving $\frac{1}{n} \to 0$) is our first example of what I call an *epsilon proof* (or an ϵ*-proof*). These proofs have the following flavor:

- You let $\epsilon > 0$.
- Based on some given information, this ϵ may tell you some stuff that is true, or provide you with some constants like M or M_1 or N_0 or N_1 or δ_0 or η or γ for which "something" is true.
- Then, you may have to choose something like an N or δ or L, either coming from the constants you get in the previous item, or based on some independent reasoning.
- You work out something and show that it is less than ϵ. (In the context of proving a sequence $\{x_n\}$ converges, this means you are working out $|x_n - L|$ for $n \geq N$.)
- Last you draw a conclusion based on the fact that the expression worked out to be less than ϵ. This may be that a sequence converges, or that a function is continuous, etc.

The main class of proofs we learn how to write in MATH 430 are ϵ-proofs.

Convergence and divergence viewed as a two-player game

A nice way to think about convergence of a sequence is as a two-player game.

Player 1 is the ϵ player, and Player 2 is the N player.

Player 1 goes first and chooses a positive number ϵ.

Then Player 2 chooses an N.

If, for every $n \geq N$, $|a_n - L| < \epsilon$, Player 2 wins; otherwise, Player 1 wins.

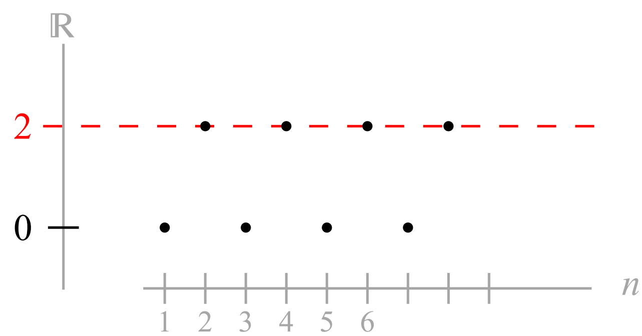
To say that the sequence converges means that Player 2 can always win.

A proof that the sequence converges is essentially a description of a strategy that Player 2 can use to win, no matter what Player 1 does (in other words, that accounts for all choices of ϵ that the first player might make).

A proof that the sequence doesn't converge to L is a description of a strategy that Player 1 can use to win, meaning

EXAMPLE

Let $b_n = 1 + (-1)^n$. Determine whether or not $b_n \to 2$.



REMARK / QUESTION

The preceding argument proves that $b_n \not\to 2$ (by finding one particular $\epsilon > 0$ such that there is no corresponding N).

But this doesn't rule out that $b_n \to$ something else.

How would you might that a sequence diverges?

One last remark

When you're *reading* an ϵ proof, choices of constants like the N or δ being made in the proof can seem like "magic".

They aren't magic–they come from scratch work that was done first, and that isn't included in the proof.

When you read an ϵ-proof, try to think about the scratch work that was done to create the argument.

Exercise

Let $x_n = \frac{n^2 - 1}{n^2 + 1}$. Prove that $\{x_n\}$ converges.

EXERCISE

Let $x_n = \sqrt{n+1} - \sqrt{n}$. Prove that $x_n \to 0$.

Properties of convergent sequences and limits

Theorem 2.24 (Limits preserve constants) *If $x_n = c$ for all n, then $x_n \to c$.*

PROOF Let $\epsilon > 0$. Given this ϵ, choose $N = \square$. Then for $n \geq N$, we have

$$|x_n - c| = |c - c| = 0 < \epsilon.$$

Thus $x_n \to c$ by definition. □

Theorem 2.25 (Limits are unique) *A convergent sequence can have at most one limit.*

PROOF We will prove this by contradiction.

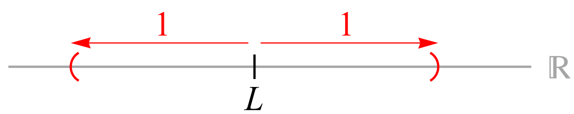
Suppose $\{x_n\}$ is a sequence with two limits L and M, where $L \neq M$.

WLOG $\square$.

Theorem 2.26 (Convergent sequences are bounded) *A convergent sequence must be bounded.*

PROOF Suppose $x_n \to L$.

Let $\epsilon = 1$. Then, $\exists\, N$ so that $n \geq N$ implies $|x_n - L| < 1$.



That means that when $n \geq N$, $x_n < L + 1$ so $|x_n| < |L + 1| \leq |L| + 1$.

Now, let

$$B =$$

It is clear that $|x_n| \leq B$ for every n. Therefore $\{x_n\}$ is bounded. □

CONSEQUENCE

Let $c_n = n^2$. $\{c_n\}$ diverges (because it is unbounded).

Main Limit Theorem

The Main Limit Theorem says that limits of sequences are preserved under arithmetic.

Theorem 2.27 (Main Limit Theorem) *Suppose $\{x_n\}$ and $\{y_n\}$ are convergent sequences. Then:*

1. *For any constant r, $\{r\, x_n\}$ converges, and $\lim\,(r\, x_n) = r\ (\lim x_n)$;*
2. *$\{x_n + y_n\}$ converges, and $\lim(x_n + y_n) = \lim x_n + \lim y_n$;*
3. *$\{x_n - y_n\}$ converges, and $\lim(x_n - y_n) = \lim x_n - \lim y_n$;*
4. *$\{x_n^2\}$ converges, and $\lim x_n^2 = (\lim x_n)^2$;*
5. *$\{x_n y_n\}$ converges, and $\lim\,(x_n y_n) = (\lim x_n)\,(\lim y_n)$;*
6. *if $M \neq 0$, then $\left\{\dfrac{x_n}{y_n}\right\}$ converges, and $\lim \left(\dfrac{x_n}{y_n}\right) = \dfrac{\lim x_n}{\lim y_n}$.*

PROOF Throughout this proof, let $L_x = \lim x_n$ and $L_y = \lim y_n$.

To prove statement (1), let $\epsilon > 0$.

Case 1: if $r = 0$, then $\{rx_n\} = \{0\}$ which converges to $0 = 0L_x = rL_x$ since limits preserve constants.

Case 2: if $r \neq 0$, since $x_n \to L_x$,

For $n \geq N$, we have

$$|(r\,x_n) - (rL)| = |r(x_n - L)| = |r|\,|x_n - L| <$$

Therefore $cx_n \to rL_x$ by definition of convergence.

To prove statement (2), let $\epsilon > 0$.

Since $x_n \to L_x$,

Since $y_n \to L_y$,

Now, let $N =$

If $n \geq N$, then

$$|(x_n + y_n) - (L_x + L_y)| =$$

For statement (3), observe $x_n + (-1)y_n = x_n - y_n$.

So by statements (1) and (2), $x_n - y_n \to L_x + (-1)L_y = L_x - L_y$.

For statement (4), let $\epsilon > 0$.

Since $\{x_n\}$ converges, $\{x_n\}$ is bounded, i.e $\exists\, B_x$ s.t. $|x_n| \leq B_x$ for all n.

Also, there is N such that whenever $n \geq N$, $|x_n - L| < \dfrac{\epsilon}{B_x + L_x}$.
For $n \geq N$, we have

$$\begin{aligned} |x_n^2 - L_x^2| &= |(x_n - L_x)(x_n + L_x)| \\ &= |x_n - L_x|\,|x_n + L_x| \\ &\leq |x_n - L_x|\,(|x_n| + L_x) \\ &< \frac{\epsilon}{B_x + L_x}(B_x + L_x) = \epsilon. \end{aligned}$$

Thus $x_n^2 \to L_x^2$ by definition.

For statement (5), observe that if you FOIL it out,

$$\frac{1}{4}(x_n + y_n)^2 - \frac{1}{4}(x_n - y_n)^2 = x_n y_n.$$

So by statements (2), (3) and (4),

$$x_n y_n \to \frac{1}{4}(L_x + L_y)^2 - \frac{1}{4}(L_{x-y})^2 = L_x L_y.$$

Last, for statement (6), let's start by proving $\dfrac{1}{y_n} \to \dfrac{1}{L_y}$.
Assume for now that $L_y > 0$ (we'll take care of the situation where $L_y < 0$ in a minute). Let $\epsilon > 0$.
Since $L_y \neq 0$, there is N_1 such that if $n \geq N_1$,

$$|y_n - L_y| < \frac{L_y}{2},$$

meaning that for $n \geq N_1$, $y_n > \dfrac{L_y}{2} > 0$.
Thus for $n \geq N_1$, $0 < \dfrac{1}{y_n} < \dfrac{2}{L_y}$.
Furthermore, there is N_2 such that if $n \geq N_2$,

$$|y_n - L_y| < \frac{\epsilon}{2} L_y^2.$$

Let $N = \max\{N_1, N_2\}$. For $n \geq N$,

$$\left|\frac{1}{y_n} - \frac{1}{L_y}\right| = \left|\frac{L_y}{y_n L_y} - \frac{y_n}{y_n L_y}\right| = \frac{|L_y - y_n|}{y_n L_y} = \frac{|y_n - L_y|}{L_y}\frac{1}{y_n} < \frac{\frac{\epsilon}{2} L_y^2}{L_y}\left(\frac{2}{L_y}\right) = \epsilon.$$

Therefore $\dfrac{1}{b_n} \to \dfrac{1}{M}$ as wanted.

If $L_y < 0$, then $-y_n \to -L_y$. Since $-L_y > 0$, we can apply the previous argument to $\{-y_n\}$ to conclude $\dfrac{1}{-y_n} \to \dfrac{1}{-L_y}$. Thus $\dfrac{1}{y_n} \to \dfrac{1}{L_y}$ by statement (1).

To finish up statement (6), observe $\dfrac{x_n}{y_n} = x_n\left(\dfrac{1}{y_n}\right)$, so by statement (5) and what we just proved,

$$\frac{x_n}{y_n} \to L_x\left(\frac{1}{L_y}\right) = \frac{L_x}{L_y}. \square$$

Theorem 2.28 (Limits preserve soft inequalities) *Let $\{x_n\}$ and $\{y_n\}$ be convergent sequences. If there exists N so that $x_n \le y_n$ for all $n \ge N$, then $\lim x_n \le \lim y_n$.*

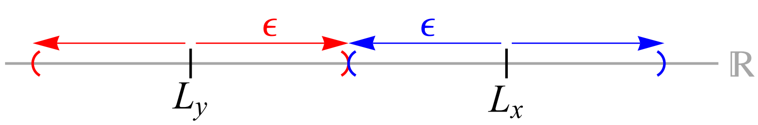
PROOF Let $L_x = \lim x_n$ and let $L_y = \lim y_n$.

Suppose not, i.e. $L_x > L_y$.

Now, let $\epsilon = \dfrac{1}{2}(L_x - L_y)$.

Then since $x_n \to L_x$ and $y_n \to L_y$, there exist N_x and N_y such that

$$n \ge N_x \Rightarrow |x_n - L_x| < \epsilon \Rightarrow x_n > L_x - \epsilon;$$
$$n \ge N_y \Rightarrow |y_n - L_y| < \epsilon \Rightarrow y_n < L_y + \epsilon.$$



For $n \ge \max\{N_x, N_y, N\}$, we have

$$x_n > L_x - \epsilon = L_x - \frac{1}{2}(L_x - L_y) = \frac{1}{2}L_x + \frac{1}{2}L_y = L_y + \frac{1}{2}(L_x - L_y) = L_y + \epsilon > y_n,$$

contradicting $x_n \le y_n$.

This proves the result by contradiction. $\square$

Squeeze Theorems

Theorem 2.29 (Squeeze Theorem (version 1)) *Suppose that $\{x_n\}$, $\{y_n\}$ and $\{z_n\}$ are sequences such that $x_n \le y_n \le z_n$ for all n. If $\lim x_n = \lim z_n = L$, then $y_n \to L$.*

PROOF HW

Hint: If we knew $\{y_n\}$ converged, this is immediate from Theorem 2.28.
But we aren't assuming $\{y_n\}$ converges–we have to prove this with an ϵ-proof.

The next version of the Squeeze Theorem is often more useful in proofs.

Theorem 2.30 (Squeeze Theorem (version 2)) *Suppose $\{x_n\}$ is a sequence and there exists N so that $|x_n - L| \le a_n$ for all $n \ge N$. If $a_n \to 0$, then $x_n \to L$.*

PROOF Let $d_n = |x_n - L|$. By hypothesis, we have

$$0 \le d_n \le a_n$$

so by (version 1 of) the Squeeze Theorem, ☐. That means that given any $\epsilon > 0$, there is ☐ so that ☐ implies

$$|d_n - 0| < \epsilon$$

By definition of convergence, $x_n \to L$. □

CONSEQUENCE

- Suppose $\{z_n\}$ is some sequence with $|z_n - r| \le \dfrac{1}{n}$. We can immediately conclude that ☐.
- Suppose $\{w_n\}$ is some sequence with $|w_n| \le \dfrac{3}{n^2}$. We can immediately conclude that ☐.

EXAMPLE

Let $y_n = 2^{-n} = \frac{1}{2^n} = \left(\frac{1}{2}\right)^n$. Prove $y_n \to 0$.

PROOF The proof relies on the following claim:

Claim: For $n \geq 2$, $\left(\frac{1}{2}\right)^n \leq \frac{1}{n}$.
Assuming this claim is true, we have

$$|y_n - 0| = |y_n| = y_n = \left(\frac{1}{2}\right)^n \leq \frac{1}{n},$$

and since $\frac{1}{n} \to 0$, it follows from the Squeeze Theorem that $y_n \to 0$.

It remains to prove the claim. To do this, first note that for any $k \geq 2$, $\frac{1}{k} \leq \frac{1}{2}$ so

$$\frac{k-1}{k} = 1 - \frac{1}{k} \geq 1 - \frac{1}{2} = \frac{1}{2}.$$

Therefore

CONSEQUENCE

For any $r \geq 2$, $r^{-n} \leq 2^{-n}$, so by the Squeeze Theorem $r^{-n} = \left(\frac{1}{r}\right)^n \to 0$.

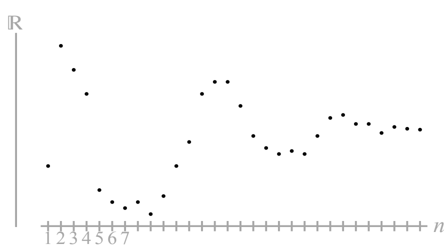
2.4 Cauchy sequences

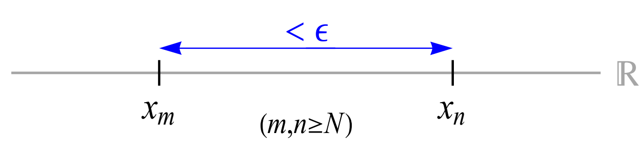
We now want to address the second question in our chart we had earlier:

	Q1: Do the numbers in the sequence get closer and closer to a single number? *(Does the sequence converge?*	**Q2:** Do the numbers in the sequence get closer and closer to each other?
$a_n = \frac{1}{n}$	**YES** ($a_n \to 0$)	
$b_n = 1 + (-1)^n$	We think NO.	
$c_n = n^2$	**NO** (unbounded)	
$\{d_n\} = \{1.4, 1.41, 1.414, \dots\}$	We think YES in $\mathbb{R}$ but NO in $\mathbb{Q}$.	

Definition 2.31 *Let F be either $\mathbb{Q}$ or $\mathbb{R}$, and let $\{x_n\}$ be a sequence in F. $\{x_n\}$ is called a* **Cauchy** *sequence (in F) if for every $\epsilon > 0$ ($\epsilon \in F$), there is N such that if $m, n \geq N$, then $|x_m - x_n| < \epsilon$.*

Idea: In a Cauchy sequence, the numbers in the sequence are getting closer and closer to each other. Notice that this is an intrinsic property of the sequence (it makes no reference to any limit of the sequence).





Theorem 2.32 (Convergent sequences are Cauchy) *If $\{x_n\}$ converges, then $\{x_n\}$ is Cauchy.*

PROOF Let $L = \lim x_n$. Let $\epsilon > 0$. There is N such that

$$n \geq N \Rightarrow |x_n - L| <$$

To show $\{x_n\}$ is Cauchy, suppose $m, n \geq N$. Then

$$|x_m - x_n|$$

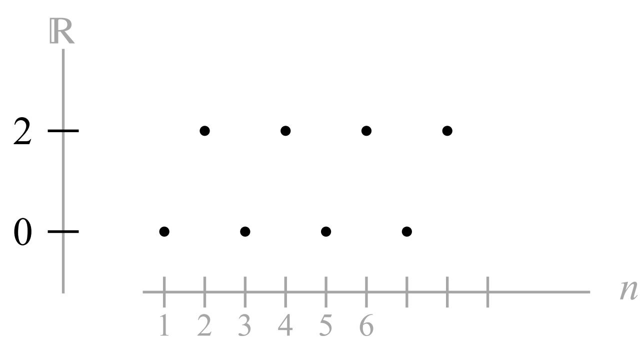
So $\{x_n\}$ is Cauchy, as wanted. □

CONSEQUENCE

For $a_n = \dfrac{1}{n}$, $\{a_n\}$ is Cauchy (since $\{a_n\}$ converges).

EXAMPLE

Let $b_n = 1 + (-1)^n$. Prove that $\{b_n\}$ is not a Cauchy sequence.



CONSEQUENCE

$\{b_n\}$ diverges (if it converged, it would be Cauchy).

QUESTION

We know every convergent sequence is Cauchy. What about the converse of this:

Does every Cauchy sequence converge?

If so, then "Cauchy sequences" and "convergent sequences" are exactly the same things. Let's consider this question in the context of one of our examples:

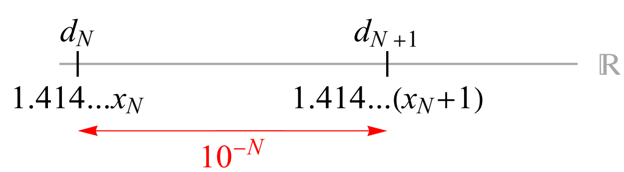
EXAMPLE

Let d_n be the largest rational number which can be written with at most n decimal places whose square is less than or equal to 2. Prove $\{d_n\}$ is a Cauchy sequence.

(Recall that $\{d_n\} = \{1, 1.4, 1.41, 1.414, \ldots\}$, so $\{d_n\}$ is an increasing sequence.)

SCRATCH WORK

More generally, for any N, we see that for all $n \geq N$, $d_N \leq d_n \leq d_N + 10^{-N}$:



PROOF

EXAMPLE, CONTINUED

Let d_n be the largest rational number which can be written with at most n decimal places whose square is less than or equal to 2. (Recall $\{d_n\} = \{1, 1.4, 1.41, 1.414, \ldots\}$.) **Does $\{d_n\}$ converge in $\mathbb{Q}$?**

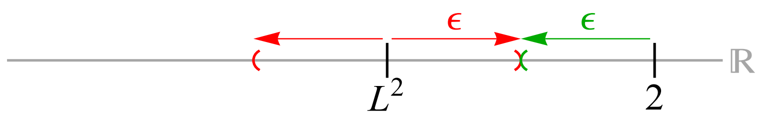
Solution: Suppose that $\{d_n\}$ does converge in $\mathbb{Q}$.

That means $d_n \to L$, where L is a rational number.

Therefore $d_n^2 \to L^2$, and since $d_n^2 \le 2$ for all n, $L^2 \le 2$ since limits preserve $\le$.

Since $L \in \mathbb{Q}$, $L^2 \ne 2$ (Hippasus), so it must be the case that $L^2 < 2$.

Now, choose $\epsilon = \frac{1}{2}(2 - L^2)$, which is positive.



Since $d_n^2 \to L$, $\exists\, N$ such that if $n \ge N$, then $|d_n^2 - L^2| < \epsilon$.

However, let's choose n so that $n \ge N$ and $10^{-n} < \frac{\epsilon}{5}$.

(This is doable since we know $10^{-n} \to 0$.)

Next, let's do this estimate, whose purpose will be seen in a minute:

$$\begin{aligned}
\boxed{\left|(d_n + 10^{-n})^2 - L^2\right|} &= \left|L^2 - (d_n + 10^{-n})^2\right| \\
&= \left|L^2 - d_n^2 - 2 \cdot d_n 10^{-n} + 10^{-2n}\right| \\
&\le \left|L^2 - d_n^2\right| + 10^{-n}\left|2d_n + 10^{-n}\right| \\
&< \epsilon + 10^{-n}\left|2d_n + 10^{-n}\right| \\
&\le \epsilon + 10^{-n}\left|4 + 1\right| \\
&= \epsilon + 10^{-n} \cdot 5 \\
&\boxed{<} \epsilon + \left(\frac{\epsilon}{5}\right) 5 = \boxed{2\epsilon}.
\end{aligned}$$

We have shown $|(d_n + 10^{-n})^2 - L^2| < 2\epsilon$, so

$$(d_n + 10^{-n})^2 < L^2 + 2\epsilon = L^2 + 2\left(\frac{2 - L^2}{2}\right) = 2. \tag{2.1}$$

This is a contradiction! d_n is supposed to be the largest number with $\le n$ decimal places whose square is at most 2, but $d_n + 10^{-n}$ is also a number with $\le n$ decimal places whose square, by (2.1) above, is less than 2.

Therefore $\{d_n\}$ cannot converge in $\mathbb{Q}$.

CONSEQUENCE

We've found a sequence $\{d_n\}$ which is Cauchy but does not converge in $\mathbb{Q}$. **Does this sequence $\{d_n\}$ converge in $\mathbb{R}$?**

Turns out, the answer is YES, but it's not because of anything about that sequence, it is because of an assumption we make about $\mathbb{R}$ (that isn't true in $\mathbb{Q}$):

Definition 2.33 *An ordered field F is called* **complete** *if every Cauchy sequence in F converges to a limit which is in F.*

Assumption #4 about the real numbers

$\mathbb{R}$ is complete.

Note: $\mathbb{Q}$ is **not** complete (the sequence $\{d_n\}$ discussed above is Cauchy, but does not converge to a limit in $\mathbb{Q}$). **Thus completeness is the big difference between $\mathbb{R}$ and $\mathbb{Q}$: $\mathbb{R}$ is complete, but $\mathbb{Q}$ is not.**

Putting our assumptions about the real numbers together, we are assuming that

$\mathbb{R}$ is a ______________, _________________ ______________ __________ .

This leads to three questions:

1. Is there such a thing?
2. If so, how many such things are there (maybe lots)?
3. Do we need any other assumptions about $\mathbb{R}$ to distinguish it from other such things?

Rigorous proofs of the answers to these questions are beyond the scope of this course, but I will tell you that **there is a complete Archimedean ordered field** and **(up to field isomorphism) there is only one complete Archimedean ordered field**. So it is valid to say:

Definition 2.34 *The set $\mathbb{R}$ is the complete Archimedean ordered field. Elements of $\mathbb{R}$ are called* **real numbers**. *Elements of $\mathbb{R}$ that are not in $\mathbb{Q}$ are called* **irrational numbers**.

Getting back to the sequence $\{d_n\}$, we showed that $\{d_n\}$ is Cauchy. By completeness, that means there is a real number L such that $d_n \to L$. The work we did on the previous page shows L^2 must equal 2, meaning that we now know **there is a real number L such that** $L^2 = 2$.

Similar arguments shows the following:

Theorem 2.35 (Existence of square roots) *Let x be any non-negative real number. Then there is another real number $\sqrt{x}$ such that $\left(\sqrt{x}\right)^2 = x$.*

PROOF HW

Hints: Let y_n be the largest rational number that can be written with at most n decimal places whose square is less than or equal to x.

Show $\{y_n\}$ is a Cauchy sequence (similar to how we showed $\{d_n\}$ was Cauchy).

By completeness, this will mean $\{y_n\}$ has a limit. Call this limit $\sqrt{x}$.

Explain why $(\sqrt{x})^2 \le x$.

Explain why it cannot be that $(\sqrt{x})^2 < L$ (by deriving a contradiction similar to the one obtained for $\{d_n\}$). □

Theorem 2.36 (Existence of n^{th} roots) *Let x be any non-negative real number, and let $n \in \{1, 2, 3, \ldots\}$. Then there is another real number $\sqrt[n]{x}$ such that $\left(\sqrt{x}\right)^n = x$.*

PROOF The proof is similar to that of Theorem 2.35 and is omitted. □

Other properties of Cauchy sequences

Theorem 2.37 (Cauchy sequences are bounded) *Let $\{x_n\}$ be a Cauchy sequence. Then $\{x_n\}$ is bounded.*

PROOF Since $\{x_n\}$ is Cauchy, given $\epsilon = 1$, there is N such that if $m, n \ge N$,

$$|x_n - x_m| < 1.$$

In particular, this means that for $n \ge N$, $|x_n - x_N| < 1$, so $|x_n| < |x_N| + 1$. Now, let

$$B = \max\left\{|x_0|, |x_1|, |x_2|, \ldots, |x_{N-1}|, |x_N| + 1\right\}.$$

It is clear that $|x_n| \le B$ for all n, so $\{x_n\}$ is bounded. □

CONSEQUENCE

Let $c_n = n^2$. $\{c_n\}$ is not Cauchy (if it was, it would be bounded).

Theorem 2.38 (The Cauchy property is preserved under arithmetic) *Let $\{x_n\}$ and $\{y_n\}$ be Cauchy sequences. Then:*

1. *$\{r\,x_n\}$ is Cauchy, for any constant $r \in \mathbb{R}$.*
2. *$\{x_n + y_n\}$ is Cauchy.*
3. *$\{x_n - y_n\}$ is Cauchy.*
4. *$\{x_n^2\}$ is Cauchy.*
5. *$\{x_n y_n\}$ is Cauchy.*

Remark: For sequences in $\mathbb{R}$, "Cauchy sequence" and "convergent sequence" are the same thing, so this result would follow immediately from what we proved about arithmetic of convergent sequences (and it would also follow that $\left\{\frac{a_n}{b_n}\right\}$ is Cauchy if $\lim b_n \neq 0$). But this theorem doesn't assume that the context is $\mathbb{R}$, so we can't use completeness.

PROOF The first three statements are HW problems.

For statement (4), assume $\{x_n\}$ is Cauchy and let $\epsilon > 0$.
Since $\{x_n\}$ is Cauchy $\{a_n\}$ is bounded, i.e. $\exists\, B$ s.t. $|x_n| \leq B$ for all n.
Also, $\exists\, N$ such that for $m, n \geq N$, $|x_m - x_n| < \frac{\epsilon}{2B}$.
So for $m, n \geq N$, we have

$$\begin{aligned} |x_m^2 - x_n^2| = |(x_m - x_n)(x_m + x_n)| &= |x_m - x_n|\,|x_m + x_n| \\ &\leq |x_m - x_n|\,(|x_m| + |x_n|) \\ &< \frac{\epsilon}{2B}(B + B) = \epsilon. \end{aligned}$$

Thus $\{x_n^2\}$ is Cauchy by definition.

For statement (5), recall that $\frac{1}{4}(x_n + y_n)^2 - \frac{1}{4}(x_n - y_n)^2 = x_n y_n$.
So by statements (1), (2), (3) and (4), $\{x_n y_n\}$ is Cauchy. □

2.5 Suprema and infima

We have seen that while both $\mathbb{Q}$ and $\mathbb{R}$ are Archimedean ordered fields with a notion of distance, there is a big difference between $\mathbb{Q}$ and $\mathbb{R}$: $\mathbb{R}$ is , whereas $\mathbb{Q}$ is not. In this section, we explore some important consequences of this fact.

Suprema and infima

Definition 2.39 *Let $E \subseteq \mathbb{R}$.*

- *An* **upper bound** *for E is a real number B such that $x \leq B$ for all $x \in E$.*
  If E has an upper bound, we say E is **bounded above***.*
- *A* **lower bound** *for E is a real number B such that $x \geq B$ for all $x \in E$.*
  If E has a lower bound, we say E is **bounded below***.*

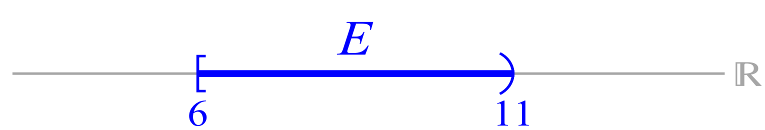
Lemma 2.40 *Let $E \subseteq \mathbb{R}$. E is bounded if and only if E is both bounded above and bounded below.*

PROOF (⇒) If E is bounded, then there is B such that $|x| \leq B$ for all $x \in E$.
Thus is an upper bound for E and is a lower bound for E.

(⇐) Suppose E is bounded below by l and bounded above by u.
Then E is bounded by $\max\{|l|, |u|\}$. □

EXAMPLES

- $E = [6, 11)$



- $F = [5, \infty)$

Definition 2.41 *Given any set $S \subseteq \mathbb{R}$, we define*

$$-E = \{-x : x \in E\}.$$

Lemma 2.42 *Let $S \subseteq \mathbb{R}$.*

- $-(-E) = E$.
- *If E is bounded above by B, then $-E$ is bounded below by $-B$.*
- *If E is bounded below by B, then $-E$ is bounded above by $-B$.*

PROOF These are straightforward arguments. For the first statement:

$$x \in -(-E) \Leftrightarrow -x \in -E \Leftrightarrow -(-x) \in E \Leftrightarrow x \in E.$$

For the second statement, suppose E is bounded above by B.
Now consider $x \in -E$. $-x \in E$ so $-x \leq E$. Thus $x \geq -E$.
Therefore $-B$ is a lower bound for $-E$.

The third statement is left as a HW problem. □

Definition 2.43 *Let $E \subseteq \mathbb{R}$.*

- *A real number s is called a* **supremum** *of E, or a* **least upper bound** *of E, if*
  1. *s is an upper bound for E, and*
  2. *if t is any upper bound for E, then $s \leq t$.*

  In this situation we write $s = \sup E$.

- *A real number i is called an* **infimum** *of E, or a* **greatest lower bound** *of E, if*
  1. *i is a lower bound for E, and*
  2. *if v is any lower bound for E, then $i \geq v$.*

  In this situation we write $i = \inf E$.

PICTURE

Lemma 2.44 (Suprema and infima are unique) *Let $E \subseteq \mathbb{R}$. E can have at most one supremum, and at most one infimum.*

PROOF Suppose s and s' are both suprema of E.

That means they are both upper bounds of E.

By the second part of the definition of supremum, since s is a supremum, s is less than or equal to any upper bound of E (such as s'), so $s \le s'$.

But since s' is a supremum, by the same logic in reverse $s' \le s$.

Since $s \le s'$ and $s' \le s$, $s = s'$.

The uniqueness of the infimum of a set that is bounded below has a similar proof. □

Lemma 2.45 (Reversing Lemma) *Let $E \subseteq \mathbb{R}$.*

1. *If $s = \sup E$, then $-s = \inf(-E)$.*
2. *If $i = \inf E$, then $-i = \sup(-E)$.*

PROOF For statement (1), let $s = \sup E$. To show $-s = \inf(-E)$, we need to show two things:

1. We need to show

   To do this, let $x \in -E$.

   Therefore

   Therefore

   Thus $x \ge -s$, meaning $-s$ is a lower bound of $-E$.

2. We need to show

   To do this, suppose v is <u>any</u> lower bound of $-E$.

   By Lemma 2.42, $-v$ is

   Since $s = \sup E$, we know

   Thus $v \le -s$. Therefore $-s$ is the greatest lower bound of $-E$, so $-s = \inf(-E)$.

Statement (2) has a similar proof, and is left as HW. □

EXAMPLE

Let $a, b \in \mathbb{R}$ be such that $a < b$. Determine, with proof, the supremum and infimum of the interval $E = [a, b)$.

Similar arguments as to those on the previous page show:

Lemma 2.46 (Suprema and infima of intervals) *Let $a, b \in \mathbb{R}$. Then:*

- $\sup(a,b) = \sup(a,b] = \sup[a,b) = \sup[a,b] = \sup(-\infty,b) = \sup(-\infty,b] = b$.
- $\inf(a,b) = \inf(a,b] = \inf[a,b) = \inf[a,b] = \inf(a,\infty) = \inf[a,\infty) = a$.

Now for a very important consequence of completeness:

Theorem 2.47 (Supremum Property) *Let $E \subseteq \mathbb{R}$ be nonempty. If E is bounded above, then $\sup E$ exists.*

Note: The supremum property fails in $\mathbb{Q}$. For instance consider the set

$$E = \{x \in \mathbb{Q} : x^2 < 2\} = (-\infty, \sqrt{2}) \cap \mathbb{Q}$$

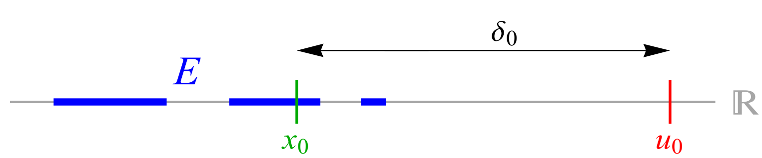
E has no supremum in $\mathbb{Q}$ (by an argument similar to the one on the previous page, the supremum would have to be a number s such that $s^2 = 2$).

PROOF OF THE SUPREMUM PROPERTY First, here's a summary of how the proof works. We'll recursively construct a sequence $\{u_n\}$ of real numbers, all of which are upper bounds for E, and this sequence will turn out to be a Cauchy sequence. By completeness, this sequence has a limit which we'll call s. Last, we'll show that $s = \sup E$.

Now for the details: since E is bounded above, it has an upper bound u_0.
Now, let x_0 be any number which is not an upper bound of E
(doable since $E \neq \varnothing$: for instance, take any $z \in E$ and let $x_0 = z - 1$).
Let $\delta_0 = u_0 - x_0$; notice $\delta_0 > 0$.

Now, recursively construct sequences $\{u_n\}$ and $\{x_n\}$ as follows: to define u_{n+1} and x_{n+1} from u_n and x_n,

1. Set $m_n = \frac{1}{2}(u_n + x_n)$ to be the midpoint between u_n and x_n.
2. If m_n is an upper bound for S, then set $u_{n+1} = m_n$ and $x_{n+1} = x_n$.
3. If m_n isn't an upper bound for S, then set $u_{n+1} = u_n$ and $x_{n+1} = m_n$.



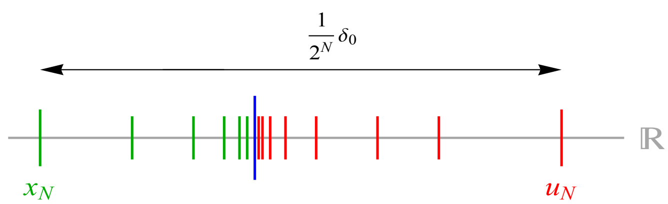
Observations:

1. $\{x_n\}$ is an increasing sequence of numbers, none of which are upper bounds of E;
2. $\{u_n\}$ is a decreasing sequence of upper bounds of E;
3. every x_n is less than every u_m;
4. $|u_{n+1} - x_{n+1}| = \frac{1}{2}|u_n - x_n|$, meaning $|u_n - x_n| = \frac{1}{2^n}\delta_0$; so by the Squeeze Theorem $(u_n - x_n) \to 0$ (and $(x_n - u_n) \to 0$ also).

Claim 1: $\{u_n\}$ is a Cauchy sequence.

Proof of Claim 1: Since $\frac{1}{2^n}\delta_0 \to 0$, given $\epsilon > 0$ we can find N so that $\frac{1}{2^N}\delta_0 < \epsilon$. Now suppose $m, n \geq N$. Notice that whenever $m, n \geq N$,

$$u_m, u_n \in [x_N, u_N].$$



By applying observation (4) above, that means

$$|u_m - u_n| \leq |u_N - x_N| = \frac{1}{2^N}\delta_0 < \epsilon.$$

This proves Claim 1.

By completeness, since $\{u_n\}$ is Cauchy, there is a real number $s = \lim u_n$.

Claim 2: $s = \lim x_n$.

Proof of Claim 2: Applying observation (4) above,

$$\lim x_n = \lim[u_n + (x_n - u_n)] = \lim u_n + \lim(x_n - u_n) = s + 0 = s.$$

Claim 3: $s = \sup E$.

Proof of Claim 3: Let $x \in E$. By construction, $x \leq u_n$ for every n.

Since limits preserve $\leq$, that means $x \leq \lim u_n = s$.

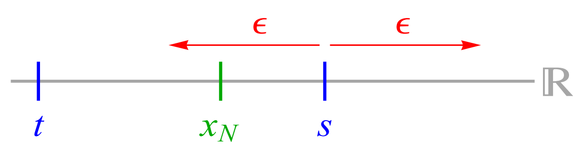
Thus s is an upper bound for E.

Second, let $t < s$. Set $\epsilon = \frac{1}{2}(s-t)$.

Since $x_n \to s$, there is N such that $|x_n - L| < \epsilon$ for all $n \geq N$.

This means

$$\begin{aligned} & |x_N - s| < \epsilon \\ \Rightarrow \quad & s - x_N < \frac{1}{2}(s-t) \\ \Rightarrow \quad & x_N > s - \frac{1}{2}(s-t) = \frac{1}{2}(s+t) > \frac{1}{2}(t+t) = t. \end{aligned}$$



Since x_N is not a lower bound of E, neither is t.

We've proven no number less than s is an upper bound for E, so this means s must be the least upper bound of E, i.e. $s = \sup E$ as wanted. □

Theorem 2.48 (Infimum property) *Let $E \subseteq \mathbb{R}$ be nonempty. If E is bounded below, then $\inf E$ exists.*

PROOF Since E is bounded below, $-E$ is bounded above.

So by the Supremum Property, $\sup E$ exists.

Finally, by the Reversing Lemma, $\inf E = -\sup(-E)$. □

Let's introduce some notation that may help with subsequent results:

Definition 2.49 *If E is not bounded above, we write $\sup E = \infty$, but this doesn't mean that the supremum of E actually exists.*

If E is not bounded below, we write $\inf E = -\infty$. Again, this doesn't mean that the infimum of E actually exists.

EXAMPLE

What is the supremum of $\varnothing$?

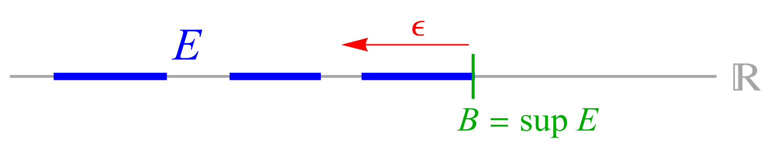
> **Theorem 2.50** *If $S \neq \varnothing$, then $\inf S \leq \sup S$.*

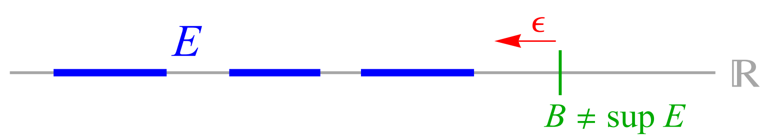
PROOF Let $x \in S$. Since infima are lower bounds, $\inf S \leq x$. Since suprema are upper bounds, we have $x \leq \sup S$. Apply transitivity of $\leq$. □

Other characterizations of suprema and infima

> **Lemma 2.51** *Let B be an upper bound of nonempty $E \subseteq \mathbb{R}$. Then*
>
> $$B = \sup E \quad \Leftrightarrow \quad \forall\, \epsilon > 0, \exists\, x \in E \text{ such that } B - \epsilon < x.$$

PICTURES TO EXPLAIN





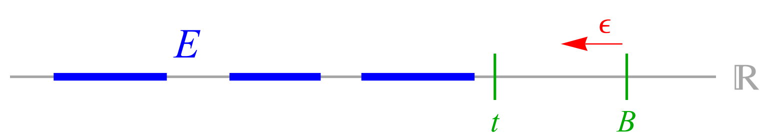
PROOF (⇒) Suppose $B = \sup E$ and let $\epsilon > 0$.

If there is no $x \in E$ such that $L - \epsilon < x$, then $B - \epsilon$ would be an upper bound for E strictly less than B, contradicting $B = \sup E$.

(⇐) Suppose not, i.e. that there is an upper bound t of E with $t < B$.

Let $\epsilon = \frac{1}{2}(B - t)$ and notice

$$B - \epsilon = B - \frac{1}{2}(B - t) = \frac{1}{2}B + \frac{1}{2}t = t + \frac{1}{2}(B - t) = t + \epsilon > t.$$



By hypothesis, $\exists\, x \in E$ with $B - \epsilon < x$.
For this x, $x > t$, contradicting the fact that t is an upper bound of E.
Therefore $B = \sup E$. □

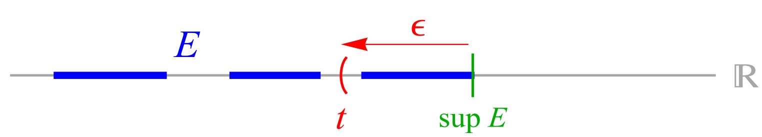
HOMEWORK

Formulate and prove a lemma analagous to Lemma 2.51 for infima, rather than suprema.

Lemma 2.52 *Let $E \subseteq \mathbb{R}$ be a set which is bounded above. Then, for every $t < \sup E$,*

$$(t, \sup E] \cap E \neq \varnothing.$$

PROOF Let $t < \sup E$. Then, define $\epsilon = \sup E - t > 0$.



By Lemma 2.51, there is $x \in E$ such that $\sup E - \epsilon < x$.
This means $\sup E - (\sup E - t) = t < x$.
This x belongs to $(t, \sup E] \cap E$, so $(t \sup E] \cap E = \varnothing$. □

HOMEWORK

Formulate and prove a lemma analagous to Lemma 2.52 for infima, rather than suprema.

Lemma 2.53 *Let $E \subseteq \mathbb{R}$ be a set which is bounded above. Then there exists an increasing sequence $\{x_n\}$ of points in E with $x_n \to \sup E$.*

Let $E \subseteq \mathbb{R}$ be a set which is bounded below. Then there exists a decreasing sequence $\{x_n\}$ of points in E with $x_n \to \sup E$.

PROOF The proof of the first statement is HW.

For the second statement, if E is bounded below, then $-E$ is bounded above.
By the first statement, there is an increasing sequence $\{x_n\} \subseteq -E$ with

$$x_n \to \sup(-E).$$

Thus $-x_n \to -\sup(-E) = \inf E$.
Since each $x_n \in -E$, $-x_n \in E$, so $\{-x_n\}$ is the desired sequence. □

Lemma 2.54 *If $E \subseteq \mathbb{R}$ is bounded above, then*

$$\sup S = \inf\{t : t \text{ is an upper bound for } S\}.$$

If $E \subseteq \mathbb{R}$ is bounded below, then

$$\inf S = \sup\{v : v \text{ is a lower bound for } S\}.$$

PROOF The first statement is HW.

Hints: Let

$$U = \{t : t \text{ is an upper bound of } E\}$$

and let $s = \sup E$.

To prove $s = \inf U$, you need to show two things: first, that s is a lower bound for U and second, if v is any lower bound for U, then $v < s$.

For the second statement, apply the Reversing Lemma:

$$\begin{aligned}
\sup\{v : v \text{ is a lower bound for } E\} &= -\inf\left[-\{v : v \text{ is a lower bound for } E\}\right] \\
&= -\inf\{v : v \text{ is an upper bound for } -E\} \\
&= -\sup(-E) \\
&\qquad \text{(by the first statement, applied to } -E) \\
&= \inf E. \quad \square
\end{aligned}$$

2.6 Other consequences of completeness

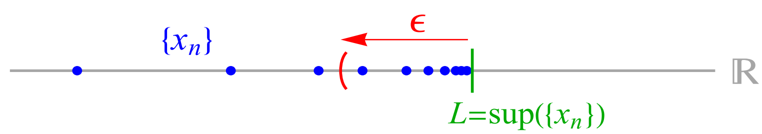
Monotone Convergence Theorem

Suprema and infima can also be used to tell us something about certain kinds of sequences of real numbers:

Theorem 2.55 (Monotone Convergence Theorem) *Let $\{x_n\}$ be a sequence of real numbers which is increasing and bounded above. Then $x_n \to \sup(\{x_n\})$.*

Remark: The important conclusion here is that the sequence converges to *something* (the fact that the limit is $\sup(\{x_n\})$ is sometimes useful, but less important).

PROOF Let $L = \sup(\{x_n\})$. To prove $x_n \to L$, let $\epsilon > 0$.
By Lemma 2.51, there is N such that $L - \epsilon < x_N$.



Since $\{x_n\}$ is increasing, $x_n > L - \epsilon$ for all $n \geq N$.
At the same time, $x_n \leq L$ since suprema are upper bounds.
Thus we have, for all $n \geq N$,

$$L - \epsilon < x_n \leq L,$$

implying $|x_n - L| < \epsilon$. Thus $x_n \to L$ by definition. □

Corollary 2.56 (Monotone Convergence Theorem) *Let $\{x_n\}$ be a sequence of real numbers which is decreasing and bounded below. Then $x_n \to \inf(\{x_n\})$.*

PROOF Suppose $\{x_n\}$ is decreasing and bounded below.
Then, $\{-x_n\}$ is increasing and bounded above.
So $-x_n \to \sup(\{-x_n\}) = -\inf(\{x_n\})$, meaning $x_n \to \inf(\{x_n\})$. □

Archimedean properties

The *Archimedean properties* of the real numbers generally refer to the idea that $\mathbb{R}$ contains arbitrarily large numbers, and arbitrarily small positive numbers.

The first version of the Archimedean Property, which we assume without proof, ensures that the real numbers contain arbitrarily large whole numbers:

Corollary 2.57 *$\mathbb{R}$ is unbounded.*

PROOF Suppose not, i.e. $\mathbb{R}$ is bounded, say by B.
That means $\mathbb{R} \subseteq [-B, B]$.
But by the Archimedean Property, there is $n \in \mathbb{N} \subseteq \mathbb{R}$ with $n > B$.
Thus $n \notin [-B, B]$, contradicting $\mathbb{R} \subseteq [-B, B]$. □

The second version of the Archimedean Property ensures that the real numbers contain arbitrarily small positive numbers:

Theorem 2.58 (Archimedean Property II) *Given any $x \in (0, \infty)$, there is $n \in \mathbb{N}$ such that $\frac{1}{n} < x$.*

PROOF Let $x > 0$. Then $\frac{1}{x} > 0$.

Apply the first Archimedean Property to find $n \in \mathbb{N}$ with $n > \frac{1}{x}$.

Take reciprocals of both sides to get $\frac{1}{n} < x$. □

The third version of the Archimidean Property says that every positive real number can be squeezed between two whole numbers:

Theorem 2.59 (Archimedean Property III) *Let $x \in (0, \infty)$. Then there is $n \in \mathbb{N}$ such that $n \le x < n + 1$.*

PROOF HW

Hint: Consider the set E of natural numbers which are $\le x$.

Show this set is nonempty and bounded above, and proceed from there.

The Density Theorem

Theorem 2.60 (Density Theorem) *Let $a, b \in \mathbb{R}$ be such that $a < b$. Then:*

1. *$\exists\, x \in \mathbb{Q}$ s.t. $a < x < b$; and*
2. *$\exists\, x \in (\mathbb{R} - \mathbb{Q})$ s.t. $a < x < b$.*

This theorem tells us that the rational numbers are *dense* in $\mathbb{R}$, i.e. that a picture of the real numbers where the rationals are indicated looks like this:

Interestingly, the irrational numbers are also dense in $\mathbb{R}$!

PROOF For the first statement, assume $a < b$.

Therefore $b - a > 0$, so by the Archimedean Property (II), $\exists\, n \in \mathbb{N}$ s.t. $\dfrac{1}{n} < b - a$, meaning $a + \dfrac{1}{n} < b$.

Then let

$$E = \left\{ p \in \mathbb{Z} : \frac{p}{n} \le b \right\}.$$

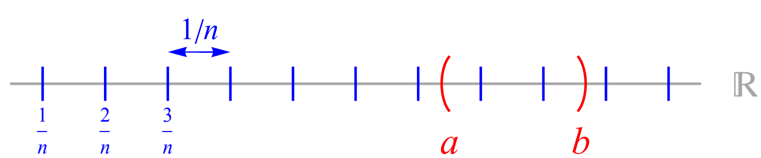
Claim: $E \ne \varnothing$.

Proof of Claim: If $b \ge 0$, then $0 \in E$.

If $b < 0$, by the Archimedean Property (I) $\exists\, p \in \mathbb{N}$ s.t. $p > -bn$.

Then $-p < bn$, so $-p \in E$.

Either way, $E \ne \varnothing$. This proves the claim.

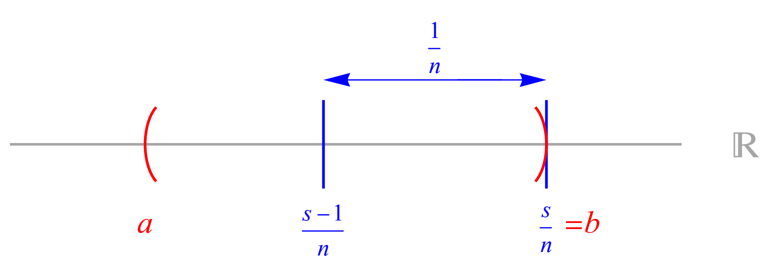


Next, if $p \in E$, then $\dfrac{p}{n} \le b$ so $p \le bn$, meaning E is bounded above by bn.

That means $s = \sup E$ exists (and is an integer), and $\dfrac{s}{n} \le b$.

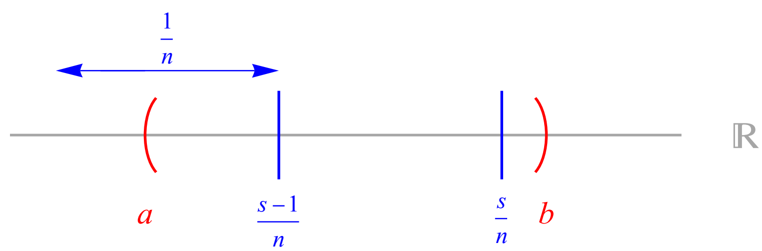
There are two cases:

Case 1: $\dfrac{s}{n} = b$. In this case, set $x = \dfrac{s-1}{n}$.



$x \in \mathbb{Q}$, $x < b$, and $x = \dfrac{s}{n} - \dfrac{1}{n} = b - \dfrac{1}{n} > a$ from above, so $a < x < b$ as wanted.

Case 2: $\frac{s}{n} < b$. In this case, set $x = \frac{s}{n}$.



Clearly $x \in \mathbb{Q}$ and $x < b$.

Last, if $x \le a$, then $x + \frac{1}{n} = \frac{s+1}{n} < a + \frac{1}{n} < b$, so $s + 1 \in E$, contradicting $s = \sup E$.

Therefore $a < x < b$ as wanted.

The second statement is a HW problem.

Hint: Instead of directly finding an irrational between a and b, use the first part of the Density Theorem to find a rational number between two other real numbers.

Then use a formula of that rational number to obtain an irrational number between a and b. □

2.7 Subsequences

It is often useful to build a new sequence from a given one by picking out certain elements of the sequence (like picking out every other one, or ones with certain properties, etc.) This is called constructing a *subsequence*.

Definition 2.61 *Let $\{x_n\}_{n=m}^{\infty}$ be a sequence and let $m \leq n_1 < n_2 < n_3 < \ldots$ be a strictly increasing sequence of integers. The sequence*

$$\{x_{n_k}\}_{k=1}^{\infty} = \{x_{n_1}, x_{n_2}, x_{n_3}, \ldots\}$$

is called a **subsequence** *of $\{x_n\}$. A real number L is called a* **subsequential limit** *of $\{x_n\}$ if there is a subsequence $\{x_{n_k}\}$ with $a_{n_k} \stackrel{k\to\infty}{\longrightarrow} L$.*

EXAMPLE

Let $b_n = 1 + (-1)^n$, i.e. $\{b_n\} = \{2, 0, 2, 0, 2, 0, \ldots\}$.

Note: subsequences have to have infinitely many terms, and in particular, as $k \to \infty$, $n_k \to \infty$ as well. This means that given any N, there is always a K such that if $k \geq K$, $n_k \geq N$.

Theorem 2.62 *Let $\{x_n\}$ be a sequence of real numbers.*

1. *If $\{x_n\}$ is Cauchy, then any subsequence of $\{x_n\}$ is also Cauchy.*
2. *If $x_n \to L$, then any subsequence of $\{x_n\}$ must also converge to L.*
3. *If $\{x_n\}$ has two different subsequential limits, then $\{x_n\}$ diverges.*

PROOF For the first statement, let $\{x_{n_k}\}$ be a subsequence of $\{x_n\}$.

Let $\epsilon > 0$.

Since $\{x_n\}$ is Cauchy, $\exists\, N$ s.t. $m, n \geq N$ implies $|x_m - x_n| < \epsilon$.

Choose K such that $k \geq K$ implies $n_k > N$.

Then, for any $j, k \geq K$, n_j and n_k are $\geq N$, so $|x_{n_j} - x_{n_k}| < \epsilon$.

Therefore $\{x_{n_k}\}$ is Cauchy.

For statement (2), let $\epsilon > 0$. Since ____________, there is N such that $n \geq N$ implies

________________ . Choose _____ such that __________ implies __________ .

Then, for any ____________, ____________________, so $a_{n_k} \to L$.

The last statement is the contrapositive of the second. □

EXAMPLE

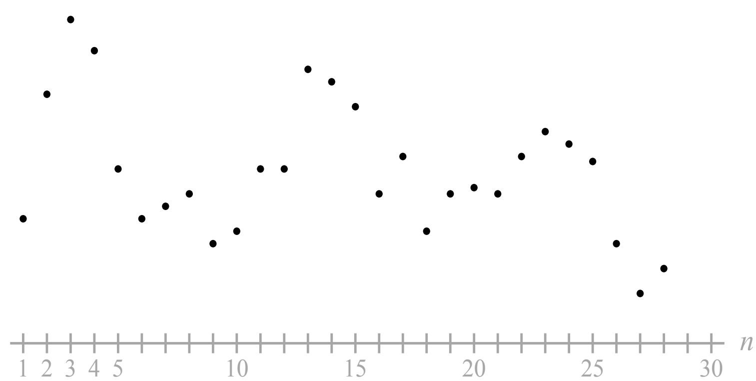
Let $b_n = 1 + (-1)^n$, i.e. $\{b_n\} = \{2, 0, 2, 0, 2, 0, \ldots\}$. Prove $\{b_n\}$ diverges.

Subsequence existence theorems

Theorem 2.63 (Monotone Subsequence Theorem) *Every sequence of real numbers has a monotone subsequence.*

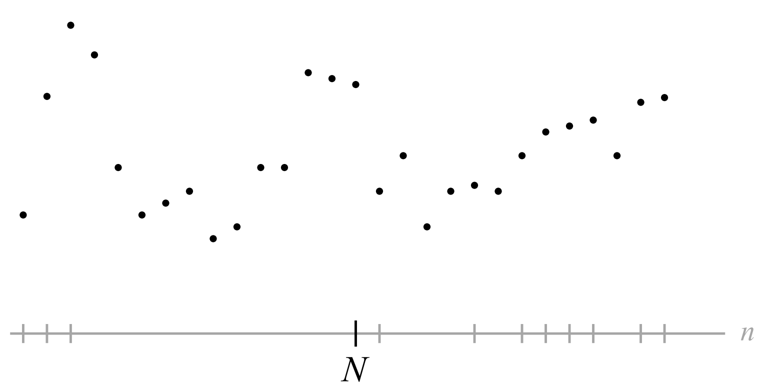
PROOF Let $\{x_n\}$ be a sequence of real numbers.

Define an index n to be a **peak** if $x_n \geq x_m$ for every $m \geq n$.



Case 1: If there are infinitely many indices which are peaks, set $n_k = k^{th}$ peak. By the definition of peak, $x_{n_k} \geq x_{n_{k+1}}$, so we have described a decreasing subsequence $\{x_{n_k}\}$ of $\{a_n\}$.

Case 2: If there are only finitely many peaks, let N be the largest peak (N is an index of the sequence).



Let $n_1 = N + 1$. Since n_1 isn't a peak, $\exists\, n_2 > n_1$ with $x_{n_2} > x_{n_1}$.
Since n_2 is not a peak, there is $n_3 > n_2$ with $x_{n_3} > x_{n_2}$, etc.
Continuing in this way, we get an increasing subsequence $\{x_{n_k}\}$. □

Theorem 2.64 (Bolzano-Weierstrass Theorem) *Every bounded sequence of real numbers has a convergent subsequence.*

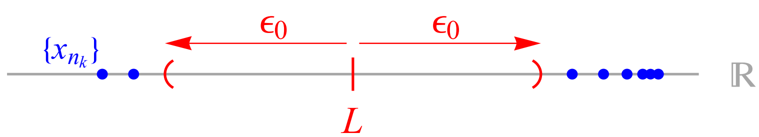
PROOF HW

Hint: This isn't hard if you research the previous theorems of this chapter.

Our next theorem here says that if $a_n \not\to L$, then you can build a subsequence of $\{a_n\}$ that "avoids" L:

Theorem 2.65 (Avoidance Theorem) *Let $\{x_n\}$ be a sequence of real numbers. Then*

$$x_n \not\to L \quad \Leftrightarrow \quad \exists\, \epsilon_0 > 0, \exists \text{ subsequence } \{x_{n_k}\} \text{ s.t. } \forall\, k, |x_{n_k} - L| \geq \epsilon_0.$$



PROOF ($\Rightarrow$) Suppose $x_n \not\to L$.

Then, for some $\epsilon_0 > 0$, there is no N such that $|x_n - L| < \epsilon_0$ for all $n \geq N$.

That means there are infinitely many n such that $|x_n - L| \geq \epsilon_0$.

Let n_k be the k^{th} such n; this gives a subsequence $\{x_{n_k}\}$ with the desired properties.

($\Leftarrow$) We prove the result by contradiction.

Let $\epsilon_0 > 0$ and $\{x_{n_k}\}$ be a subsequence s.t. $|x_{n_k} - L| \geq \epsilon_0$ for all k.

If $x_n \to L$, then there would be N such that $n \geq N$ would imply $|x_n - L| < \epsilon_0$.

But $n_k \to \infty$ as $k \to \infty$, so eventually one of these x_n would be an x_{n_k}, violating the hypothesis. □

Theorem 2.66 *If $\{x_n\}$ is a bounded sequence of real numbers so that every convergent subsequence of $\{x_n\}$ converges to L, then $x_n \to L$.*

PROOF Suppose not. Then, by the Avoidance Theorem, there is $\epsilon_0 > 0$ and there is a subsequence $\{x_{n_k}\}$ such that for all k, $|x_{n_k} - L| \geq \epsilon_0$.

By the Bolzano-Weierstrass Theorem, $\{x_{n_k}\}$ has a convergent subsequence $\{x_{n_{k_l}}\}$ which converges to some L'.

But since $\left|x_{n_{k_l}} - L\right| \geq \epsilon_0$ for all l, and limits preserve $\geq$, $|L' - L| \geq \epsilon_0 > 0$ so $L \neq L'$.

This is a contradiction. □

2.8 Limits superior and inferior

Definition 2.67 *Let $\{x_n\}$ be a sequence of real numbers.*

- *If $\{x_n\}$ is bounded above, let*

$$\overline{\lim}\, x_n = \lim_{n\to\infty} \left(\sup\{x_m : m \geq n\}\right).$$

  $\overline{\lim}\, x_n$ is called the **limit superior** *(or just* **lim sup***) of $\{x_n\}$ and is also denoted* $\limsup a_n$.

  If $\{x_n\}$ is not bounded above, we write $\overline{\lim}\, x_n = \infty$, though this does not mean the limit superior actually exists.

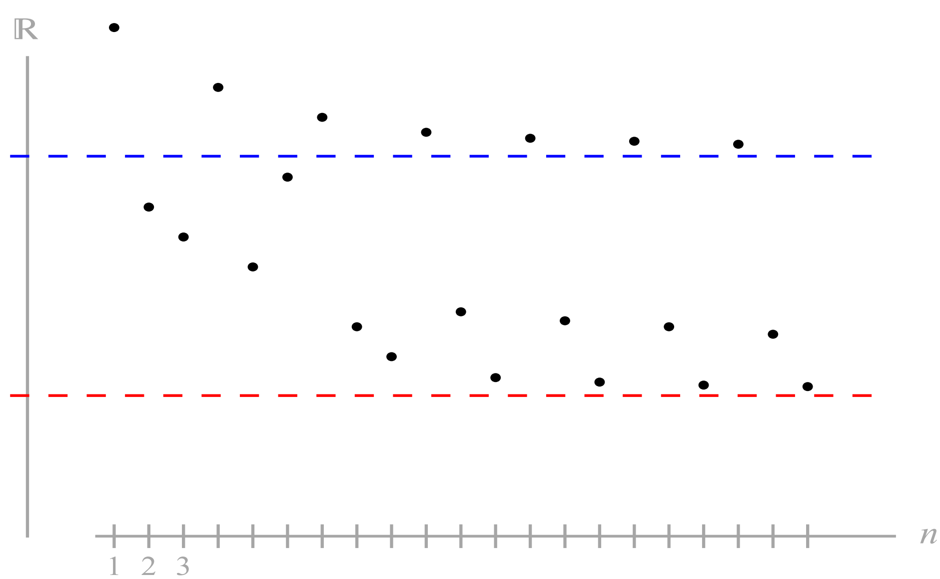
- *If $\{x_n\}$ is bounded above, let*

$$\underline{\lim}\, x_n = \lim_{n\to\infty} \left(\inf\{x_m : m \geq n\}\right).$$

  $\underline{\lim}\, x_n$ is called the **limit inferior** *(or just* **lim inf***) of $\{x_n\}$ and is also denoted* $\liminf a_n$.

  If $\{x_n\}$ is not bounded below, we write $\overline{\lim}\, x_n = -\infty$, though this does not mean the limit inferior actually exists.

A PICTURE TO EXPLAIN



Theorem 2.68 *Let $\{x_n\}$ be a bounded sequence of real numbers.*

1. $\overline{\lim}\, x_n$ *exists.*
2. $\underline{\lim}\, x_n$ *exists.*
3. *If $\{x_n\}$ is bounded, then $\underline{\lim}\, x_n \leq \overline{\lim}\, x_n$.*

PROOF For the first statement, for each n let $s_n = \sup\{x_m : m \geq n\}$.

Since $\{x_n\}$ is bounded below, so is $\{s_n\}$.

Also, notice that for each n,

$$\{x_m : m \geq n\} \supseteq \{x_m : m \geq n+1\}.$$

Therefore $s_n = \sup\{x_m : m \geq n\} \geq \sup\{x_m : m \geq n+1\} = s_{n+1}$, meaning that the sequence $\{s_n\}$ is decreasing.

Since $\{s_n\}$ is decreasing and bounded below, by the Theorem,

$\lim s_n$ exists, i.e. $\lim(\sup\{x_m : m \geq n\})$ exists, i.e. $\overline{\lim}\, x_n$ exists.

The second statement is HW.

For statement (3), if we let s_n be as above and let $i_n = \inf\{x_m : m \geq n\}$, we have

$$i_n \leq s_n$$

Since limits preserve $\leq$,

$$\lim i_n \leq \lim s_n$$
$$\lim \inf\{x_m : m \geq n\}) \leq \lim(\sup\{x_m : m \geq n\})$$
$$\underline{\lim}\, x_n \leq \overline{\lim}\, x_n. \; \square$$

Theorem 2.69 *Let $\{x_n\}$ be a bounded sequence of real numbers and let S be the set of subsequential limits of $\{x_n\}$. Then*

$$\overline{\lim}\, x_n = \sup S \quad \textit{and} \quad \underline{\lim}\, x_n = \inf S.$$

PROOF To prove the first statement, we will to establish two claims:

Claim 1: $\overline{\lim}\, x_n$ is an upper bound of S.

Claim 2: $\overline{\lim}\, x_n \in S$. (This ensures no number $< \overline{\lim}\, x_n$ is an upper bound of S.)

Proof of Claim 1: Let $s \in S$. That means $\exists$ subsequence $\{x_{n_k}\}$ with $x_{n_k} \xrightarrow{k\to\infty} s$.

Let $\epsilon > 0$.

Then $\exists K$ s.t. $k \geq K$ implies $|x_{n_k} - s| < \epsilon$, i.e. $x_{n_k} > s - \epsilon$.

Also, since $n_k \to \infty$, $\exists k \geq K$ s.t. $n_k > n$, and for this k,

$$\sup\{x_m : m \geq n\} \geq x_{n_k} > s - \epsilon.$$

Since limits preserve $\geq$, $\overline{\lim}\, x_n = \lim \sup\{x_m : m \geq n\} \geq s - \epsilon$.

Last, since $\epsilon > 0$ is arbitrary, $\overline{\lim}\, x_n \geq s$.

Therefore $\overline{\lim}\, x_n$ is an upper bound for S.

Proof of Claim 2: Let $L = \overline{\lim}\, x_n = \lim(\sup\{x_m : m \geq n\})$.

We will construct a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ that converges to L by defining n_1, then n_2, then n_3 and so on.

To define n_1, observe that by definition of convergence (with $\epsilon = \frac{1}{4}$), $\exists N_1$ s.t.

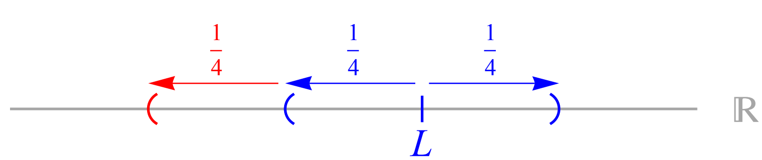
$$n \geq N_1 \Rightarrow |\sup\{x_m : m \geq n\} - L| < \frac{1}{4},$$

i.e. $L - \frac{1}{4} < \sup\{x_m : m \geq n\} < L + \frac{1}{4}$.

By a characterization of sup (with $\epsilon = \frac{1}{4}$), $\exists n_1 \geq n \geq N_1$ s.t.

$$\begin{aligned} L - \frac{1}{4} - \frac{1}{4} &< x_{n_1} < L + \frac{1}{4} \\ L - \frac{1}{2} &< x_{n_1} < L + \frac{1}{4}, \end{aligned}$$

i.e. $|x_{n_1} - L| < \frac{1}{2}$.



To define n_2, repeat this procedure with a smaller ϵ.

In particular, use $\epsilon = \dfrac{1}{8}$ to find N_2 (WLOG $N_2 > n_1$) s.t.

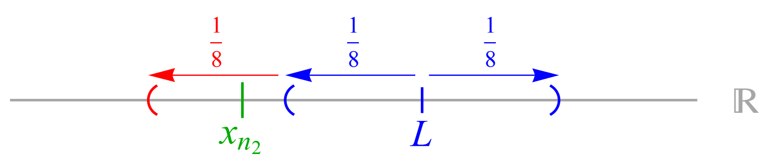
$$n \geq N_2 \Rightarrow |\sup\{x_m : m \geq n\} - L| < \frac{1}{6},$$

i.e. $L - \dfrac{1}{4} < \sup\{x_m : m \geq n\} < L + \dfrac{1}{8}$.

By a characterization of sup (with $\epsilon = \dfrac{1}{8}$), $\exists\, n_2 \geq n \geq N_2 > n_1$ s.t.

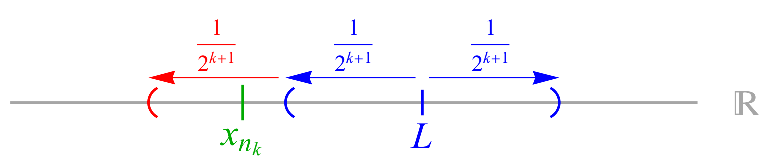
$$\begin{aligned} L - \frac{1}{8} - \frac{1}{8} &< x_{n_2} < L + \frac{1}{8} \\ L - \frac{1}{4} &< x_{n_2} < L + \frac{1}{8}, \end{aligned}$$

i.e. $|x_{n_2} - L| < \dfrac{1}{4}$.



More generally, for each k we can (using $\epsilon = \dfrac{1}{2^{k+1}}$) define $n_{k+1} > n_k$ so that

$$|x_{n_k} - L| < \frac{1}{2^k}.$$



By the Squeeze Theorem, $x_{n_k} \to L$, so $L \in S$, proving Claim 2.

The proof of the second statement ($\underline{\lim}\, x_n = \inf S$) is left as HW. □

Theorem 2.70 *Let $\{x_n\}$ be a bounded sequence of real numbers. TFAE:*

1. $\underline{\lim}\, x_n = \overline{\lim}\, x_n = L$.
2. $x_n \to L$.

PROOF Throughout the proof, let S be the set of subsequential limits of $\{x_n\}$.

(1) $\Rightarrow$ (2): $S \neq \varnothing$ by the [   ] Theorem.

By hypothesis,

$$\inf S = \underline{\lim}\, x_n = L = \overline{\lim}\, x_n = \sup S,$$

so $S = \{L\}$. In other words, every convergent subsequence of $\{x_n\}$ converges to L. By Theorem 2.66, $x_n \to L$.

(2) $\Rightarrow$ (1): Suppose $x_n \to L$.

Thus every subsequence $\{x_{n_k}\}$ also converges to L, so $S = \{L\}$.

Therefore $L = \inf S = \underline{\lim}\, x_n$ and $L = \sup S = \overline{\lim}\, x_n$. □

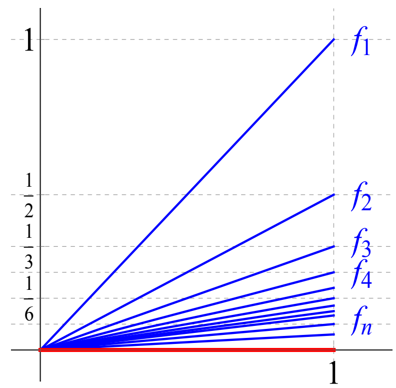
Remark: To apply the (1) $\Rightarrow$ (2) direction of Theorem 2.69, it is sufficient to show $\overline{\lim}\, x_n \leq \underline{\lim}\, x_n$ (since we know the opposite inequality always holds).

2.9 Sequences of functions

To this point, we have focused on studying equences $\{x_n\}$ of *numbers* (determining which sequences converge, which are Cauchy, etc.).
It is also useful to discuss the convergence of sequences of other types of mathematical objects (vectors, matrices, random variables, etc.); in this course, we care about sequences of *functions*.

EXAMPLE

Let $\{f_n\}$ be the sequence of functions $[0,1] \to \mathbb{R}$ defined by $f_n(x) = \dfrac{x}{n}$.



QUESTION

What does it mean for such a sequence to *converge*?

Pointwise convergence

Definition 2.71 *Let $E \subseteq \mathbb{R}$ and $\{f_n\}$ be a sequence of functions $E \to \mathbb{R}$.*
We say $\{f_n\}$ **converges (pointwise) (on E)** *to $f : E \to \mathbb{R}$, and write $f_n \to f$ on E, if $f_n(x) \to f(x)$ for all $x \in E$.*
Equivalently, $\forall\, x \in E$ and $\forall\, \epsilon > 0\ \exists\, N = N(x,\epsilon) \in \mathbb{N}$ s.t.

$$n \geq N \Rightarrow |f_n(x) - f(x)| < \epsilon.$$

In this context, f is called the **(pointwise) limit** *of $\{f_n\}$.*

Good things about pointwise convergence:

1. It is usually easy to compute the pointwise limit of a sequence of functions.

   For example, if $f_n(x) = \dfrac{x}{n}$, then $f_n \to f$ where

2. It preserves soft inequalities: if $f_n(x) \le g(x)$ for all x and $f_n \to f$, then $f(x) \le g(x)$ for all x.

3. There is a completeness property: if, for every $x \in E$, $\{f_n(x)\}$ is a Cauchy sequence, then $\exists f : E \to \mathbb{R}$ such that $f_n \to f$.

Bad things about pointwise convergence:

1. If all you know is $f_n \to f$, you usually can't conclude anything about f (like whether it is continuous, differentiable, or integrable) from information coming from the f_n.

2. There is no notion of "distance" between two functions that is consistent with pointwise convergence.

Uniform convergence

GOAL

Come up with a notion of "convergence" of a sequence of functions that avoids the drawbacks of pointwise convergence.

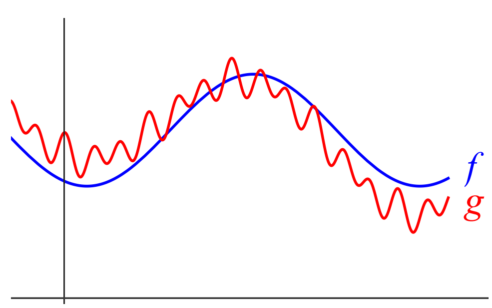
To do this, let's think about convergent sequences of numbers.
We can colloquially restate the idea that $x_n \to L$ by saying

"when n is large, x_n becomes arbitrarily close to L".

So if we have a sequence of functions $\{f_n\}$ that "converges" to f, we might say

"when n is large, f_n becomes arbitrarily close to f."

This begs a question: what does it mean for one function to be "close"?
For real numbers x and y, they are within ϵ of one another if $|x - y| < \epsilon$.
What do you think it means for two *functions* f and g to be within a "distance" of $< \epsilon$ from one another?



With this in mind, we might say that $\{f_n\}$ converges to f if

$$\text{"when } n \text{ is large, } |f_n(x) - f(x)| < \epsilon \text{ for all } x\text{".}$$

This leads to the following definition.

Definition 2.72 *Let $E \subseteq \mathbb{R}$ and let $\{f_n\}$ be a sequence of functions $E \to \mathbb{R}$. We say $\{f_n\}$* **converges uniformly (on E)** *to $f : \mathbb{R} \to \mathbb{R}$, and write $f_n \rightrightarrows f$* **on** *E, if $\forall\, \epsilon > 0$, $\exists\, N = N(\epsilon) \in \mathbb{N}$ s.t.*

$$n \geq N \Rightarrow |f_n(x) - f(x)| < \epsilon \text{ for all } x \in E.$$

In this context, f is called the **uniform limit** *of $\{f_n\}$.*

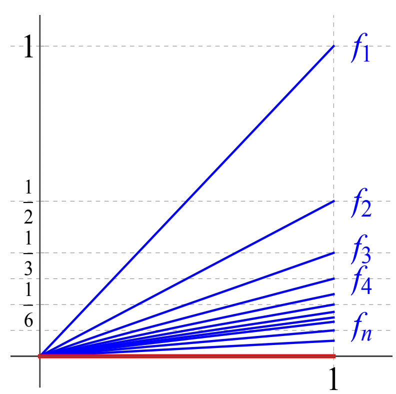
Lemma 2.73 (Uniform convergence implies pointwise convergence) *Let $E \subseteq \mathbb{R}$ and let $\{f_n\}$ be a sequence of functions $E \to \mathbb{R}$. If $f_n \rightrightarrows f$ on E, then $f_n \to f$ on E.*

PROOF This is immediate from the definitions. □

The converse of Lemma 2.73 is false. Consider these examples:

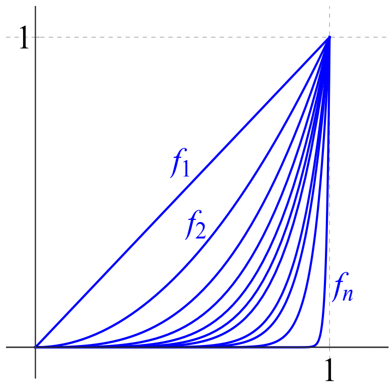
EXAMPLE A

Let $\{f_n\}$ be the sequence of functions $f_n : [0,1] \to \mathbb{R}$ defined by $f_n(x) = \dfrac{x}{n}$. Earlier, we observed $f_n \to 0$. Does $f_n \rightrightarrows 0$?



EXAMPLE B

Let $\{f_n\}$ be the sequence of functions $f_n : [0,1] \to \mathbb{R}$ defined by $f_n(x) = x^n$.
Find a function $f : [0,1] \to \mathbb{R}$ so that $f_n \to f$. Does $f_n \rightrightarrows f$?



Completeness of uniform convergence

Definition 2.74 *Let $E \subseteq \mathbb{R}$ and let $\{f_n\}$ be a sequence of functions $E \to \mathbb{R}$. We say $\{f_n\}$ is* **uniformly Cauchy (on E)** *if $\forall\, \epsilon > 0$, $\exists\, N = N(\epsilon) \in \mathbb{N}$ s.t.*

$$m, n \geq N \Rightarrow |f_n(x) - f_m(x)| < \epsilon \text{ for all } x \in E.$$

Theorem 2.75 *A sequence $\{f_n\}$ of functions $E \to \mathbb{R}$ is uniformly Cauchy if and only if it is uniformly convergent.*

PROOF By the definition of uniformly Cauchy, each $\{f_n(x)\}$ is a Cauchy sequence of real numbers, hence converges to some $f(x)$ by completeness of $\mathbb{R}$. This defines a function $f : E \to \mathbb{R}$ so that $f_n \to f$ on E.

Now fix $\epsilon > 0$.

$\{f_n\}$ being uniformly Cauchy implies $\exists\, N = N(\epsilon)$ such that

$$m, n \geq N \text{ implies } |f_m(x) - f_n(x)| < \frac{\epsilon}{2} \text{ for all } x \in E.$$

That means that for all $x \in E$,

$$-\frac{\epsilon}{2} < f_m(x) - f_n(x) < \frac{\epsilon}{2}.$$

Fix n and take limits on all these terms as $m \to \infty$; since limits preserve soft inequalities we have

$$-\frac{\epsilon}{2} \leq f(x) - f_n(x) \leq \frac{\epsilon}{2}$$

which implies (for all $x \in E$) that

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

i.e. $|f_n(x) - f(x)| < \epsilon$ for all $x \in E$.

This proves $f_n \rightrightarrows f$ on E. □

2.10 Chapter 2 Summary

Note: items marked with a red star (★) are only in the long version of my lecture notes.

DEFINITIONS TO KNOW

Nouns

- $\mathbb{R}$ is the complete Archimedean ordered field.
- The **absolute value** $|x$ of $x \in \mathbb{R}$ is x if $x \geq 0$ and $-x$ if $x < 0$.
- The **distance** between x and y is $|x - y|$.
- An **upper bound** of a set $E \subseteq \mathbb{R}$ is a number B so that $\forall\, x \in E, x \leq B$.
  A **lower bound** of a set $E \subseteq \mathbb{R}$ is a number B so that $\forall\, x \in E, x \geq B$.
- The **supremum** of a set $E \subseteq \mathbb{R}$ is its least upper bound, i.e. a number s so that s is an upper bound of E and if t is any upper bound of E, then $t \geq s$.
  The **infimum** of a set $E \subseteq \mathbb{R}$ is its greatest lower bound, i.e. a number i so that i is a lower bound of E and if v is any lower bound of E, then $v \leq i$.
- A **subsequence** of $\{x_n\}_n$ is a sequence $\{x_{n_k}\}_k$ where $\{n_k\}$ is a strictly increasing sequence of indices.
- L is a **subsequential limit** of $\{x_n\}$ if $\exists$ subsequence $\{x_{n_k}\}$ with $x_{n_k} \overset{k\to\infty}{\longrightarrow} L$.
- (★) The **limit superior** of $\{x_n\}$ is $\overline{\lim}\, x_n = \lim\left(\sup\{x_m : m \geq n\}\right)$.
  (★) The **limit inferior** of $\{x_n\}$ is $\underline{\lim}\, x_n = \lim\left(\inf\{x_m : m \geq n\}\right)$.

Adjectives that describe subsets of $\mathbb{R}$ (including sequences)

- E is **bounded above** if it has an upper bound.
  E is **bounded below** if it has a lower bound.
  E is **bounded** if it is bounded above and bounded below (equivalently, if $\exists B \in \mathbb{R}$ so that $|x| \leq B$ for all $x \in E$).

Adjectives that describe sequences

- $\{x_n\}$ is **increasing** if $x_n \leq x_{n+1}$ for all n.
  $\{x_n\}$ is **decreasing** if $x_n \geq x_{n+1}$ for all n.
  $\{x_n\}$ is **monotone** if it is either increasing or decreasing.
- $\{x_n\}$ **converges** to L if $\forall \epsilon > 0$, $\exists\, N$ so that $n \geq N$ implies $|x_n - L| < \epsilon$.
- $\{x_n\}$ **diverges** if it does not converge to any $L \in \mathbb{R}$.
- $\{x_n\}$ is **Cauchy** if $\forall\, \epsilon > 0$, $\exists\, N$ so that $m, n \geq N$ implies $|x_m - x_n| < \epsilon$.

Adjectives that describe sequences of functions

- (★) $\{f_n\}$ **converges pointwise** to f on E if $\forall x \in E$, $f_n(x) \to f(x)$.
- (★) $\{f_n\}$ **converges uniformly** to f on E if $\forall \epsilon > 0$, $\exists N$ so that $n \geq N$ implies $|f_n(x) - f(x)| < \epsilon \, \forall x \in E$.
- (★) $\{f_n\}$ is **uniformly Cauchy** on E if $\forall \epsilon > 0$, $\exists N$ so that $m, n \geq N$ implies $|f_n(x) - f_m(x)| < \epsilon \, \forall x \in E$.

Theorems with names

Triangle inequality (△ ineq):

$|x + y| \leq |x| + |y|$;

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

Completeness of $\mathbb{R}$: A sequence of real numbers converges if and only if it is Cauchy. (Convergent sequences are always Cauchy, but the converse isn't true for the rational numbers.)

Main Limit Theorem: Limits are preserved under arithmetic (so is the Cauchy property).

Squeeze Theorem:

If $x_n \leq y_n \leq z_n$ and $\lim x_n = \lim z_n = L$, then $y_n \to L$.

If $|x_n - L| \leq a_n$ and $a_n \to 0$, then $x_n \to L$.

Reversing Lemma: $-\sup E = \inf(-E)$ and $-\inf E = \sup(-E)$.

Supremum Property: If nonempty $E \subseteq \mathbb{R}$ is bounded above, then $\sup E$ exists.

Infimum Property: If nonempty $E \subseteq \mathbb{R}$ is bounded below, then $\inf E$ exists.

Monotone Convergence Theorem (MCT): If $\{x_n\}$ is increasing and bounded above, then $x_n \to \sup\{x_n\}$. If $\{x_n\}$ is decreasing and bounded below, then $x_n \to \inf\{x_n\}$.

Archimedean Properties:

I. $\forall x \in \mathbb{R}$, $\exists n \in \mathbb{N}$ so that $n > x$.

II. $\forall x > 0$, $\exists n \in \mathbb{N}$ so that $\frac{1}{n} < x$.

III. $\forall x > 0$, $\exists n \in \mathbb{N}$ so that $n \leq x < n + 1$.

Density Theorem: If $a < b$, then the interval (a, b) contains both a rational number and an irrational number.

Monotone Subsequence Theorem (MST): Every sequence has a monotone subsequence.

Bolzano-Weierstrass Theorem (B-W): Every bounded sequence has a convergent subsequence.

Avoidance Theorem: If $x_n \not\to L$, then $\exists\ \epsilon_0 > 0$ and $\exists$ subsequence $\{x_{n_k}\}$ so that $|x_{n_k} - L| \geq \epsilon_0$ for all k.

OTHER THEOREMS TO REMEMBER

- Convergent sequences (so also Cauchy sequences) are bounded.
- Limits preserve soft inequalities $\leq$ and $\geq$.
- If $E \subseteq \mathbb{R}$ is bounded above, then $\exists\ \{x_n\} \subseteq E$ so that $x_n \to \sup E$.

  If $E \subseteq \mathbb{R}$ is bounded below, then $\exists\ \{x_n\} \subseteq E$ so that $x_n \to \inf E$.
- If $x_n \to L$, then any subsequence $\{x_{n_k}\}$ also converges to L.
- A sequence with two different subsequential limits must diverge.
- If the subsequential limit set of a bounded sequence consists of a single number, then the sequence converges to that number.
- (★) If $\overline{\lim}\, x_n \leq \underline{\lim}\, x_n$, then $\overline{\lim}\, x_n = \underline{\lim}\, x_n$ and $\{x_n\}$ converges to this common value.
- (★) A sequence of functions that converges uniformly must converge pointwise.
- (★) A sequence of functions converges uniformly if and only if it is uniformly Cauchy.

STANDARD PROOF TECHNIQUES

To prove that $\{x_n\}$ converges, do <u>one</u> of these things:

1. Apply the MCT.
2. Apply the Main Limit Theorem.
3. Use the Squeeze Theorem (usually, this means showing $|x_n - L| \leq a_n$ for some $\{a_n\}$ where $a_n \to 0$).
4. Prove it directly (let $\epsilon > 0$; from scratch work figure out N so that $n \geq N$ implies $|x_n - L| < \epsilon$).
5. Prove $\{x_n\}$ is Cauchy.
6. (★) Prove $\overline{\lim}\, x_n \leq \underline{\lim}\, x_n$.

7. Show the sequence is bounded, and that every subsequence $\{x_{n_k}\}$ converges to L.
8. Show $\{x_n\}$ is a subsequence of a convergent sequence.

To prove that $\{x_n\}$ **diverges**, do one of these things:

1. Show $\{x_n\}$ is unbounded.
2. Show $\{x_n\}$ has two different subsequential limits (or that a subsequence of $\{x_n\}$ diverges).
3. Prove $\{x_n\}$ isn't Cauchy.

To prove $s = \sup E$, do both of these things:

1. Show s is an upper bound of E (let $x \in E$ and argue why $x \leq s$).
2. Show s is the least upper bound, by doing one of these things:
   a) Assume t is an upper bound of E and proving $s \leq t$.
   b) Show $s \in E$.
   c) Let $\epsilon > 0$ and from scratch work, find a number $x \in E \cap (s - \epsilon, s]$.

To prove $i = \inf E$, do both of these things:

1. Show i is an lower bound of E (let $x \in E$ and argue why $x \geq i$).
2. Show i is the greatest lower bound, by doing one of these things:
   a) Assume v is a lower bound of E and proving $i \geq v$.
   b) Show $i \in E$.
   c) Let $\epsilon > 0$ and from scratch work, find a number $x \in E \cap [i, i + \epsilon)$.

2.11 Chapter 2 Homework

Exercises from Section 2.1

1. Prove that there is no rational number x so that $x^2 = 5$.

2. Prove that there is no rational number x so that $2^x = 3$. (You may not assume anything about logarithms in this problem.)

3. Prove that there is no total ordering on the field $\mathbb{C}$ of complex numbers which makes $\mathbb{C}$ into an ordered field.

   Hint: Suppose that there is a total ordering $\leq$ on $\mathbb{C}$ which makes $\mathbb{C}$ into an ordered field. Derive a contradiction, starting with the observation that either $i > 0$ or $i < 0$.

4. Prove that in any ordered field, $1 > 0$. (The trick here is not to assume what you are to prove–only use facts about ordered fields given in §2.1.)

5. Consider the equation $|x - 11| < |x + 5|$. Rather than solving this equation algebraically, let's think about it this way: if x is a solution of this equation, that means the distance from x to $\boxed{}$ is less than the distance from x to $\boxed{}$? Draw a number line and think about the set of x for which this holds; that's the solution of the equation. Write that solution set as an inequality.

6. Describe the solution set of $|x + 12| > |x + 4|$.

Exercises from Section 2.2

7. Prove Theorem 2.14, which says that for $x, y \in \mathbb{R}$, $|x\,y| = |x|\,|y|$.

8. Prove the fourth statement of Theorem 2.16, which says that for $x, y, r \in \mathbb{R}$, $|rx - ry| = |r|\,|x - y|$.

9. Let $x, y \in \mathbb{R}$. Prove $||x| - |y|| \leq |x - y|$.

10. Classify each of these sets as bounded or unbounded:

    a) $[3, \infty)$
    b) $\mathbb{Q} \cap (-3, 5)$
    c) $\{3^n : n \in \mathbb{Z}\}$
    d) $\{2^{-n} : n \in \mathbb{N}\}$
    e) $\{2^{-n} : n \in \mathbb{Z}\}$
    f) $\varnothing$

Exercises from Section 2.3

11. For each given sequence,

    i. Determine whether or not the sequence is bounded. If it is bounded, give an explicit bound.

    ii. Determine if the sequence is monotone; if it is, classify it as increasing or decreasing.

    iii. If the sequence is monotone, determine whether it is strictly increasing/decreasing.

    a) $\left\{\dfrac{n+1}{n}\right\}_{n=1}^{\infty}$

    b) $\left\{(-2)^{\frac{1}{2}(n^2+n)}\right\}_{n=1}^{\infty}$

    c) $\left\{\cos\dfrac{\pi n}{2}\right\}_{n=0}^{\infty}$

    d) $\{a_n\}_{n=1}^{\infty}$, where $\{a_n\}$ is the largest rational number with denominator $\leq n$ such that $a_n < \pi$

    e) $\{b_n\}_{n=1}^{\infty}$, where $\{b_n\}$ is the largest rational number with denominator n such that $b_n < \pi$

12. Let $q_0 = \dfrac{1}{2}$ and for each $n \geq 1$, set $q_n = \dfrac{1}{q_{n-1}+2}$.

    a) Simplify q_1, q_2, q_3, and q_4.

    b) Prove that $\{q_n\}$ is bounded.

    Hints: Clearly $q_n \geq 0$ for all n. We claim that $q_n \leq \dfrac{1}{2}$ for all n. To prove this claim, suppose not; then let N be the smallest index so that $q_N > \dfrac{1}{2}$. Show that $N \neq 0$, then $N \geq 1$ so $N-1 \in \mathbb{N}$. Since N is the smallest index so that $q_N > \dfrac{1}{2}$, it must be that $q_{N-1} \leq \frac{1}{2}$. Explain why these last two inequalities contradict one another.

13. Let $x_n = \dfrac{3n+2}{n-1}$. Prove that $\{x_n\}$ converges (using only the definition of convergence, not any theorems that follow later in the notes).

14. Use the definition of convergence to prove that $\dfrac{2n}{n+1} \to 2$.

15. Use the definition of convergence to prove that $\dfrac{n^2-1}{2n^2+3} \to \dfrac{1}{2}$.

16. Let $x_n = \dfrac{1}{n^3}$.

    a) Prove $\{x_n\}$ converges using only the Main Limit Theorem and the fact that $\dfrac{1}{n} \to 0$.

    b) Prove $\{x_n\}$ converges directly, using an ϵ-proof.

17. Let $\{x_n\}$ and $\{y_n\}$ be two sequences of real numbers where $\{x_n\}$ converges but $\{y_n\}$ diverges. Prove that $\{x_n + y_n\}$ diverges.

    Hint: A proof by contradiction is short (use the fact that the difference of two convergent sequences converges).

18. Prove the Squeeze Theorem (Theorem 2.29, which says that if $\{x_n\}, \{y_n\}$ and $\{z_n\}$ are sequences such that $x_n \le y_n \le z_n$ for all n, and if $\lim x_n = \lim z_n = L$, then $y_n \to L$.

Exercises from Section 2.4

19. Prove Theorem 2.35, which says that for any non-negative real number x, there is another real number $\sqrt{x}$ such that $\left(\sqrt{x}\right)^2 = x$.

20. Prove the first three statements of Theorem 2.38 (without using completeness).

21. Prove or disprove: if $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences of real numbers, then $\left\{\dfrac{x_n}{y_n}\right\}$ is a Cauchy sequence.

22. Give an explicit example of a sequence $\{x_n\}$ of real numbers such that the sequence $\{|x_n|\}$ converges but the sequence $\{x_n\}$ diverges.

Exercises from Section 2.5

23. Prove the third statement of Lemma 2.42, which says that if $E \subseteq \mathbb{R}$ is bounded below by B, then $-E$ is bounded above by $-B$.

24. Prove the second statement of the Reversing Lemma (Lemma 2.45), which says that if $i = \inf E$, then $-i = \sup(-E)$.

25. Formulate and prove a lemma analagous to Lemma 2.51 for infima, rather than suprema.

26. Prove the first statement of Lemma 2.53, which says that if $E \subseteq \mathbb{R}$ is a set which is bounded above, then there exists an increasing sequence $\{x_n\}$ of points in E with $x_n \to \sup E$.

Hint: For each n, apply Lemma 2.51 with $\epsilon = \frac{1}{n}$. This defines a sequence $\{x_n\}$; prove $x_n \to \sup E$.

27. Prove the first statement of Lemma 2.52, which says that If $E \subseteq \mathbb{R}$ is bounded above, then $\sup S = \inf\{t : t \text{ is an upper bound for } S\}$.

28. Let S_1 and S_2 be any two subsets of $\mathbb{R}$. Define the *sum* of these sets to be the set

$$S_1 + S_2 = \{s_1 + s_2 : s_1 \in S_1, s_2 \in S_2\}$$

and the *difference* of these two sets to be the set

$$S_1 - S_2 = \{s_1 - s_2 : s_1 \in S_1, s_2 \in S_2\}.$$

a) Suppose $S_1 \subseteq \mathbb{R}$ and $S_2 \subseteq \mathbb{R}$ are both bounded above. Show that $S_1 + S_2$ is bounded above, and prove or disprove: $\sup(S_1 + S_2) = \sup S_1 + \sup S_2$.

b) Suppose $S_1 \subseteq \mathbb{R}$ and $S_2 \subseteq \mathbb{R}$ are both bounded. Show that $S_1 - S_2$ is bounded, and prove or disprove: $\sup(S_1 - S_2) = \sup S_1 - \sup S_2$.

Exercises from Section 2.6

29. Prove the third Archimedean Property (Theorem 2.59, which says that if $x \in (0, \infty)$, then there is $n \in \mathbb{N}$ such that $n \leq x < n + 1$.

30. Prove that for any $x \in \mathbb{R}$, there is $n \in \mathbb{Z}$ such that $n \leq x < n + 1$.

    Hint: If $x > 0$, this follows from the third Archimedean Property (preceding exercise). Prove two other cases: $x = 0$ and $x < 0$. For $x < 0$, apply the third Archimedean Property to $-x$.

31. Prove the second statement of the Density Theorem (Theorem 2.60), which says that if $a < b$, then there exists $x \in \mathbb{R} - \mathbb{Q}$ so that $a < x < b$.

    Hint: Use the first part of the Density Theorem to find a rational number in the interval $(a + \sqrt{2}, b + \sqrt{2})$.

32. Consider the sequence $\{x_n\}$ of real numbers defined recursively by setting $x_1 = 2$ and then defining $x_{n+1} = 2 - \frac{1}{x_n}$ for all $n \geq 1$.

    a) Write out the first five terms of this sequence.

    b) Prove that $\{x_n\}$ converges.
    Hint: The Monotone Convergence Theorem may be helpful.

33. Consider the sequence $\{x_n\}$ of real numbers defined recursively by setting $x_1 = 2$ and then defining $x_{n+1} = \sqrt{x_n + 3}$ for all $n \geq 1$. Prove that $\{x_n\}$ converges.

34. Consider the sequence $\{x_n\}$ where $x_n = \dfrac{n^2+2}{n^2+4}$.

   a) Prove $\{x_n\}$ converges using the MCT.

   b) Prove $\{x_n\}$ converges by establishing the inequality $1 - \dfrac{2}{n^2} \le x_n \le 1$ and applying the Squeeze Theorem.

   c) Prove $\{x_n\}$ converges directly (using an ϵ-proof).

   d) Prove $\{x_n\}$ converges by rewriting it with some algebra and applying the Main Limit Theorem (other than the Main Limit Theorem, assume nothing other than $\dfrac{1}{n} \to 0$).

35. Prove that for any $x \in \mathbb{R}$, there is a sequence $\{x_n\}$ of rational numbers that converges to x.

36. Let $\{x_n\}$ be a sequence of positive rational numbers with $\dfrac{x_{n+1}}{x_n} \to L$.

   a) Show that if $\{x_n\}$ converges, then $L \le 1$.
   Hint: Prove this by contradiction: assume $L > 1$ and show the sequence is unbounded.

   b) Prove that if $L < 1$, then $\{x_n\}$ converges. To what does $\{x_n\}$ converge?

   c) Prove, by constructing examples, that if $L = 1$, it is possible for $\{x_n\}$ to converge, and possible for $\{x_n\}$ to diverge.

37. Let $\{x_n\}$ be a sequence of real numbers. We say that $\{x_n\}$ **Cesàro converges (to L)** if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} x_k = L.$$

(The notion of Cesàro convergence is useful in the study of Markov chains in MATH 416.)

   a) Prove that if $x_n \to L$, then $\{x_n\}$ Cesàro converges to L.
   Hint: Let $\epsilon > 0$ and choose N so that $n \ge N$ implies $|x_n - L| < \epsilon$. Now, for $n \ge N$ take $\sum_{k=1}^{n} x_k$ and split this sum into two parts: the terms from $k = 1$ to N plus the terms from $k = N+1$ to n. After dividing by n, the first part clearly converges to something. Bound the second part based on the fact that $|x_n - L| < \epsilon$; this will show that $\left| \dfrac{1}{n} \sum_{k=1}^{n} x_k - L \right| < \epsilon$.

   b) Give an example of a sequence $\{x_n\} \subseteq \mathbb{R}$ which Cesàro converges but diverges.
   Hint: Look at our prototype examples of sequences.

c) Give an example of a bounded sequence which does not Cesáro converge.
Hint: Build a sequence that starts with a 1, then has some 0s, then has some 1s, then some 0s, etc. If you choose the right number of 0s (and 1s) in each block of terms, you can force the sequence $\left\{\frac{1}{n}\sum_{k=1}^{n} x_k\right\}$ to have subsequences converging to two different limits.

38. Prove **Polya's Lemma**, which says that if $\{x_n\} \subseteq \mathbb{R}$ is a subadditive sequence of nonnegative numbers (*subadditive* means that for all $m, n \in \mathbb{N}$, $x_{m+n} \leq x_m + x_n$), then the sequence $\left\{\frac{1}{n}x_n\right\}$ converges.
Hint: First, use the subadditivity to show that $x_n \leq nx_1$ for all n. Then, use that to show the sequence $\left\{\frac{1}{n}x_n\right\}$ is decreasing. Since the sequence $\{x_n\}$ is assumed non-negative, the MCT applies.

Exercises from Section 2.7

39. Prove the Bolzano-Weierstrass Theorem (Theorem 2.64), which says that every bounded sequence of real numbers has a convergent subsequence.

Exercises from Section 2.8

40. For each of the following sequences, find $\overline{\lim}\, x_n$ and $\underline{\lim}\, x_n$ (no proofs required; just write the answers):

a) $x_n = \dfrac{n+4}{3n-2}$

b) $x_n = 5 + 3 \cdot (-1)^n$

c) $x_n = \sin\left(\dfrac{n\pi}{4}\right)$

d) $x_n = \begin{cases} 3 & \text{if } n \text{ is odd} \\ 3 + e^{-n} & \text{if } n \text{ is even} \end{cases}$

e) $x_n = \begin{cases} \dfrac{1}{n} & \text{if } n \text{ is prime} \\ 1 - \dfrac{1}{n} & \text{if } n \text{ is not prime} \end{cases}$

f) $x_n = \begin{cases} \dfrac{1}{n} & \text{if } n \leq 100 \\ 1 - \dfrac{1}{n} & \text{if } n > 100 \end{cases}$

41. Prove the second statement of Theorem 2.68, which says that for a bounded sequence $\{x_n\}$ of real numbers, $\underline{\lim}\, x_n$ exists.

42. Prove the second statement of Theorem 2.69, which says that for a bounded sequence $\{x_n\}$ of real numbers with subsequential limit set S, $\underline{\lim}\, x_n = \inf S$.

43. Let $\{x_n\}$ and $\{y_n\}$ be two bounded sequences of real numbers.

a) Prove that $\overline{\lim}\,(x_n + y_n) \leq \overline{\lim}\, x_n + \overline{\lim}\, y_n$.

b) Show by giving an explicit counterexample that it is *not necessarily the case* that $\overline{\lim}\,(x_n + y_n) = \overline{\lim}\,x_n + \overline{\lim}\,y_n$.

Exercises from Section 2.9

44. Let $\{f_n\}$ and $\{g_n\}$ be sequences of functions, each taking $\mathbb{R}$ to $\mathbb{R}$, with $f_n \rightrightarrows f$ and $g_n \rightrightarrows g$. Prove $f_n + g_n \rightrightarrows f + g$.

45. Let $\{f_n\}$ be a sequence of functions $\mathbb{R} \to \mathbb{R}$, with $f_n \rightrightarrows f$. Prove $rf_n \rightrightarrows rf$ for any constant $r \in \mathbb{R}$.

46. Show by constructing a specific counterexample that if $\{f_n\}$ and $\{g_n\}$ are sequences of functions, each taking $\mathbb{R}$ to $\mathbb{R}$, with $f_n \rightrightarrows f$ and $g_n \rightrightarrows g$, it is not necessarily the case that $f_n g_n \rightrightarrows fg$.

47. Let $\{f_n\}$ be the sequence of functions $[0,1] \to \mathbb{R}$ defined by $f_n(x) = \dfrac{x}{1+nx^2}$. Find the pointwise limit f of $\{f_n\}$ and determine whether or not $\{f_n\}$ converges uniformly to f on $[0,1]$.

48. Let $\{g_n\}$ be the sequence of functions $[0,1] \to \mathbb{R}$ defined by $g_n(x) = \dfrac{nx}{1+nx^2}$. Find the pointwise limit g of $\{g_n\}$ and determine whether or not $\{g_n\}$ converges uniformly to g on $[0,1]$.

49. Let $\{h_n\}$ be the sequence of functions $[0,1] \to \mathbb{R}$ defined by $h_n(x) = \dfrac{nx}{1+n^2x^2}$. Find the pointwise limit h of $\{h_n\}$ and determine whether or not $\{h_n\}$ converges uniformly to h on $[0,1]$.

Chapter 3

Topology of $\mathbb{R}$

3.1 Open and closed sets

Loosely speaking, *topology* is a branch of mathematics which is sort of like "abstract geometry": it studies the properties of shapes and other sets that are preserved under stretching, compressing, shifting, twisting, and other "continuous" deformations of the space.
This subject depends on the notion of an *open subset* of a space, and the idea is that the sets which are "open" remain open when you stretch/rotate/twist the space.

In this course, we care most about calculus, not topology. But you can't really do calculus without understanding some of the topology of $\mathbb{R}$, so we'll discuss some fundamental topological concepts in this chapter.

Open balls

Definition 3.1 *Let $x \in \mathbb{R}$ and $\epsilon > 0$. The set*

$$B_\epsilon(x) = \{y \in \mathbb{R} : |y - x| < \epsilon\}$$

is called the **open ball (of radius ϵ, centered at x)**.

EXAMPLES

$B_3(8) =$ $\mathbb{R}$

$B_{1/6}(0) =$ $\mathbb{R}$

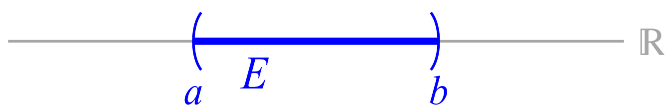
Lemma 3.2 (Open balls are bounded open intervals) *A set $E \subseteq \mathbb{R}$ is an open ball if and only if $E = (a, b)$ for real numbers a and b with $a < b$.*

PROOF (⇒) Suppose E is an open ball, i.e. $B = B_\epsilon(x)$ for some $x \in \mathbb{R}$ and $\epsilon > 0$. Then

$$E =$$

as wanted.

(⇐) Suppose $E = (a, b)$.



Then $E = B_\epsilon(x)$ for

$x =$ and $\epsilon =$

so E is an open ball, as desired. □

Open sets

Definition 3.3 *Let $E \subseteq \mathbb{R}$. E is called* **open** *if for every $x \in E$, there is $\epsilon > 0$ such that $B_\epsilon(x) \subseteq E$.*

Idea: Open sets are those where every point in the set has "room to breathe" without leaving the set.

EXAMPLES

$E = (0, 1)$

$E = [0, 1)$

Other open sets:

Other sets that are not open:

Theorem 3.4 (Open balls are open sets) *Let $E \subseteq \mathbb{R}$ be an open ball. Then E is an open set.*

PROOF HW (this is similar to the example $E = (0, 1)$ shown on the previous page).

Theorem 3.5 (Unions of open sets are open) *Let $\{E_\alpha\}$ be a collection of open sets (finite, countable or uncountably many sets). Then $\bigcup_\alpha E_\alpha$ is open.*

PROOF Let $x \in \bigcup_\alpha E_\alpha$.

Then $x \in E_\alpha$ for some α.

Since E_α is open, $\exists\, \epsilon > 0$ s.t. $B_\epsilon(x) \subseteq E_\alpha \subseteq E$. □

Theorem 3.6 (Intersections of finitely many open sets are open) *Let $E_1, \ldots, E_n$ be subsets of $\mathbb{R}$, each of which is open. Then $\bigcap_{k=1}^{n} E_k$ is open.*

PROOF Let $x \in \bigcap_{k=1}^{n} E_k$.

That means $x \in E_k$ for all $k \in \{1, 2, \ldots, n\}$.

Since each E_k is open, for each k, $\exists\, \epsilon_k > 0$ s.t. $B_{\epsilon_k}(x) \subseteq E_k$.

Let $\epsilon = \min\{\epsilon_1, \epsilon_2, \ldots, \epsilon_n\}$.

Then if $y \in B_\epsilon(x)$, $y \in B_\epsilon(x) \subseteq B_{\epsilon_k}(x) \subseteq E_k$ for all k.

Thus $B_\epsilon(x) \subseteq \bigcap_{k=1}^{n} E_k$. □

WARNING: An intersection of countably infinitely many open sets may not be open. As an example, let

$$E_n = \left(-\frac{1}{n}, \frac{1}{n}\right).$$

Each E_n is open, but $\bigcap_{n=1}^{\infty} E_n = \{0\}$, which is not open.

TO PROVE A SET E IS OPEN:

Let $x \in E$.

Write down a formula for $\epsilon > 0$ (often coming from some scratch work).

Prove that $B_\epsilon(x) \subseteq E$. □

Alternatively, prove E is the union of sets already known to be open (such as open balls). □

Alternatively, prove E is the intersection of finitely many sets already known to be open. □

Closed sets

Definition 3.7 *A set $E \subseteq \mathbb{R}$ is called* **closed** *if E^C is open.*

WARNING: Sets are not doors!

Doors are open, or closed, but never both and never neither.

Sets can be open, closed, both (we use the word **clopen** for this) or neither.

"closed" does NOT mean "not open".

EXAMPLES

$(0,1)^C =$

$[a, b]$

$\{x\}$

$\mathbb{R}$

$\varnothing$

$\mathbb{Q}$

$(0,1)$

Theorem 3.8 (Intersections of closed sets are closed) *Let $\{E_\alpha\}$ be a collection of closed sets (finite, countable or uncountably many sets). Then $\bigcap_\alpha E_\alpha$ is closed.*

PROOF Since each E_α is closed, each E_α^C is open.

Since unions of open sets are open, $\bigcup_\alpha E_\alpha^C$ is open.

Thus $\bigcap_\alpha E_\alpha = \left[\bigcup_\alpha E_\alpha^C\right]^C$ is the complement of an open set, hence closed. □

Theorem 3.9 (Unions of finitely many closed sets are closed) *Let $E_1, \ldots, E_n$ be subsets of $\mathbb{R}$, each of which is closed. Then $\bigcup_{k=1}^n E_k$ is closed.*

PROOF Since each E_k is closed, each E_k^C is open.

Since an intersection of finitely many open sets is open, $\bigcap_{k=1}^n E_k^C$ is open.

So $\bigcup_{k=1}^n E_k = \left[\bigcap_{k=1}^n E_k^C\right]^C$ is the complement of an open set, hence closed. □

Sequential closedness

Definition 3.10 *Let $E \subseteq \mathbb{R}$. E is called* **sequentially closed** *if for every sequence $\{x_n\}$ of numbers in E that converges to $x \in \mathbb{R}$, it must be the case that $x \in E$.*

EXAMPLES

$(0, 1)$

$\mathbb{Q}$

$[0, 1]$

Theorem 3.11 (Closed sets are the same as sequentially closed sets) *Let $E \subseteq \mathbb{R}$. Then E is closed if and only if E is sequentially closed.*

PROOF (⇒) Suppose E is closed.

To show E is sequentially closed, we let $\{x_n\} \subseteq E$ be a convergent sequence, with $x_n \to x$ where $x \in \mathbb{R}$; we need to show $x \in E$.

We will prove this by contradiction: suppose not, i.e. $x \notin E$, i.e. $x \in E^C$.

Since E^C is open, $\exists$ ____________ s.t. _________________ $\subseteq E^C$.

Since $x_n \to x$, $\exists$ _________ s.t. ________________________ $\Rightarrow x_N \in B_\epsilon(x) \subseteq E^C$.

Contradiction! $x_N \in E^C$, but $\{x_n\} \subseteq E$.

Therefore $x \in E$, meaning E is sequentially closed.

(⇐) Suppose E is sequentially closed.

To show E is closed, we will show E^C is open. So let $x \in E^C$.

Again we argue by contradiction.

Suppose there is no $\epsilon > 0$ such that $B_\epsilon(x) \subseteq E^C$.

Then, for every n, $\exists\, x_n \in B_{1/n}(x) - E^C$, meaning $x_n \in E$ and $|x_n - x| < \dfrac{1}{n}$.

By the ________________________, $x_n \to x$, and since E is assumed to be sequentially closed, $x \in E$.

This is a contradiction, so $\exists\, \epsilon > 0$ s.t. $B_\epsilon(x) \subseteq E^C$.

This makes E^C open, so therefore E is closed. □

Theorem 3.12

Let $E \subseteq \mathbb{R}$ be open. Then $\sup E \notin E$ and $\inf E \notin E$.

Let $F \subseteq \mathbb{R}$ be a closed set which is bounded above. Then $\sup F \in F$.

Let $F \subseteq \mathbb{R}$ be a closed set which is bounded below. Then $\inf F \in F$.

PROOF HW

TO PROVE A SET E IS CLOSED:

Prove E^C is open (see above). □

Alternatively, prove E is sequentially closed:
Let $\{x_n\} \subseteq E$ be s.t. $x_n \to x$, and prove $x \in E$. □

Alternatively, prove E is the intersection of sets already known to be closed. □

Alternatively, prove E is the union of finitely many sets already known to be closed. □

3.2 Intervals, betweenness and connectedness

In Chapter 2 we defined *intervals* by giving a laundry list of the types of sets that we call intervals:

QUESTION

Is there an *invariant* of those sets, i.e. a property that those sets have that other sets don't have (apart from being an "interval")?

To answer this question, let's look at an example of a set that isn't an interval:

$$E = (0,1) \cup (2,3)$$

Heuristically, what makes this E not be an interval?

This idea is captured formally with the concept of *betweenness*:

Betweenness

Definition 3.13 *Let $E \subseteq \mathbb{R}$. We say E has the* **betweenness property** *if for every $y, z \in E$ with $y \leq z$, $[y,z] \subseteq E$.*

EXAMPLE

$E = (0,1) \cup (2,3)$ does not have betweenness:

Lemma 3.14 *Let $E \subseteq \mathbb{R}$. E is an interval if and only if E has the betweenness property.*

PROOF ($\Rightarrow$) Suppose $E \subseteq \mathbb{R}$ is an interval.

We prove this with 10 cases, depending on what type of interval E is:

Case 1: If $E = \emptyset$, then E has betweenness vacuously.

Case 2: If $E = \mathbb{R}$, then E obviously has betweenness, since every $[y, z]$ is a subset of $\mathbb{R}$.

Case 3: If $E = [a, b]$, then let $y, z \in E$ with $y \leq z$. This means $a \leq y \leq z \leq b$.
To verify that E has betweenness, we need to show $[y, z] \subseteq E$.
Toward that end, let $x \in [y, z]$. Then $y \leq x \leq z$, so $a \leq x \leq b$, so $x \in [a, b] = E$ as wanted.

Case 4: If $E = (a, b]$, repeat Case 3, changing each red $[$ to $($ and each red $\leq$ to $<$.

Case 5: If $E = [a, b)$, repeat Case 3, changing each green $]$ to $)$ and each green $\leq$ to $<$.

Case 6: If $E = (a, b)$, repeat Case 3, making the changes indicated in both Cases 4 and 5.

Case 7: If $E = [a, \infty)$, let $y, z \in E$ with $y \leq z$. This means $a \leq y \leq z$.
As above, we need to show $[y, z] \subseteq E$, so let $x \in [y, z]$.
Then $a \leq y \leq x \leq z$, so $a \leq x$, so $x \in [a, \infty) = E$.

Case 8: If $E = (a, \infty)$, repeat Case 7, changing each red $[$ to $($ and each red $\leq$ to $<$.

Case 9: If $E = (-\infty, b]$, let $y, z \in E$ with $y \leq z$. This means $y \leq z \leq b$.
As above, we need to show $[y, z] \subseteq E$, so let $x \in [y, z]$.
Then $y \leq x \leq z \leq b$, so $x \leq b$, so $x \in (-\infty, b] = E$.

Case 10: If $E = (-\infty, b)$, repeat Case 9, changing each green $]$ to $)$ and each green $\leq$ to $<$.

($\Leftarrow$) Suppose $E \subseteq \mathbb{R}$ has the betweenness property.

Claim: $(\inf E, \sup E) \subseteq E \subseteq [\inf E, \sup E]$.
Once the claim is proven, it follows that E must be an interval.

Proof of Claim: For the first inclusion, let $x \in (\inf E, \sup E)$.
Then $x > \inf E$ and $x < \sup E$.

By a characterization of inf and sup,
$\exists\, y \in [\inf E, x) \bigcap E$ and $\exists\, z \in (x, \sup E] \bigcap E$.

By the betweenness property, $[y, z] \subseteq E$.
We have $x \in [y, z] \subseteq E$.
For the second inclusion ($E \subseteq [\inf E, \sup E]$), let $x \in E$.
Then, $\inf E \le x \le \sup E$, so $x \in [\inf E, \sup E]$.
This proves the claim, and therefore the theorem. □

Connectedness

Definition 3.15 *Let $E \subseteq \mathbb{R}$. A* **disconnection** *of E is a pair of sets U and V with all four of these properties:*

1. U and V are open;
2. U and V are disjoint, *i.e.* $U \cap V = \varnothing$;
3. U and V both hit E, *meaning* $U \cap E \ne \varnothing$ *and* $V \cap E \ne \varnothing$;
4. U and V cover E, *meaning* $E \subseteq U \cup V$.

$E \subseteq \mathbb{R}$ is called **connected** *if it does not have a disconnection.*
E is called **disconnected** *if it has a disconnection.*

EXAMPLES

$E = (0, 1) \cup (2, 3)$

$F = \mathbb{R} - \{3\}$

$\mathbb{Q}$

$\mathbb{R} - \mathbb{Q}$

$[0, 1]$

Lemma 3.16 *If $E \subseteq \mathbb{R}$ is connected, then E has betweenness.*

PROOF We prove the contrapositive. Suppose E does not have betweenness.
That means $\exists y, z \in E$ with $y \leq z$ but $[y, z] \nsubseteq E$.
Therefore, $\exists x \in [y, z]$ with $x \notin E$.

Now let $U = (-\infty, x)$ and $V = (x, \infty)$.

1. U and V are open (HW);
2. U and V are disjoint;
3. U and V both hit E, since $y \in U \cap E$ and $z \in V \cap E$; and
4. U and V cover E, since $E \subseteq \mathbb{R} - \{x\} \subseteq U \cup V$.

Therefore $\{U, V\}$ is a disconnection of E. □

PREVIEW

We will see later that the converse of Lemma 3.16 is true, i.e. if $E \subseteq \mathbb{R}$ has betweenness, then E is connected. But proving this is harder, because to prove a set *is* connected requires ruling out *all possible disconnections*. To do this, we need a way of describing all open sets in $\mathbb{R}$, since a disconnection is a pair of open sets.

Fortunately there is a theorem that describes all open subsets of $\mathbb{R}$:

Theorem 3.17 (Lindelöf's Theorem) *A subset of $\mathbb{R}$ is open if and only if it is the union of countably many disjoint open intervals.*

To prove Lindelöf's Theorem, we need to discuss *equivalence relations*. Recall that a relation on a set E is a symbol you put between two elements of E that produces a true or false statement.

Definition 3.18 *Let E be a set. A relation $\sim$ on E is called an* **equivalence relation** *if it has three properties:*

1. $\sim$ is reflexive: $\forall\, x \in E$, $x \sim x$.
2. $\sim$ is symmetric: $\forall\, x, y \in E$, $x \sim y$ *implies* $y \sim x$.
3. $\sim$ is transitive: $\forall x, y, z \in E$, $x \sim y$ *and* $y \sim z$ *imply* $x \sim z$.

A prototype example of an equivalence relation is "=".

Definition 3.19 *Let $\sim$ be an equivalence relation on set E.*
For any $x \in E$, the **equivalence class** *of x, denoted $[x]$, is the set of all $y \in E$ such that $y \sim x$.*

Lemma 3.20 *Given any equivalence relation on any set E:*

1. *any two equivalence classes either coincide or are disjoint; and*
2. *the union of all the equivalence classes is E.*

PROOF To prove (1), suppose $[x]$ and $[y]$ are not disjoint; then they both contain some $z \in E$.

Since $z \in [x]$, $z \sim x$, and since $z \in [y]$, $z \sim y$. So by transitivity, $x \sim y$.

Now, if $a \in [x]$, $a \sim x$; by transitivity $a \sim y$ so $a \in [y]$. This proves $[x] \subseteq [y]$.

At the same time, if $a \in [y]$, $a \sim y$; by transitivity $a \sim x$ so $a \in [x]$. This proves $[y] \subseteq [x]$.

Therefore $[x] = [y]$.

For statement (2),

Clearly, $[x] \subseteq E$ so $\bigcup_{x \in E} [x] \subseteq E$.

By reflexivity, $x \sim x$ so $x \in [x]$. Thus $E = \bigcup_{x \in E} \{x\} \subseteq \bigcup_{x \in E} [x]$.

Therefore $E = \bigcup_{x \in E} [x]$. □

EXAMPLE

Consider "$\equiv_3$", the relation on $\mathbb{Z}$ denoting congruence $\mod 3$ (this means we say $x \equiv_3 y$ if x and y have the same remainder when divided by 3).

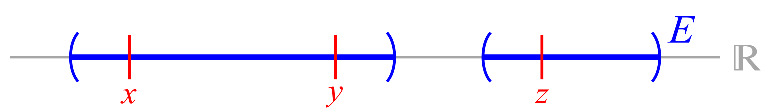
The concept is that an equivalence relation on a set E partitions E into disjoint equivalence classes. With that in mind, we prove Lindelöf's Theorem by defining an equivalence relation on an arbitrary open set E and showing that the equivalence classes are open intervals. Now for the details:

PROOF OF LINDELÖF'S THEOREM

($\Leftarrow$) This is immediate, since unions of open sets are open.

($\Rightarrow$) Let $E \subseteq \mathbb{R}$ be open. Define a relation $\sim$ on E by saying

$$x \sim y \Leftrightarrow [\min\{x,y\}, \max\{x,y\}] \subseteq E.$$



Claim 1: $\sim$ is an equivalence relation.

Proof of Claim 1: That $\sim$ is reflexive and symmetric is obvious.

To prove transitivity, suppose $x \sim y$ and $y \sim z$.
Thus $[\min\{x,y\}, \max\{x,y\}] \subseteq E$ and $[\min\{y,z\}, \max\{y,z\}] \subseteq E$.
Therefore$[\min\{x,y,z\}, \max\{x,y,z\}] \subseteq E$.
Therefore $[\min\{x,z\}, \max\{x,z\}] \subseteq E$, meaning $x \sim z$.
Thus $\sim$ is transitive.

Claim 2: The $\sim$ -equivalence classes are intervals.

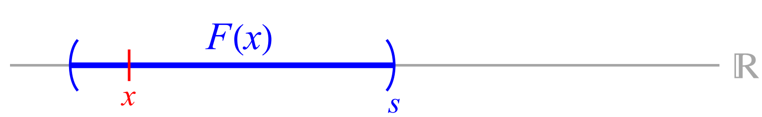
Proof of Claim 2: Let $F(x)$ denote the equivalence class of $x \in E$.

We will show $F(x)$ has betweenness.
To do this, suppose $y, z \in F(x)$ with $y \leq z$.
By definition of $\sim$, $[\min\{y,x\}, \max\{y,x\}]$ and $[\min\{z,x\}, \max\{z,x\}]$ are both subsets of E, meaning $[\min\{y,z,x\}, \max\{y,z,x\}]$ is also a subset of E, so $[y,z] \subseteq F(x)$.
Therefore $F(x)$ has betweenness, so it is an interval.

Claim 3: The $\sim$ -equivalence classes are <u>open</u> intervals.

Proof of Claim 3: Again, let $F(x)$ be the equivalence class of $x \in E$.

If $F(x)$ is bounded above, let $s = \sup F(x)$.



If $s \in F(x)$, then $s \in E$, and since E is open, $\exists\, \epsilon > 0$ s.t. $B_\epsilon(s) \subseteq E$.
But then, $s + \dfrac{\epsilon}{2} \in F(x)$, making s not an upper bound of $F(x)$, contradicting $s = \sup F(x)$. Therefore $s \notin F(x)$.

An argument similar to that in the previous four lines proves that if $F(x)$ is bounded below, $\inf F(x) \notin F(x)$.

Thus $F(x)$ is an interval that does not contain either its sup or its inf, so it is an open interval, as desired.

Claim 4: There are only countably many $\sim$ -equivalence classes.

Proof of Claim 4: By the Density Theorem, each equivalence class (being an open interval) must contain a rational number.

There are only countably many rational numbers, so there can only be countably many equivalence classes.

Finally, by Lemma 3.20 E is the disjoint union of its $\sim$ -equivalence classes, proving the $(\Rightarrow)$ direction. □

We finish this section with a theorem that sums up our work on intervals.

A word on notation: TFAE means "The following are equivalent", meaning that if any one of the statements is true, they are all true, and if any one of the statements is false, they are all false.

Theorem 3.21 *Let $E \subseteq \mathbb{R}$. TFAE:*

1. *E is an interval.*
2. *E has the betweenness property.*
3. *E is connected.*

PROOF (1) ⇔ (2) is Lemma 3.14.

(3) ⇒ (2) is Lemma 3.16.

(2) ⇒ (3): Suppose not, i.e. that $\exists\, E \subseteq \mathbb{R}$ that has betweenness but is disconnected, say by open sets U and V.

Since U and V are open, we can write them as the disjoint union of countably many open intervals:

$$U = \bigcup_{k=1}^{\infty}(a_k, b_k) \quad \text{and} \quad V = \bigcup_{k=1}^{\infty}(c_k, d_k).$$

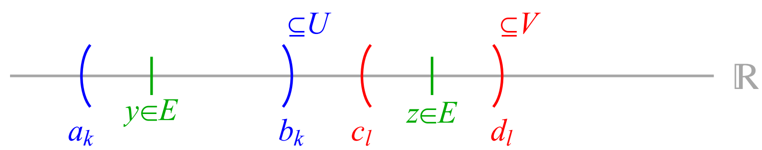
Since $\{U, V\}$ is a disconnection of E:

- U hits $E \Rightarrow \exists\, k \in \mathbb{N}$ and $y \in \mathbb{R}$ so that $y \in (a_k, b_k) \cap E$.
- V hits $E \Rightarrow \exists\, l \in \mathbb{N}$ and $z \in \mathbb{R}$ so that $z \in (c_l, d_l) \cap E$.
- U and V are disjoint, so (a_k, b_k) and (c_l, d_l) are disjoint.

Therefore, either $a_k < y < b_k < c_l < z < d_l$ or $c_l < z < d_l < a_k < y < b_k$.

WLOG (without loss of generality), the first situation holds (otherwise, switch the names of U and V).

Since E has betweenness, $b_k \in E$. So we have a picture like this:

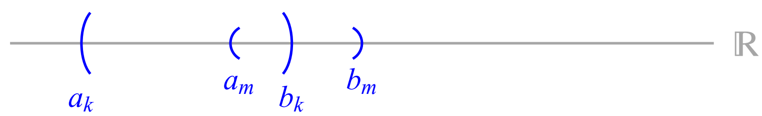


Since $E \subseteq U \cup V$, this gives two possibilities:

Case 1: $b_k \in U$.

This means $b_k \in (a_m, b_m)$ for some m.

$m \neq k$ since $b_k \notin (a_k, b_k)$.



However, in this situation, because all the intervals are open,

$$(a_m, b_m) \cap (a_k, b_k) \neq \varnothing,$$

contradicting the fact that the (a_k, b_k) are <u>disjoint</u> open intervals, so this case is impossible.

Case 2: $b_k \in V$.

This means $b_k \in (c_m, d_m)$ for some m.

However, this would imply (again since the intervals are open) that

$$(c_m, d_m) \cap (a_k, b_k) \neq \varnothing, \text{ i.e. } U \cap V \neq \varnothing,$$

which is impossible. So this case can't happen either.

In either case, we have a contradiction to the assumption that E has a disconnection.

Therefore E is connected, as wanted. □

3.3 Compactness

Motivation

Arbitrary subsets of $\mathbb{R}$ have suprema and infima, but a supremum may be ∞ or may not be in the set you're taking the supremum of. An infimum may be $-\infty$ or not in the set you're taking the infimum of.

Example $E = (0, 1)$

However, a finite set $E \subseteq \mathbb{R}$, E always has a *maximum* and a *minimum* (and that maximum and minimum are members of E).

Example $F = \{1, 6, 8, 12, 25\}$

We want to generalize this application of "finiteness" by characterizing other subsets of $\mathbb{R}$ that always contain their maximum and their minimum.

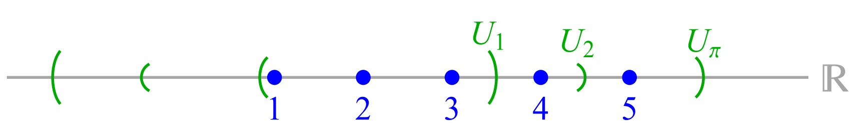
Open covers and subcovers

Definition 3.22 *Let $E \subseteq \mathbb{R}$.*
An **open cover(ing)** *of E is a set $\{U_\alpha\}_\alpha$ of open sets whose union contains (i.e. "covers" E), i.e.*

$$E \subseteq \bigcup_\alpha U_\alpha.$$

Let $\{U_\alpha\}$ be an open cover of E. A **subcover** *(of $\{U_\alpha\}$) is another open cover of E consisting of some (maybe all) of the U_α.*
E is called **compact** *if every open cover of E has a finite subcover.*

Example 1

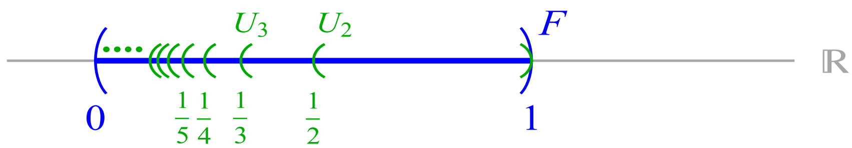


$E = \{1, 2, 3, 4, 5\}$

For each $\alpha \in \mathbb{R}$, let $U_\alpha = \left(\alpha - \frac{5}{2}, \alpha + \frac{5}{2}\right)$.

$\{U_\alpha\}$ is an open cover of E, since $E \subseteq \bigcup_\alpha U_\alpha$:

EXAMPLE 2



$F = (0, 1)$

For each $n \in \mathbb{N}$, let $U_n = \left(\frac{1}{n}, 1\right)$.

$\{U_n\}$ is an open cover of F, since $F \subseteq \bigcup_{n=1}^{\infty} U_n$:

However, there is no finite subcover of $\{U_n\}$:

EXAMPLE 3

Let $E \subseteq \mathbb{R}$ be any set. An open cover of E can be obtained by taking

Theorem 3.23 (Union of finitely many compact sets is compact) *If $E_1, E_2, \ldots, E_n$ are each compact subsets of $\mathbb{R}$, then $\bigcup_{k=1}^{n} E_k$ is also compact.*

PROOF HW

Hints: to show a set is compact, start with an arbitrary open cover $\{U_\alpha\}$ of that set. You have to show that there must be a finite subcover. Here, you start with an open cover $\{U_\alpha\}$ of $\bigcup_{k=1}^{n} E_k$. Notice that $\{U_\alpha\}$ is also an open cover of each E_k. Apply compactness of E_k to find finite subcovers of each E_k, and put them together to get a finite subcover of $\bigcup_{k=1}^{n} E_k$.

Theorem 3.24 (Intersection of compact sets is compact) *If $\{E_\alpha\}$ is a collection of compact subsets of $\mathbb{R}$, then $\bigcap_\alpha E_\alpha$ is also compact.*

PROOF HW

Countable subcovers

It turns out that every open cover of any subset of $\mathbb{R}$ has a countable subcover. We prove that in the next two results:

Lemma 3.25 *Let $E \subseteq \mathbb{R}$. E is* **separable**, *meaning there is a countable set $C \subseteq E$ such that for every $x \in E$ and every $\epsilon > 0$, there is $y \in C$ such that $|y - x| < \epsilon$.*

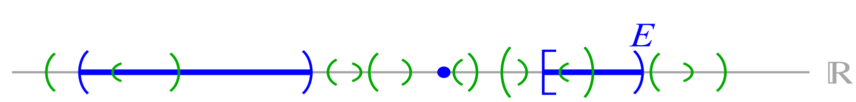
PROOF For each n, use the idea of Example 3 above to construct this countable open cover of $\mathbb{R}$ by balls with rational centers:

$$\{B_{1/n}(q) : q \in \mathbb{Q}\}.$$

(This cover is countable because there are only countably many choices of $q \in \mathbb{Q}$ and countably many choices of $n \in \mathbb{N}$.)

Now, for each set $B_{1/n}(q)$ in the cover that intersects E, select one point in that ball that is also in E.

This produces a countable set C_n of points in E.



Now let $C = \bigcup_{n=1}^{\infty} C_n$.

C, being the countable union of countable sets, is countable.

It remains to show $\forall\, x \in E$, $\forall \epsilon > 0$, $\exists\, y \in C$ s.t. $|y - x| < \epsilon$.

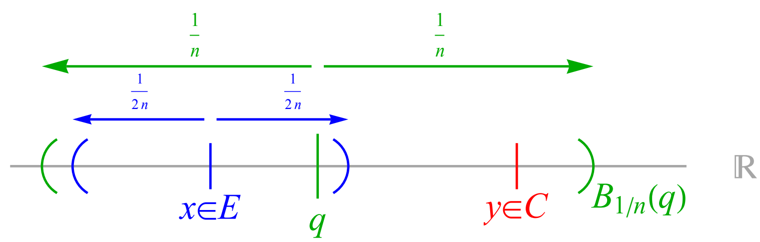
To verify this, first let $x \in E$ and $\epsilon > 0$.

Given ϵ, choose $n > \frac{3}{2\epsilon}$, so that $\frac{3}{2n} < \epsilon$.

Next, by the Density Theorem, $\exists$ rational number $q \in \left(x - \frac{1}{2n}, x + \frac{1}{2n}\right)$.

That means $|x - q| < \frac{1}{2n}$, so $|x - q| < \frac{1}{n}$, i.e. $x \in B_{1/n}(q)$.

Since $B_{1/n}(q) \cap E \neq \varnothing$, $\exists\, y \in C_n$, i.e. $y \in B_{1/n}(q) \bigcap E$.



For this y, $|y - x| \le |y - q| + |q - x| < \frac{1}{n} + \frac{1}{2n} = \frac{3}{2n} < \epsilon$.

This proves the theorem. □

Theorem 3.26 *Let $E \subseteq \mathbb{R}$. For any cover of E by open sets $\{U_\alpha\}$, there is a countable subcover.*

PROOF Let $C \subseteq E$ be as in Lemma 3.25 (meaning that $\forall\, x \in E$ and $\forall\, \epsilon > 0$, $\exists\, y \in C$ with $|x - y| < \epsilon$).

Now consider the open sets

$$\mathcal{B} = \{B_q(c) : c \in C, q \in \mathbb{Q} \bigcap (0, \infty)\}.$$

Since C and $\mathbb{Q}$ are countable, there are countably many sets in $\mathcal{B}$.

Next, let $\mathcal{B}'$ be the collection of sets in $\mathcal{B}$ which are contained entirely within a single U_α (where $\{U_\alpha\}$ is the open cover given in the theorem).

$\mathcal{B}'$ is a countable collection of open sets; label these sets as $F_1, F_2, F_3, \ldots$.

Claim: $\mathcal{B}' = \{F_1, F_2, F_3, \ldots\}$ is a cover of E.

Proof of Claim: Let $x \in E$.

Since $\{U_\alpha\}$ is a cover of E, $\exists\, \alpha$ s.t. $x \in U_\alpha$.

This U_α is open, so $\exists$ rational number $\epsilon > 0$ s.t. $B_\epsilon(x) \subseteq U_\alpha$.

Applying Corollary 3.25, there is $y \in C$ such that $|y - x| < \dfrac{\epsilon}{4}$. Notice

$$B_{\epsilon/2}(y) \subseteq B_{\epsilon/2+|y-x|}(x) \subseteq B_{\epsilon/2+\epsilon/4}(x) = B_{3\epsilon/4}(x) \subseteq B_{\epsilon}(x) \subseteq U_{\alpha},$$

so $B_{\epsilon/2}(y)$ is one of the members of $\mathcal{B}'$ (say F_N).
At the same time $x \in B_{\epsilon/2}(y)$, so $x \in F_N$.
This shows every $x \in E$ belongs to some F_N, proving the claim.

For each set F_n in $\mathcal{B}'$, we now choose a U_n from the $\{U_\alpha\}$ which contains F_n (such a U_α must exist by the definition of $\mathcal{B}'$).
This yields a countable subcollection $\{U_1, U_2, U_3, \ldots\}$ where $F_n \subseteq U_n$ for all n.
If $x \in E$, then

$$x \in \bigcup_{n=1}^{\infty} F_n \subseteq \bigcup_{n=1}^{\infty} U_n,$$

so $\{U_n\}_{n=1}^{\infty}$ is a countable subcover of the $\{U_\alpha\}$. □

So for any open cover of any subset of $\mathbb{R}$, finding a countable subcover is trivial (now that we have proved this theorem). But finding a finite subcover might be impossible, as we have seen with the set $(0,1)$ and as we now see with the set $\mathbb{R}$:

EXAMPLE

Prove that $\mathbb{R}$ is not compact.

Sequential compactness

Definition 3.27 *A set $E \subseteq \mathbb{R}$ is called* **sequentially compact** *if every sequence of numbers in E has a subsequence which converges to a number in E.*

EXAMPLES

$\mathbb{R}$

$(0, 1]$

$\mathbb{Q}$

$[0, 1]$

Theorem 3.28 *Let $E \subseteq \mathbb{R}$. E is compact if and only if it is sequentially compact.*

Proof ($\Rightarrow$) Assume E is compact and let $\{x_n\}$ be a sequence in E.

Case 1: There are only finitely many different numbers in $\{x_n\}$.

In this situation, $\exists\, x \in E$ s.t. for infinitely many n, $x_n = x$.

Then $\exists$ subsequence $\{x_{n_k}\}$ where $x_{n_k} = x$ for all k.

This subsequence converges to $x \in E$.

Case 2: $\{x_n\}$ has infinitely many different elements.

In this situation, we make the following claim:

Claim: $\exists\, x \in E$ s.t. for every $\epsilon > 0$, $\exists\, x_n$ s.t. $|x_n - x| < \epsilon$.

Proof of claim: Suppose not, i.e. $\forall\, y \in E$, $\exists\, \epsilon(y) > 0$ s.t. $B_{\epsilon(y)}(y) \cap \{x_n\} = \{y\}$.

Now, $\{B_{\epsilon(y)}(y) : y \in E\}$ is an open cover of E.

By compactness there is a finite subcover

$$\{B_{\epsilon(y_k)}(y_k) : 1 \leq k \leq n\}.$$

But then,

$$\begin{aligned}
\{x_n\}_{n=1}^{\infty} &= \{x_n\}_{n=1}^{\infty} \cap E \quad \text{(since } \{x_n\} \subseteq E\text{)} \\
&\subseteq \left[\bigcup_{k=1}^{n} B_{\epsilon(y_k)}(y_k)\right] \cap \{x_n\}_{n=1}^{\infty} \quad \text{(since } \{B_{\epsilon(y_k)}(y_k)\} \text{ covers } E\text{)} \\
&= \bigcup_{k=1}^{n} \left[B_{\epsilon(y_k)}(y_k) \cap \{x_n\}_{n=1}^{\infty}\right] \\
&= \bigcup_{k=1}^{n} \{y\} \quad \text{(by the red remark above)} \\
&= \{y\},
\end{aligned}$$

contradicting the fact that $\{x_n\}$ has infinitely many different elements.

Applying the claim with $\epsilon = 1$, $\exists\, n_1 \in \mathbb{N}$ s.t. $x_{n_1} \neq x$ and $|x_{n_1} - x| < 1$.

Applying the claim again with $\epsilon = \min\left\{\frac{1}{2}, |x_{n_1} - y|\right\}$ to show

$$\exists\, n_2 > n_1 \text{ s.t. } x_{n_2} \neq x \text{ and } |x_{n_2} - x| < \frac{1}{2}.$$

For each k, apply the claim again with $\epsilon = \min\left\{\frac{1}{k+1}, |x_{n_k} - y|\right\}$ to show

$$\exists\, n_{k+1} > n_k \text{ s.t. } x_{n_{k+1}} \neq x \text{ and } |x_{n_{k+1}} - x| < \frac{1}{k+1}.$$

By the Squeeze Theorem, $x_{n_k} \to x \in E$.

Therefore E is sequentially compact.

($\Leftarrow$) Assume E is sequentially compact.

Start with an arbitrary open cover $\{U_\alpha\}$ of E; by Theorem 3.26, $\exists$ countable subcover $\{U_1, U_2, U_3, \ldots\}$.

Suppose not, i.e. that there is no finite subcover of $\{U_1, U_2, U_3, \ldots\}$.

Then, $\forall\, n \geq 0$, $\{U_1, \ldots, U_n\}$ does not cover E, so $\exists\, x_n \in E - (U_1 \cup U_2 \cup \cdots \cup U_n)$.

E is assumed sequentially compact, so $\{x_n\}$ has a subsequence $\{x_{n_k}\}$ which converges to some $x \in E$.

This x must belong to U_N for some N, since $\{U_1, U_2, \ldots\}$ cover E.

As U_N is open, there is $\epsilon_0 > 0$ such that $B_{\epsilon_0}(x) \subseteq U_N$.

On the other hand, for every $n \geq N$, $x_n \notin U_N$, so $x_n \notin B_{\epsilon_0}(x)$, i.e. $|x_n - x| \geq \epsilon_0$.

This contradicts $x_{n_k} \to x$, so in fact $\{U_1, U_2, \ldots\}$ must have a finite subcover.

This means E is compact as wanted. □

Theorem 3.29 (Heine-Borel Theorem) *Let $E \subseteq \mathbb{R}$. E is compact if and only if E is closed and bounded.*

PROOF ($\Rightarrow$) Assume E is compact; that means E is sequentially compact.

To show E is closed, suppose $\{x_n\} \subseteq E$ is some sequence with $x_n \to x \in \mathbb{R}$.

By sequential compactness, $\exists$ subsequence $\{x_{n_k}\}$ which converges to a number in E.

But this number must be x (since any subsequence has the same limit as the original convergent sequence), so $x \in E$.

Thus E is sequentially closed, hence closed.

To show E is bounded, let $\epsilon > 0$.

Observe that $\{B_\epsilon(x) : x \in E\}$ is an open cover of E.

By compactness, there is a finite subcover, i.e. $E \subseteq \bigcup_{k=1}^{n} B_\epsilon(x_k)$.

Last, let $b = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$.

For every $x \in E$, $|x| \leq b + \epsilon$, so E is bounded.

($\Leftarrow$) Let E be closed and bounded.

Then, let $\{x_n\}$ be a sequence in E.

Since E is bounded, by the ______________________________Theorem, $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$.

Since E is closed, it is sequentially closed, so $x = \lim x_{n_k}$ belongs to E.

We have proven E is sequentially compact, and therefore E is compact. □

Corollary 3.30 *Let* $E \subseteq \mathbb{R}$.
If E *is compact, then* $\sup E \in E$ *and* $\inf E \in E$.
In other words, $\sup E = \max E$ *and* $\inf E = \min E$, *so* E *contains its maximum and minimum.*

PROOF Let $s = \sup E$; for every n, we can choose $x_n \in \left(s - \frac{1}{n}, s\right] \bigcap E$.

Thus $|x_n - s| \leq \frac{1}{n}$ so $x_n \to s$.

Since E is compact, it is closed, hence sequentially closed, so $s \in E$.

The proof that $\inf E \in E$ is similar. □

The Nested Interval Theorem

From work in the last two sections, a subset of $\mathbb{R}$ is connected and compact if and only if it is a closed and bounded interval $[a, b]$.
Now for a result which says something about what happens when you intersect certain types of these sets:

Theorem 3.31 (Nested Interval Theorem) *Let* $\{I_n\}$ *be a sequence of nonempty compact intervals in* $\mathbb{R}$. *If* $I_{n+1} \subseteq I_n$ *for every* n, *then*

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

VOCABULARY

A sequence of sets is called **nested** if the sets either get bigger and bigger, or smaller and smaller. For this theorem, the sets get smaller and smaller:

Note: All the hypotheses of this theorem are important for the conclusion to be true (HW).

We're going to prove the Nested Interval Theorem two different ways: one way that uses completeness, and one way that uses compactness.

FIRST PROOF As each I_n is a closed bounded interval, we can write $I_n = [a_n, b_n]$ with $a_n \le b_n$.

The fact that $I_{n+1} \subseteq I_n$ means $\{a_n\}$ is an increasing sequence bounded above by b_1 and $\{b_n\}$ is a decreasing sequence bounded below by a_1.

By the MCT, $a = \lim a_n = \sup a_n$ and $b = \lim b_n = \inf b_n$ exist, and $a \le b$ since limits preserve soft inequalities.

So $[a, b] \neq \varnothing$.

Claim: $[a, b] = \bigcap_{n=1}^{\infty} I_n$.

Proof of claim: ($\subseteq$) Let $x \in [a, b]$.

Then $x \ge a = \sup a_n$ so $x \ge a_n\ \forall\, n$, and $x \le b = \inf b_n$, so $x \le b_n\ \forall\, n$.

Thus $x \in [a_n, b_n] = I_n$ for all n.

Therefore $x \in \bigcap_{n=1}^{\infty} I_n$ as wanted.

($\supseteq$) Let $x \in \bigcap_{n=1}^{\infty} I_n = \bigcap_{n=1}^{\infty} [a_n, b_n]$.

Thus $x \ge a_n$ for all n, so $x \ge \sup a_n = a$.

Similarly, $x \le b_n$ for all n, so $x \le \inf b_n = b$.

Thus $x \in [a, b]$ as wanted. □

SECOND PROOF Suppose not, that $\bigcap_{n=1}^{\infty} I_n = \varnothing$.

Then since each I_n is compact, each I_n is closed.

So each I_n^C is open, and by DeMorgan's Law,

$$\bigcup_{n=1}^{\infty} I_n^C = \left(\bigcap_{n=1}^{\infty} I_n \right)^C = \varnothing^C = \mathbb{R}.$$

In other words, $\{I_n^C\}$ is an open covering of $\mathbb{R}$, hence an open covering of the compact set I_1.

By compactness, $\exists$ finite subcover, i.e.$\exists\, N$ s.t. $\bigcup_{n=1}^{N} I_n^C \supseteq I_1$. Thus

$$\left(\bigcap_{n=1}^{N} I_n \right)^C = \bigcup_{n=1}^{N} I_n^C \supseteq I_1$$

so

$$I_N = \bigcap_{n=1}^{N} I_n \subseteq I_1^C.$$

But this contradicts $I_N \subseteq I_{N-1} \subseteq \cdots \subseteq I_2 \subseteq I_1$, since $I_N \neq \varnothing$.

The result follows by contradiction. □

Corollary 3.32 *Let $I_n = [a_n, b_n]$ be a sequence of nonempty, closed, bounded intervals in $\mathbb{R}$. If $I_{n+1} \subseteq I_n$ for every n and $(b_n - a_n) \to 0$, then there is a unique real number x such that*

$$\bigcap_{n=1}^{\infty} I_n = \{x\}.$$

PROOF Let $x = \lim b_n$ (which exists by the MCT).

Since $\lim(b_n - a_n) = 0$, $\lim b_n = \lim a_n = x$, so by the first proof of the Nested Interval Theorem,

$$\bigcap_{n=1}^{\infty} I_n = [\sup a_n, \inf b_n] = [\lim a_n, \lim b_n] = [x, x] = \{x\}. \square$$

Corollary 3.33 $\mathbb{R}$ *is uncountable.*

PROOF First, we will prove $[0, 1]$ is uncountable.

Suppose not, i.e. $\exists$ an injection $f : [0, 1] \to \mathbb{N}$. Write $x_n = f^{-1}(n)$ so that

$$[0, 1] = \{x_1, x_2, x_3, x_4, \dots\}.$$

Now, let $I_0 = [0, 1]$ and for each $n \geq 1$, choose a closed, nonempty interval $I_n \subseteq I_{n-1}$, which has $\frac{1}{3}$ the length of I_{n-1}, such that $x_n \notin I_n$.

By the Nested Interval Theorem, $\bigcap_{n=1}^{\infty} I_n \neq \varnothing$, so $\exists\, x \in \bigcap_{n=1}^{\infty} I_n \subseteq [0, 1]$.

But this x cannot equal x_n for any n: $x \in I_n$, but $x_n \notin I_n$.

This contradicts the assumption that $[0, 1] = \{x_1, x_2, x_3, x_4, \dots\}$.

Therefore $[0, 1]$ must be countable.

If $\mathbb{R}$ was countable, then any subset of $\mathbb{R}$ (such as $[0, 1]$) would be countable.

Thus $\mathbb{R}$ is also uncountable. $\square$

3.4 Chapter 3 Summary

DEFINITIONS TO KNOW

Acronyms

- **WLOG** means "without loss of generality"
- **TFAE** means "the following are equivalent"

Nouns

- An **open ball** $B_\epsilon(x)$ is a set $\{y \in \mathbb{R} : |y - x| < \epsilon\}$; equivalently, an open ball in $\mathbb{R}$ is a bounded open interval (a, b).
- A set $E \subseteq \mathbb{R}$ has **betweenness** if $\forall\, a, b \in E$ with $a \leq b$, $[a, b] \subseteq E$.
- A **disconnection** of set $E \subseteq \mathbb{R}$ is a pair of disjoint open sets $\{U, V\}$ which both hit E and whose union covers E.
- A relation $\sim$ on set E is called an **equivalence relation** if it is reflexive ($x \sim x$), symmetric ($x \sim y$ implies $y \sim x$) and transitive ($x \sim y$ and $y \sim z$ implies $x \sim z$).

  If $\sim$ is an equivalence relation on E, the **equivalence class** of $x \in E$ is the set of things in E equivalent to x.
- An **open cover** of set E is a collection of open sets whose union contains E.

  A **subcover** of open cover $\{U_\alpha\}$ is another open cover consisting of some (maybe all) of the U_α.

Adjectives that describe subsets of $\mathbb{R}$

- E is **open** if $\forall\, x \in E$, $\exists\, \epsilon > 0$ s.t. $B_\epsilon(x) \subseteq E$.
- E is **closed** if its complement is open.
- E is **clopen** if it is closed and open.
- E is **sequentially closed** if for every sequence in E that converges to a limit in $\mathbb{R}$, the limit must be in E.
- E is **connected** if it does not have a disconnection.
- E is called **compact** if every open cover of E has a finite subcover.
- E is called **sequentially compact** if every sequence of numbers in E has a subsequence which converges to a limit in E.

THEOREMS WITH NAMES

Lindelöf's Theorem $E \subseteq \mathbb{R}$ is open $\Leftrightarrow$ E is the union of countably many disjoint open intervals.

Heine-Borel Theorem $E \subseteq \mathbb{R}$ is compact $\Leftrightarrow$ E is closed and bounded.

($\Leftrightarrow$ E is sequentially compact, although this part isn't "Heine-Borel")

Nested Interval Theorem If $\{I_n\}$ is a sequence of nonempty, closed bounded intervals with $I_{n+1} \subseteq I_n$ for all n, then $\bigcap_n I_n \neq \varnothing$.

In this setting, if $I_n = [a_n, b_n]$ and $b_n - a_n \to 0$, then $\exists\, x \in \mathbb{R}$ s.t. $\bigcap_n I_n = \{x\}$.

OTHER THEOREMS TO REMEMBER

- Open balls are open sets; the union of any number of open sets is open; the intersection of finitely many open sets is open.
- The intersection of any number of closed sets is closed; the union of finitely many closed sets is closed.
- $E \subseteq \mathbb{R}$ is closed $\Leftrightarrow$ E is sequentially closed.
- Open sets do not contain their infimum or supremum.
- Closed sets (therefore also compact sets) contain their infimum and their supremum.
- For $E \subseteq \mathbb{R}$, E is an interval $\Leftrightarrow$ E has betweenness $\Leftrightarrow$ E is connected.
- The intersection of any number of compact sets is compact; the union of finitely many compact sets is compact.
- Every open cover of any subset of $\mathbb{R}$ has a countable subcover.
- $\mathbb{R}$ is uncountable.

STANDARD PROOF TECHNIQUES

To prove that $E \subseteq \mathbb{R}$ is open, do <u>one</u> of these things:

1. Show E is the union of sets already known to be open (like open intervals).
2. Show E is the intersection of finitely many sets already known to be open.
3. Use the definition: let $x \in E$ and write down a formula for $\epsilon > 0$ coming from scratch work; then prove $B_\epsilon(x) \subseteq E$.

To prove that $E \subseteq \mathbb{R}$ is closed, do one of these things:

1. Show E is the intersection of sets already known to be closed (like singletons or closed intervals).
2. Show E is the union of finitely many sets already known to be closed.
3. Show E is sequentially closed: take $\{x_n\} \subseteq E$ with $x_n \to x$, and prove $x \in E$.
4. Show E^C is open (see above).

To prove that $E \subseteq \mathbb{R}$ is connected, do one of these things:

1. Show E is an interval.
2. Show E has betweenness: let $x, y \in E$ with $x \leq y$ and prove $[x, y] \subseteq E$.
3. Show E has no disconnection (usually by assuming not and deriving a contradiction).

To prove that $E \subseteq \mathbb{R}$ is compact, do one of these things:

1. Show E is closed and bounded.
2. Show E is the intersection of sets already known to be compact.
3. Show E is the union of finitely many sets already known to be compact.
4. Show E is sequentially compact: take $\{x_n\} \subseteq E$ and prove there is a subsequence $\{x_{n_k}\}$ s.t. $x_{n_k} \to x \in E$.
5. Use the definition: let $\{U_\alpha\}$ be an open cover of E and prove that it has a finite subcover.

3.5 Chapter 3 Homework

Exercises from Section 3.1

1. Prove that for any $x \in \mathbb{R}$, if $0 < \delta < \epsilon$, then $B_\delta(x) \subseteq B_\epsilon(x)$.

2. Prove that for any $x, y \in \mathbb{R}$ and any $\epsilon > 0$, $B_\epsilon(x) \subseteq B_{\epsilon+|x-y|}(y)$.

3. Prove Theorem 3.4, which says that every open ball in $\mathbb{R}$ is an open set.

4. Prove that for any $x \in \mathbb{R}$, the sets $(-\infty, x)$ and (x, ∞) are open.

5. Give a specific example, with proof, of countably many closed sets whose union is not closed.

6. Let $a \leq b$. Prove that $[a, b]$ is closed.

   Note: Having done this problem, you will have proven that any singleton (i.e. set with one element) $\{a\}$ is closed, since $\{a\} = [a, a]$.

   Also, you will have proven that any finite set is closed, since any finite set is the union of finitely many singletons (i.e. $\{x_1, ..., x_n\} = \bigcup_{k=1}^{n} \{x_k\}$).

7. Let $E \subseteq \mathbb{R}$. Prove that E is open if and only if $-E$ is open (recall that $-E = \{-x : x \in E\}$.

8. Let $E \subseteq \mathbb{R}$. Prove that E is closed if and only if $-E$ is closed.

9. Prove the first statement of Theorem 3.12, which says that if $E \subseteq \mathbb{R}$ is open, then $\sup E \notin E$ and $\inf E \notin E$.

10. Prove the second statement of Theorem 3.12, which says that if $E \subseteq \mathbb{R}$ is a closed set that is bounded above, then $\sup E \in E$.

11. Prove the third statement of Theorem 3.12, which says that if $E \subseteq \mathbb{R}$ is a closed set that is bounded below, then $\inf E \in E$.

12. Without proof, characterize each of these sets as "open", "closed", "clopen", or "neither open nor closed":

    a) $\{0\}$
    b) $E = \{0, 1\}$
    c) $[0, 1)$
    d) $F = \left\{\frac{1}{n} : n \in \mathbb{N}\right\}$
    e) $G = \left\{\frac{1}{n} : n \in \mathbb{N}\right\} \cup \{0\}$
    f) $[0, \infty)$
    g) $\mathbb{R} - \{0\}$
    h) $\mathbb{Z}$
    i) $\mathbb{R} - \mathbb{Q}$
    j) $\varnothing$

Exercises from Section 3.2

13. For each set described in Exercise 12, determine whether or not the set is connected. If the set is disconnected, write down an explicit disconnection of the set; if the set is connected, you do not need to prove that it is connected.

14. Prove or disprove: if E and F are connected subsets of $\mathbb{R}$, then $E \cup F$ is connected.

15. Prove or disprove: if E and F are connected subsets of $\mathbb{R}$ and $E \cap F \neq \varnothing$, then $E \cup F$ is connected.

    Hint: Use the fact that connected subsets have the betweenness property.

16. Prove or disprove: if E and F are connected subsets of $\mathbb{R}$, then $E \cap F$ is connected.

17. Consider the set $E = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$. Show E is disconnected by finding an explicit disconnection of E (you need to prove that you have found a disconnection).

Exercises from Section 3.3

18. Use the definition of compactness (not sequential compactness or Heine-Borel) to show that $[1, \infty)$ is not compact.

19. Prove Theorem 3.23, which says that if $E_1, E_2, ..., E_n$ are compact subsets of $\mathbb{R}$, then $\bigcup_{k=1}^{n} E_k$ is compact.

20. Prove Theorem 3.24, which says that if $\{E_\alpha\}$ is a collection of compact subsets of $\mathbb{R}$, then $\bigcap_{\alpha} E_k$ is compact.

21. Prove that if $E \subseteq \mathbb{R}$ is compact and $F \subseteq E$ is closed, then F is compact.

22. For each set described in Exercise 12, determine whether or not the set is compact (no proof is required).

23. Determine, with proof, whether each set is compact:

    a) $E = \{\frac{1}{n} : n \in \mathbb{N}\}$

    b) $F = \{\frac{1}{n} : n \in \mathbb{N}\} \cup \{0\}$

24. This problem verifies that all the hypotheses of the Nested Interval Theorem are needed to draw its conclusion.

a) Give an example of a sequence $\{I_n\}$ of nonempty, closed, and bounded intervals with $\bigcap_{n=1}^{\infty} I_n = \varnothing$.

b) Give an example of a sequence $\{I_n\}$ of nonempty closed intervals with $I_{n+1} \subseteq I_n$ for every n where $\bigcap_{n=1}^{\infty} I_n = \varnothing$.

c) Give an example of a sequence $\{I_n\}$ of nonempty bounded intervals with $I_{n+1} \subseteq I_n$ for every n where $\bigcap_{n=1}^{\infty} I_n = \varnothing$.

Chapter 4

Infinite series

4.1 Convergence of infinite series

In this chapter, we discuss *infinite series*, which you first encountered in Calculus 2. Recall that an infinite series is

To accomplish this, we associate to every infinite series a sequence of numbers; summing the infinite series corresponds to taking the limit of that sequence:

Definition 4.1 *Let $\{a_n\}$ be a sequence of real numbers.*
The **sequence $\{S_N\}$ of partial sums** *associated to $\{a_n\}$ is defined as follows:*

$$\begin{aligned} S_1 &= a_1 \\ S_2 &= a_1 + a_2 \\ S_3 &= a_1 + a_2 + a_3 \\ &\vdots \\ S_N &= \sum_{n \le N} a_n = \sum_{n=1}^{N} a_n \end{aligned}$$

If the sequence $\{S_N\}$ converges to $S \in \mathbb{R}$, then we say $\sum a_n$ **converges (to S)** *and we write $\sum a_n = S$ or $\sum_{n=1}^{\infty} a_n = S$.*
If $\{S_N\}$ diverges, we say $\sum a_n$ **diverges**.

Let's discuss some basic series you first study in Calculus 2:

Theorem 4.2 (Geometric series formula) *Let $r \in (-1, 1)$. Then*

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}.$$

PROOF To show this, we need to use the definition of convergence.

Let S_N be the N^{th} partial sum of this series. Then

$$S_N = r^0 + r^1 + r^2 + \cdots + r^{n-1} + r^n.$$

Therefore $(1-r)S_N = 1 - r^{N+1}$, so $S_N = \dfrac{1 - r^{N+1}}{1-r}$.

Now, as $N \to \infty$, $S_N = \dfrac{1 - r^{N+1}}{1-r} \to$

There are a couple of related formulas that we will need:

Corollary 4.3 (Finite geometric sum formulas) *Let $r \in \mathbb{R}$, and let $M, N \in \mathbb{N}$ be such that $M \leq N$. Then*

$$\sum_{n=0}^{N} r^n = \frac{1 - r^{N+1}}{1-r} \quad \textit{and} \quad \sum_{n=M}^{N} r^n = r^M \left(\frac{1 - r^{N-M+1}}{1-r} \right).$$

PROOF We proved the first formula when proving the preceding theorem.

For the second formula,

$$\begin{aligned}
\sum_{n=M}^{N} r^n &= r^M + r^{M+1} + \cdots r^N \\
&= r^M \left(1 + r + r^2 + \cdots + r^{N-M}\right) \\
&= r^M \sum_{n=0}^{N-M} r^n \\
&= r^M \left(\frac{1 - r^{N-M+1}}{1-r} \right). \; \square
\end{aligned}$$

Theorem 4.4 *The harmonic series* $\sum_{n=1}^{\infty} \frac{1}{n}$ *diverges.*

PROOF If the series converges, then its sequence $\{S_N\}$ of partial sums converges, so any subsequence of $\{S_N\}$ also converges.

But, consider the subsequence $\{S_{2^N}\}$:

$$S_{2^N} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{7} + \frac{1}{8} + \cdots + \frac{1}{15} + \frac{1}{16} + \cdots + \frac{1}{2^5 - 1} + \frac{1}{2^5} + \cdots + \frac{1}{2^N}$$

$$>$$

$$= \frac{1}{2} + \frac{1}{2}(1) + \frac{1}{4}(2) + \frac{1}{8}(4) \cdots \frac{1}{2^N}(2^{N-1})$$

$$= \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots + \frac{1}{2}$$

$$= \frac{N}{2}.$$

Since $S_{2^N} > \frac{N}{2}$, $\{S_{2^N}\}$ is unbounded, hence cannot converge.

Thus neither does $\{S_N\}$, so $\sum \frac{1}{n}$ must diverge, as wanted. □

Theorem 4.5 *The p-series* $\sum_{n=1}^{\infty} \frac{1}{n^2}$ *converges.*

PROOF Let S_N be the N^{th} partial sum of this series.

$\{S_N\}$ is an increasing sequence, since each term of the series is positive.

So it is sufficient to show that the sequence $\{S_N\}$ is bounded above.

Notice first that

$$\begin{aligned}
S_{2^k-1} &= 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \frac{1}{25} + \cdots + \frac{1}{(2^k-1)^2} \\
&= 1 + \left[\frac{1}{4} + \frac{1}{9}\right] + \left[\frac{1}{16} + \cdots + \frac{1}{(2^3-1)^2}\right] + \left[\frac{1}{(2^3)^2} + \cdots + \frac{1}{(2^4-1)^2}\right] + \cdots + \frac{1}{(2^k-1)^2} \\
&< \\
&= 1 + \frac{1}{4}(2) + \frac{1}{16}(4) + \frac{1}{2^6}(2^3) + \frac{1}{2^8}(2^4) + \ldots + \frac{1}{2^{2k-2}}(2^{k-1}) \\
&= 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{2^3} + \frac{1}{2^4} + \cdots + \frac{1}{2^{k-1}} \\
&= \sum_{n=0}^{k-1} \left(\frac{1}{2}\right)^n \\
&< \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n \\
&= \frac{1}{1-\frac{1}{2}} = 2.
\end{aligned}$$

That means $\{S_{2^k-1}\}$ is bounded above by 2, so $\{S_N\}$ is also bounded above by 2 since $\{S_N\}$ is increasing.

By the MCT, $\{S_N\}$ converges, so $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges by definition. □

Corollary 4.6 *For any $p \geq 2$, the p-series*

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

converges.

PROOF For each N, the N^{th} partial sums of this series form an increasing sequence of numbers, the N^{th} of which is less than the N^{th} partial sum of $\sum \frac{1}{n^2}$, and subsequently less than $\sum_{n=1}^{\infty} \frac{1}{n^2}$. By the MCT, $\sum \frac{1}{n^p}$ converges. □

4.2 Decimal and base b representations

In elementary school, you learn about decimals. For example, if you write

$$x = .45028 = .45028000000\cdots \qquad \text{or} \qquad y = .131313131313\cdots,$$

you have intuition as to what that means. Now let's give that intuition some formal grounding. For the numbers x and y given above, what we really mean when we write those decimal representations is

$$x = .45028 =$$

$$y = .131313131313... =$$

Put another way, decimals are shorthand for a particular kind of infinite series:

Theorem 4.7 (Every decimal representation gives a real number) *Let $\{x_n\}$ be a sequence of numbers, each taken from the set $\{0, 1, 2, ..., 9\}$. Then the series*

$$\sum_{n=1}^{\infty} \frac{x_n}{10^n}$$

converges to a real number x. We write this as

$$x = .x_1x_2x_3\cdots \qquad \textit{or} \qquad x = .x_1x_2x_3\cdots_{[10]}$$

PROOF Let S_N be the N^{th} partial sum of the series.

Notice that $\{S_N\}$ is an increasing sequence, and

$$\begin{aligned} S_N &= \sum_{n=1}^{N} \frac{x_n}{10^n} \\ &\le \sum_{n=1}^{N} \frac{9}{10^n} \\ &= 9 \sum_{n=1}^{N} \left(\frac{1}{10}\right)^n \\ &= \frac{9}{10} \sum_{n=0}^{N-1} \left(\frac{1}{10}\right)^n \\ &= \frac{9}{10} \cdot \frac{1 - \left(\frac{1}{10}\right)^N}{1 - \frac{1}{10}} \\ &= \left[1 - \left(\frac{1}{10}\right)^N\right] \\ &\le 1. \end{aligned}$$

By the MCT, $\{S_N\}$ converges to a real number x (and $0 \le x \le 1$). □

Definition 4.8 *Let $x_0 \in \{0, 1, 2, ...\}$ and let $x_n \in \{0, 1, 2, ..., 9\}$ for all n. By*

$$x_0.x_1x_2x_3x_4\cdots,$$

we mean the nonnegative real number $x_0 + \sum_{n=1}^{\infty} \frac{x_n}{10^n}$.

EXAMPLE

4.72727272727... =

Definition 4.9 *The* **floor function** *is the function* $\lfloor\ \rfloor : \mathbb{R} \to \mathbb{R}$ *defined by setting*

$$\lfloor x \rfloor = \sup\{y : y \le x \text{ and } y \in \mathbb{Z}\}.$$

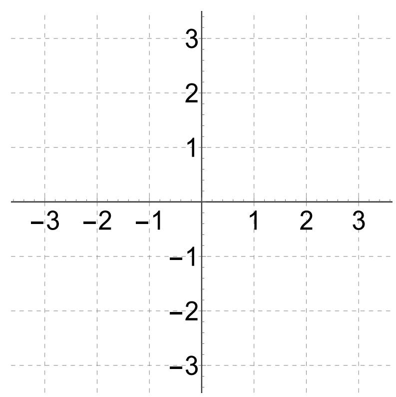
EXAMPLES

$$\lfloor 5 \rfloor = \qquad\qquad \left\lfloor \frac{17}{6} \right\rfloor =$$

$$\lfloor \pi \rfloor = \qquad\qquad \lfloor -\pi \rfloor =$$

$$\lfloor \sqrt{2} \rfloor = \qquad\qquad \lfloor -2 \rfloor =$$

The graph of the floor function looks like this:



Theorem 4.10 (Every real number has a decimal representation) *Let* $x \in \mathbb{R}$. *Then* $\exists$ *a sequence of numbers* $\{x_n\}$, *each taken from the set* $\{0, 1, 2, \ldots, 9\}$, *s.t.*

$$\begin{aligned} x &= \lfloor x \rfloor + \sum_{n=1}^{\infty} \frac{x_n}{10^n} \\ &= \lfloor x \rfloor . x_1 x_2 x_3 x_4 \cdots \\ &= \lfloor x \rfloor . x_1 x_2 x_3 x_4 \cdots_{[10]} \end{aligned}$$

This is called a **decimal** *or* **base** 10 *representation of* x.

PROOF Suppose for now that $x \in [0, 1]$ (other x's will be handled later).

To obtain a decimal representation of x, we use the Nested Interval Theorem.

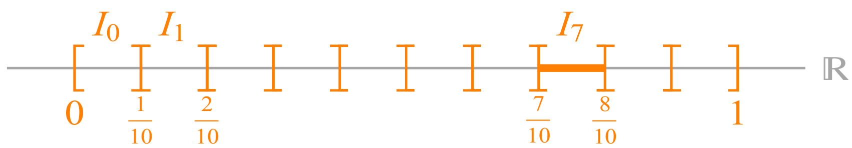
Let $D = \{0, 1, .., 9\}$; the elements of D are called **(decimal) digits**.

First step: For each $n_1 \in D$, let $I_{n_1} = \left[\frac{n_1}{10}, \frac{n_1 + 1}{10}\right]$.

$\bigcup_{n_1=0}^{9} I_{n_1} = [0, 1]$, so x must belong to some I_{n_1}.

Set $x_1 = n_1$, and let a_1 and b_1 be the the left- and right-hand endpoints of I_{n_1}:

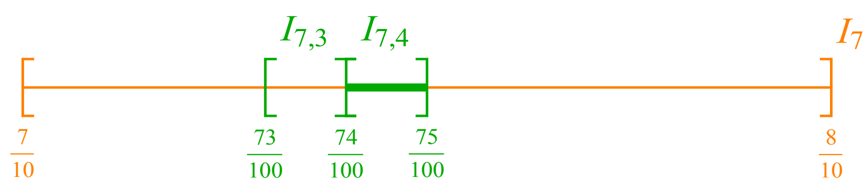
$$a_1 = \frac{n_1}{10} \quad \text{and} \quad b_1 = \frac{n_1 + 1}{10}.$$



Second step: Now, for each $n_2 \in D$, set $I_{n_1,n_2} = \left[\frac{10n_1 + n_2}{100}, \frac{10n_1 + n_2 + 1}{100}\right]$.

Since $\bigcup_{n_2=0}^{9} I_{n_1,n_2} = I_{n_1}$, x must belong to some I_{n_1,n_2}.

Set $x_2 = n_2$, and let a_2 and b_2 be the left- and right-hand endpoints of I_{n_1,n_2}.



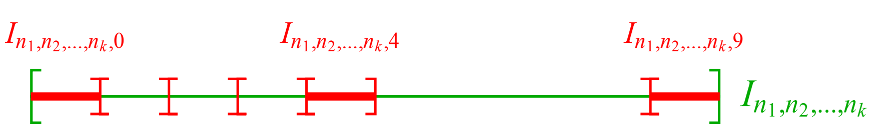
$(k+1)^{th}$ *step:* If $x \in I_{n_1,n_2,n_3,\dots,n_k}$, then for each $n_{k+1} \in D$, set

$$I_{n_1,n_2,\dots,n_k,n_{k+1}} = \left[\frac{10^k n_1 + 10^{k-1} n_2 + \cdots + 10n_k + n_{k+1}}{10^{k+1}}, \frac{10^k n_1 + \cdots + 10n_k + n_{k+1} + 1}{10^{k+1}}\right].$$

Since $\bigcup_{n_{k+1}=0}^{9} I_{n_1,n_2,\dots,n_{k+1}} = I_{n_1,n_2,\dots,n_k}$, x must belong to some $I_{n_1,n_2,\dots,n_{k+1}}$.

Set $x_{k+1} = n_{k+1}$.

Let a_{k+1} and b_{k+1} be the left- and right-hand endpoints of $I_{n_1,n_2,\dots,n_{k+1}}$.



Repeating these steps, we get a nested decreasing sequence of intervals

$$I_{x_1} \supseteq I_{x_1,x_2} \supseteq I_{x_1,x_2,x_3} \supseteq \cdots \ni x;$$

this sequence is also denoted

$$[a_1, b_1] \supseteq [a_2, b_2] \supseteq [a_3, b_3] \supseteq \cdots \ni x.$$

Now, observe for each k that $b_k = a_k + \dfrac{1}{10^k}$.

Therefore $b_k - a_k 10^{-k} \to 0$, so by the (corollary of the) Nested Interval Theorem,

$$\bigcap_{k=1}^{\infty} I_{x_1,x_2,\ldots,x_k,\ldots} = \{x\}.$$

Now consider the infinite series $\sum_{k=1}^{\infty} \dfrac{x_k}{10^k}$ (call this series (*)).

Since $a_k = \sum_{j=1}^{k} \dfrac{x_j}{10^j}$, the sequence of partial sums of (*) is $\{a_1, a_2, a_3, \ldots\}$. This

sequence converges to the unique point in $\bigcap_{k=1}^{\infty} I_{x_1,..,x_k}$, which is x.

Thus $x = \sum_{k=1}^{\infty} \dfrac{x_k}{10^k} = .x_1 x_2 x_3 x_4 \cdots$ as wanted.

This wraps up the situation when $x \in [0, 1]$.

Next, for an arbitrary $x \in [0, \infty)$, note $x = \lfloor x \rfloor + (x - \lfloor x \rfloor)$.

By the previous work, since $x - \lfloor x \rfloor \in [0, 1)$, the number $x - \lfloor x \rfloor$ has a decimal representation, which begins with the decimal point.

Stick $\lfloor x \rfloor$ in front of the decimal to get the decimal representation of x.

Last, for $x \in (-\infty, 0]$, note $-x \in [0, \infty)$.

By previous work, $-x$ has a decimal representation $x_0.x_1 x_2 \cdots$.

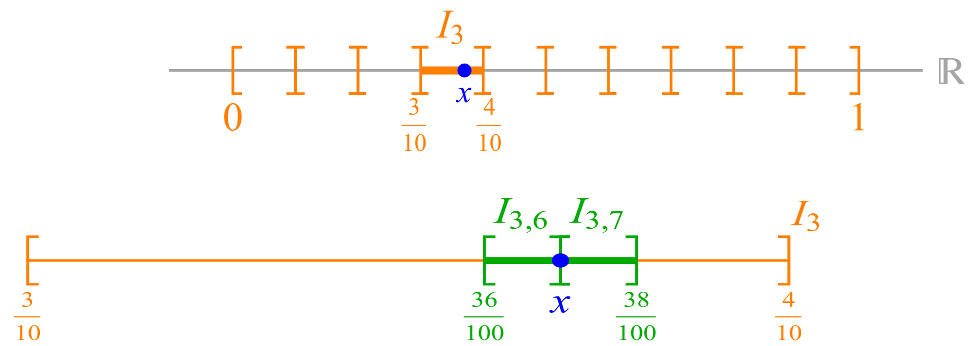
A decimal representation of $-x$ is therefore $-x_0.x_1 x_2 \cdots$. □

QUESTION

We've proven every $x \in \mathbb{R}$ has a decimal representation. Could some x have multiple different decimal representations?

EXAMPLE

Consider $x = \frac{37}{100}$.



Theorem 4.11 (Uniqueness of decimal representations) *Let $x \in [0,1]$.*
If $x = \frac{a}{10^N}$ for some $a \in \{0,1,2,...,10^N\}$, then x has exactly two decimal representations, which must be

$$x = .x_1x_2x_3\cdots x_{N-1}x_N999999999\cdots_{[10]}$$

and

$$x = .x_1x_2x_3\cdots x_{N-1}(x_N+1)000000000\cdots_{[10]}.$$

for some sequence $\{x_1, x_2, ..., x_N\}$ of numbers, each belonging to $\{0,1,2,...,9\}$, with $x_N \neq 9$.
Otherwise, x has exactly one decimal representation.

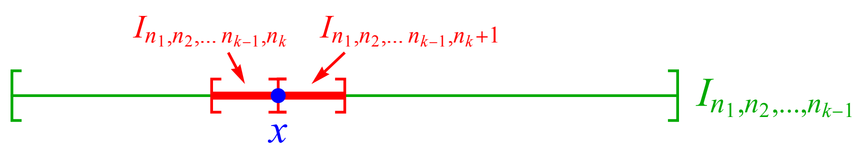
PROOF In order for x to have more than one decimal representation, x has to be an endpoint of one of the intervals $I_{n_1,\ldots,n_k}$ described in the proof of the previous theorem.

But these endpoints are precisely those which are rational numbers with denominator equal to a power of 10, meaning x has the desired form.

If x belongs to two intervals $I_{n_1,\ldots,n_k}$, choose the smallest k for which this is the case.

Then, x is the right endpoint of $I_{n_1,n_2,\ldots,n_k}$ and the left endpoint of $I_{n_1,n_2,\ldots,n_k+1}$.

At every step after the k^{th} one, x will be the right-most endpoint of $I_{n_1,\ldots,n_k,9,9,9,\ldots,9}$ and the left-most endpoint of $I_{n_1,\ldots,n_k+1,0,0,0,0,\ldots,0}$, producing the two decimal representations described in the theorem.



Last, to show that the indicated decimal representations of an x with two decimal representations are the same, observe

$$\begin{aligned}
.x_1x_2x_3\cdots x_N999999\cdots_{[10]} &= \sum_{n=0}^{N}\frac{x_n}{10^n} + \sum_{n=N+1}^{\infty}\frac{9}{10^n}\\
&= \sum_{n=0}^{N}\frac{x_n}{10^n} + \frac{9}{10^{N+1}}\left(\frac{1}{1-\frac{1}{10}}\right)\\
&= \sum_{n=0}^{N}\frac{x_n}{10^n} + \frac{9}{10^{N+1}}\left(\frac{10}{9}\right)\\
&= \sum_{n=0}^{N}\frac{x_n}{10^n} + \frac{1}{10^N}\\
&= \sum_{n=0}^{N-1}\frac{x_n}{10^n} + \frac{x_N+1}{10^N}\\
&= .x_1x_2x_3\cdots x_{N-1}(x_N+1)00000\cdots_{[10]}. \square
\end{aligned}$$

Corollary 4.12 *$\mathbb{R}$ is uncountable.*

PROOF We will again prove $[0,1]$ is uncountable, but by a slightly different method as before.

Since $[0,1] \subseteq \mathbb{R}$, it will follow that $\mathbb{R}$ is uncountable.
To show this, suppose not, i.e. that $[0,1]$ is countable.
That means $\exists$ injection $f : [0,1] \to \mathbb{N}$.
Take each number $f^{-1}(n)$ and write its decimal representation (if it has more than one decimal representation, just choose one arbitrarily):

$$\begin{aligned}
f^{-1}(1) &= .\boxed{x_{11}}x_{12}x_{13}x_{14}\cdots_{[10]}\\
f^{-1}(2) &= .x_{21}\boxed{x_{22}}x_{23}x_{24}\cdots_{[10]}\\
f^{-1}(3) &= .x_{31}x_{32}\boxed{x_{33}}x_{34}\cdots_{[10]}\\
f^{-1}(4) &= .x_{41}x_{42}x_{43}\boxed{x_{44}}\cdots_{[10]}\\
f^{-1}(5) &= .x_{51}x_{52}x_{53}x_{54}\quad\ddots\ {}_{[10]}\\
&\ \ \vdots \qquad\qquad \vdots\\
f^{-1}(n) &= .x_{n1}x_{n2}x_{n3}x_{n4}\quad\cdots\quad\boxed{x_{nn}}\cdots_{[10]}\\
&\ \ \vdots \qquad\qquad \vdots
\end{aligned}$$

Now, choose numbers $y_1, y_2, y_3, \ldots \in \{1, \ldots, 8\}$ such that $\forall\, n$,

We obtain $y = .y_1 y_2 y_3 \cdots_{[10]} \in [0,1]$.
y has only one decimal representation, since it has no 0s or 9s as digits.
Furthermore, y cannot be any of the $f^{-1}(n)$, because it is different from $f^{-1}(n)$ in the n^{th} decimal place.
This contradicts $[0,1]$ being countable. $\square$

Representation in other bases

There's nothing special about the choice of 10 as a base (other than that we have 10 fingers and 10 toes).

You can prove the same theorems we just discussed for any base $b \in \{2, 3, 4, \ldots\}$ with exactly the same arguments as before (mostly, just replace the 10s with bs):

Theorem 4.13 *Let $b \in \{2, 3, 4, 5, \ldots\}$.*

Every $x \in \mathbb{R}$ has a **base** b **representation***, meaning a sequence $\{x_n\}$ in $\{0, 1, 2, \ldots, b-1\}$ such that*

$$x = \lfloor x \rfloor . x_1 x_2 x_3 x_4 \cdots_{[b]} = \lfloor x \rfloor + \sum_{n=1}^{\infty} \frac{x_n}{b^n}.$$

If $b = 2$, we call this a **binary representation** *of x, and if $b = 3$, we call this a* **ternary representation** *of x.*

If $x = \dfrac{a}{b^N}$ for some $N \in \{1, 2, 3, \ldots\}$ and some $a \in \{0, 1, 2, \ldots, b^N\}$, then x has exactly two base b representations:

$$x = \lfloor x \rfloor . x_1 x_2 x_3 \cdots x_{N-1} x_N (b-1)(b-1)(b-1)(b-1) \cdots_{[b]}$$

and

$$x = \lfloor x \rfloor . x_1 x_2 x_3 \cdots x_{N-1} (x_N + 1) 000000 \cdots_{[b]};$$

otherwise, the base b representation of x is unique.

EXAMPLES

What are all the base 6 representations of $\dfrac{455}{216}$?

What numbers have a base 7 representation which begins $.32\ldots.._{[7]}$?

What rational number has base 3 representation $.12121212\ldots_{[3]}$?

The Cantor function

Definition 4.14 *The* **Cantor function** $c : [0,1] \to [0,1]$ *is the function defined as follows:*

Step 1: *Let $x \in [0,1]$ have ternary (base 3) representation*

$$x = .x_1x_2x_3x_4\cdots_{[3]}.$$

Step 2: *If any of the digits x_n are 1, replace all the digits after the first 1 with 0.*

Step 3: *Replace any of the digits (before the first 1) that are 2s with 1s. (In other words, divide all the digits before the first 1 by 2.)*

Step 4: *Treat the string $.y_1y_2y_3\cdots_{[2]}$ as a binary (base 2) representation of a real number. The result is $c(x)$.*

EXAMPLE

Compute $c(x)$ if $x = .020221021021012201\ldots_{[3]}$.

EXAMPLE

Compute $c\left(\dfrac{2}{3}\right)$.

Theorem 4.15 *The Cantor function is well-defined.*
(This means that if x has two different ternary representations, then $c(x)$ does not depend on which ternary representation you take in Step 1 of computing $c(x)$).

PROOF Suppose you took $x \in [0,1]$ with two different ternary representations. That means those representations must be

$$x = .x_1x_2x_3 \cdots x_{n-1}x_n 222222 \cdots_{[3]} = .x_1x_2x_3 \cdots x_{n-1}(x_n+1)0000 \cdots_{[3]}$$

for some string of ternary digits $x_1, x_2, ..., x_n$ (where $x_n \neq 2$).

Case 1: $x_k = 1$ for some $k < n$.
In this situation, the different digits in these two representations get turned into 0s in Step 1, so both representations yield the same value of $c(x)$.

Case 2: $x_k \neq 1$ for all $k < n$, but $x_n = 1$. Here, using the first form of x, we get

Step 1: $x = .x_1x_2x_3 \cdots x_{n-1}12222222222 \cdots_{[3]}$
Step 2: $.x_1x_2x_3 \cdots x_{n-1}10000000000 \cdots_{[3]}$
Step 3: $.\left(\frac{x_1}{2}\right)\left(\frac{x_2}{2}\right)\left(\frac{x_3}{2}\right)\left(\frac{x_1}{2}\right) \cdots \left(\frac{x_{n-1}}{2}\right)10000000 \cdots_{[2]}$

Using the second form of x, we get

Step 1: $x = .x_1x_2x_3 \cdots x_{n-1}20000000000 \cdots_{[3]}$
Step 2: $.x_1x_2x_3 \cdots x_{n-1}20000000000 \cdots_{[3]}$
Step 3: $.\left(\frac{x_1}{2}\right)\left(\frac{x_2}{2}\right)\left(\frac{x_3}{2}\right)\left(\frac{x_1}{2}\right) \cdots \left(\frac{x_{n-1}}{2}\right)10000000 \cdots_{[2]}$

After Step 3, the two forms produce the same binary representation. Therefore they yield the same value of $c(x)$.

Case 3: $x_k \neq 1$ for all $k < n$ and $x_n = 0$.
Here, using the first form of x, we get

Step 1: $x = .x_1x_2x_3 \cdots x_{n-1}02222222222 \cdots_{[3]}$
Step 2: $.x_1x_2x_3 \cdots x_{n-1}02222222222 \cdots_{[3]}$
Step 3: $.\left(\frac{x_1}{2}\right)\left(\frac{x_2}{2}\right)\left(\frac{x_3}{2}\right)\left(\frac{x_1}{2}\right) \cdots \left(\frac{x_{n-1}}{2}\right)0111111111 \cdots_{[2]}$
Step 4: $c(x) = \sum_{k=1}^{n-1} \frac{x_k/2}{2^k} + \sum_{k=n+1}^{\infty} \frac{1}{2^k} = \boxed{\sum_{k=1}^{n-1} \frac{x_k/2}{2^k} + \frac{1}{2^n}}.$

Using the second form of x, we get

Step 1: $x = .x_1x_2x_3\cdots x_{n-1}10000000000\cdots_{[3]}$

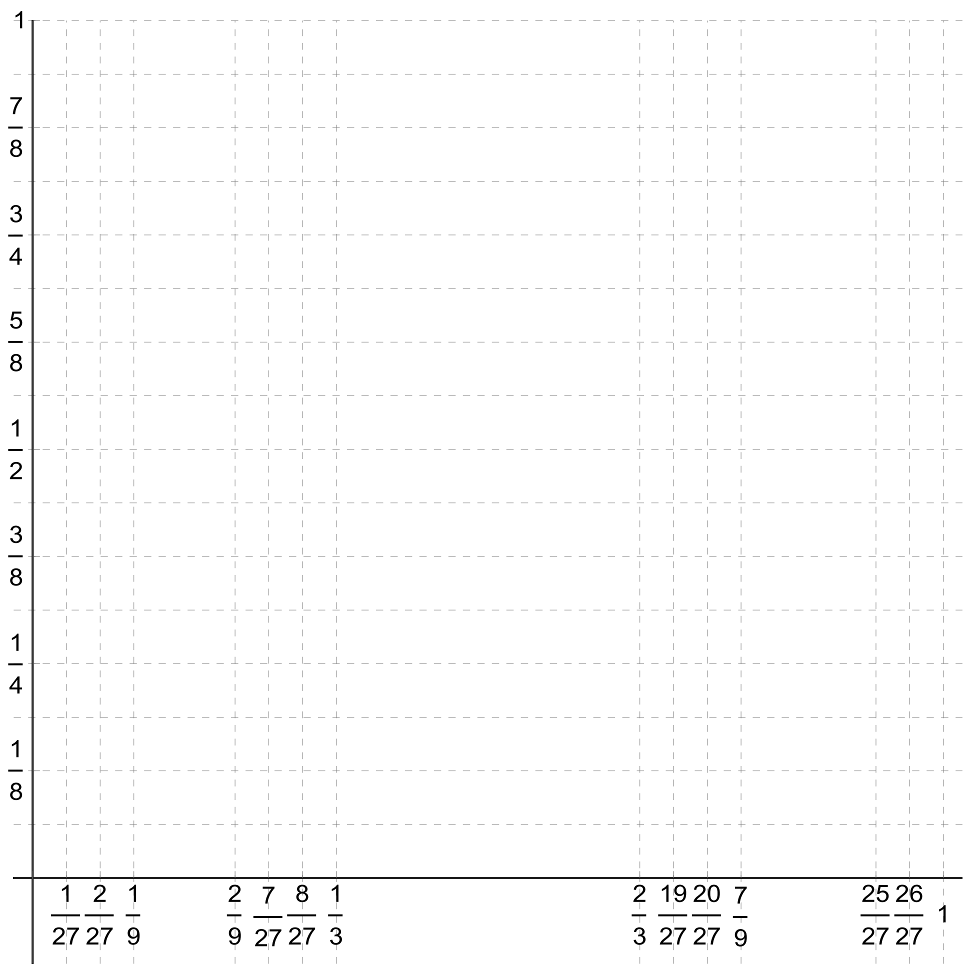
Step 2: $.x_1x_2x_3\cdots x_{n-1}10000000000\cdots_{[3]}$

Step 3: $.\left(\frac{x_1}{2}\right)\left(\frac{x_2}{2}\right)\left(\frac{x_3}{2}\right)\left(\frac{x_1}{2}\right)\cdots\left(\frac{x_{n-1}}{2}\right)10000000\cdots_{[2]}$

Step 4: $c(x) = \boxed{\sum_{k=1}^{n-1}\frac{x_k/2}{2^k} + \frac{1}{2^n}}.$

Notice that you get the same thing for $c(x)$. □

The graph of the Cantor function



Theorem 4.16 *The Cantor function $c : [0,1] \to [0,1]$ is surjective.*

PROOF Let $y \in [0,1]$. We need to find $x \in [0,1]$ such that $c(x) = y$.

To do this, write y in binary as

$$y = .y_1 y_2 y_3 \cdots_{[2]}.$$

Now, let $x_n = 2y_n$ for each n, and consider the number

$$x = .x_1 x_2 x_3 \cdots_{[3]}.$$

Notice that the ternary expansion of x has no 1s in it. So

$$c(x) = .\left(\frac{x_1}{2}\right)\left(\frac{x_2}{2}\right)\left(\frac{x_3}{2}\right)\left(\frac{x_1}{2}\right)\cdots_{[2]} = .y_1 y_2 y_3 \cdots_{[2]} = y.$$

Thus c is surjective as wanted. □

Theorem 4.17 *The Cantor function $c : [0,1] \to [0,1]$ is increasing.*

PROOF Let $x, y \in [0,1]$ be such that $x < y$. Write x and y in ternary:

$$x = .x_1 x_2 x_3 \cdots_{[3]}$$
$$y = .y_1 y_2 y_3 \cdots_{[3]}$$

WLOG, if x and/or y happen to have two ternary representations, choose the one that ends in all 0s, not all 2s.

Our goal is to show $c(x) \le c(y)$.

Now, let k be the smallest index such that $x_k \ne y_k$ (such a k exists since $x \ne y$).

For this k, since $x \le y$, $x_k < y_k$.

Case 1: $x_j = 1$ for some $j < k$.

In this situation, when doing Step 2 of the Cantor function, we get the same string of digits for x and y. Ultimately, this yields $c(x) = c(y)$.

Case 2: $x_j \ne 1$ for all $j < k$, $x_k = 1$ and $y_k = 2$.

In this case, when doing Step 2 on x and y, we obtain (respectively)

$$.x_1 x_2 x_3 \cdots x_{k-1} 100000 \cdots_{[3]} \quad \text{and} \quad .x_1 x_2 x_3 \cdots x_{k-1} 2 y_{k+1} y_{k+2} \cdots_{[3]}$$

and when the digits before the first one are halved, we get

$$.\left(\frac{x_1}{2}\right)\left(\frac{x_2}{2}\right)\left(\frac{x_3}{2}\right)\cdots\left(\frac{x_{k-1}}{2}\right) 100000 \cdots_{[2]}$$

from x and

$$.x_1x_2x\left(\frac{x_1}{2}\right)\left(\frac{x_2}{2}\right)\left(\frac{x_3}{2}\right)\cdots\left(\frac{x_{k-1}}{2}\right)1a_{k+1}a_{k+2}\cdots_{[2]}$$

from y.

Thus

$$c(x) = \sum_{j=1}^{k-1}\frac{(x_j/2)}{2^j} + \frac{1}{2^k}$$

and

$$\begin{aligned} c(y) &= \sum_{j=1}^{k-1}\frac{(x_j/2)}{2^j} + \frac{1}{2^k} + \sum_{j=k+1}^{\infty}\frac{a_j}{2^j} \\ &\geq \sum_{j=1}^{k-1}\frac{(x_j/2)}{2^j} + \frac{1}{2^k} \\ &= c(x). \end{aligned}$$

Case 3: $x_j \neq 1$ for all $j < k$, $x_k = 0$ and $y_k = 1$.

In this case, when doing Step 2 on x and y we obtain

$$.x_1x_2x_3\cdots x_{k-1}0x_{k+1}x_{k+2}\cdots_{[3]} \quad \text{and} \quad .x_1x_2x_3\cdots x_{k-1}10000\cdots_{[3]}$$

and when the digits before the first one are halved, we get

$$.\left(\frac{x_1}{2}\right)\left(\frac{x_2}{2}\right)\left(\frac{x_3}{2}\right)\cdots\left(\frac{x_{k-1}}{2}\right)0a_{k+1}a_{k+2}\cdots_{[2]}$$

from x and

$$.x_1x_2x\left(\frac{x_1}{2}\right)\left(\frac{x_2}{2}\right)\left(\frac{x_3}{2}\right)\cdots\left(\frac{x_{k-1}}{2}\right)100000\cdots_{[2]}$$

from y.

Thus

$$\begin{aligned} c(x) &= \sum_{j=1}^{k-1}\frac{(x_j/2)}{2^j} + \sum_{j=k+1}^{\infty}\frac{a_j}{2^j} \\ &\leq \sum_{j=1}^{k-1}\frac{(x_j/2)}{2^j} + \sum_{j=k+1}^{\infty}\frac{1}{2^j} \\ &= \sum_{j=1}^{k-1}\frac{(x_j/2)}{2^j} + \frac{1}{2^k} \\ &= c(y). \end{aligned}$$

In all three cases we have shown $c(x) \leq c(y)$ (under the assumption $x < y$), making c increasing. □

4.3 More on infinite series

Theorem 4.18 (Linearity of Convergence of Infinite Series) *Let $\{a_n\}$ and $\{b_n\}$ be sequences of real numbers, and let $r \in \mathbb{R}$. If $\sum a_n = S$ and $\sum b_n = T$, then*

$$\sum (a_n + b_n) = S + T \qquad \text{and} \qquad \sum r\, a_n = rS.$$

PROOF HW

Hint: Let S_N and T_N be the N^{th} partial sums of $\sum a_n$ and $\sum b_n$, respectively.
By definition of convergent series, $S = \lim S_N$ and $T = \lim T_N$.
What are the partial sums of $\sum (a_n + b_n)$? What's true about them, and why?

Theorem 4.19 (Triangle Inequality for Infinite Series) *Let $\{a_n\}$ be a sequence of real numbers. If $\sum |a_n|$ converges, then so does $\sum a_n$.*

PROOF First, denote by S_N the partial sums of $|a_n|$.
Since $|a_n| \geq 0$ for all n, S_N is an increasing sequence.
Since $\sum |a_n|$ converges (let's say to S), it must be that $S_N \leq S$ for all N.
For each n, let

$$a_n^+ = \max\{a_n, 0\} = \begin{cases} a_n & \text{if } a_n \geq 0 \\ 0 & \text{if } a_n < 0 \end{cases}$$

and

$$a_n^- = \min\{a_n, 0\} = \begin{cases} a_n & \text{if } a_n \leq 0 \\ 0 & \text{if } a_n > 0 \end{cases} .$$

Notice $a_n^+ + a_n^- = a_n$ and $a_n^+ - a_n^- = |a_n|$.
Now, consider the series $\sum a_n^+$. First, $a_n^+ \geq 0$ for all n, so the partial sums S_N^+ of $\sum a_n^+$ form an increasing sequence. Also, notice that $0 \leq a_n^+ \leq |a_n|$ for all n, so by taking partial sums, $S_N^+ \leq S_N \leq S$. By the MCT, $\{S_N^+\}$ converges, i.e. $\sum a_n^+$ converges.
Since $\sum |a_n|$ converges and $\sum a_n^+$ converges, it follows that

$$\sum a_n^- = \sum (a_n^+ - |a_n|)$$

also converges. But then $\sum a_n = \sum (a_n^+ + a_n^-)$ converges as well. □

Theorem 4.20 (Ratio Test) *Let $\{a_n\}$ be a sequence of real numbers.*

1. *If* $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} < 1$, *then* $\sum a_n$ *converges.*
2. *If* $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} > 1$, *then* $\sum a_n$ *diverges.*

If $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$ *does not exist or equals* 1, *then no conclusion can be drawn from this theorem.*

PROOF We prove the first statement here. Let $L = \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$ and suppose $L < 1$.

Let $r = \frac{1}{2}(1+L)$; notice that $r < 1$.

Since $\frac{|a_{n+1}|}{|a_n|} \to L$, there is $N \geq 0$ so that for $n \geq N$,

$$\left| \frac{|a_{n+1}|}{|a_n|} - L \right| < \frac{1}{2}(1-L).$$

which implies that when $n \geq N$, $\frac{|a_{n+1}|}{|a_n|} < r$, which rearranges into

$$|a_{n+1}| < r|a_n|.$$

Thus for $k \geq N$, $|a_k| < r^{k-N}|a_N|$.

Now, when $n \geq N$, the n^{th} partial sum of $\sum |a_n|$ is

$$\begin{aligned}
S_N = \sum_{k\leq n} |a_k| &= \sum_{k<N} |a_k| + \sum_{k=N}^{n} |a_k| \\
&\leq \sum_{k<N} |a_k| + \sum_{k=N}^{n} r^{k-N}|a_N| \\
&\leq \sum_{k<N} |a_k| + \sum_{k=N}^{\infty} r^{k-N}|a_N| \\
&= \sum_{k<N} |a_k| + |a_N| \sum_{k=N}^{\infty} r^{k-N} \\
&= \sum_{k<N} |a_k| + |a_N| \sum_{k=0}^{\infty} r^{k} && \text{(index change)} \\
&= \sum_{k<N} |a_k| + |a_N| \frac{1}{1-r}. && \text{(geometric series formula)}
\end{aligned}$$

Since all the $|a_k|$ are non-negative, $\{S_N\}$ is an increasing sequence, bounded

above by the finite number $\left[\sum_{k<N} |a_k| + |a_N| \frac{1}{1-r}\right]$.
Thus $\{S_N\}$ converges by the MCT, i.e. $\sum |a_n|$ converges.
$\sum a_n$ therefore converges by the Triangle Inequality for infinite series.

The proof of the second statement is left as HW. It has a similar proof as the first statement, except that here the goal is to show that the partial sums of $\sum a_n$ are unbounded. □

4.4 Convergence of power series; transcendental functions

First, we can repeat all the definitions of Section 4.1 in the context of series made up of *functions* rather than numbers:

Definition 4.21 *Let $\{f_n\}_{n=0}^{\infty}$ be a sequence of functions from E to $\mathbb{R}$.*
The sequence $\{S_N\}$ of partial sums associated to $\{f_n\}$ is the sequence of functions defined as follows:

$$\begin{aligned} S_0(x) &= f_0(x) \\ S_1(x) &= f_0(x) + f_1(x) \\ S_2(x) &= f_0(x) + f_1(x) + f_2(x) \\ &\vdots \\ S_N(x) &= \sum_{n\leq N} f_n(x) = \sum_{n=0}^{N} f_n(x). \end{aligned}$$

If the sequence $\{S_N\}$ converges (pointwise) to $f : E \to \mathbb{R}$, we say $\sum f_n$ **converges (pointwise) (to f) on** *E and write $\sum f_n = f$ on E.*
We say the series $\sum f_n$ converges uniformly to f if $S_N \rightrightarrows f$ on E.

We often make series of of power functions:

Definition 4.22 *Let $\{a_n\}_{n=0}^{\infty}$ be a sequence of real numbers. Then the infinite series of functions $\sum_{n=0}^{\infty} a_n x^n$ is called a* **power series**.

Theorem 4.23 *Let* $\sum_{n=0}^{\infty} a_n x^n$ *be a power series with* $\lim_{n\to\infty} \sqrt[n]{|a_n|} = 0$. *Then:*

1. $a_n x^n$ *converges pointwise on* $\mathbb{R}$ *to a function* $f : \mathbb{R} \to \mathbb{R}$.
2. *For any compact subset* E *of* $\mathbb{R}$, $\sum a_n x^n$ *converges uniformly to* f *on* E.

PROOF To get started, we first prove that the sequence $\{a_n\}$ must be bounded.

To verify this, note that by hypothesis $\sqrt[n]{|a_n|} \to 0$.

Therefore $\exists K \in \mathbb{N}$ s.t. $n \geq K$ implies $\left|\sqrt[n]{|a_n|}\right| < \frac{1}{2}$, i.e. $|a_n| < \left(\frac{1}{2}\right)^n \leq 1$.

That means that the entire sequence $\{a_n\}$ is bounded by

$$A = \max\{|a_0|, |a_1|, |a_2|, \ldots, |a_K|, 1\}.$$

Now, we prove statement (2).

Let $E \subseteq \mathbb{R}$ be compact; thus E is bounded so $E \subseteq [-B, B]$ for some $B \geq 1$.

Now, let $\epsilon > 0$. WLOG $\epsilon \in (0,1)$.

Given this ϵ, choose $\delta \in (0,1)$ so that $\frac{\delta^K}{1-\delta} < \epsilon$.

(This can be done because $\lim_{\delta\to 0} \frac{\delta^K}{1-\delta} = 0$.)

Next, choose $L \geq K$ so that $n \geq L$ implies $\sqrt[n]{|a_n|} < \frac{\delta}{B}$, i.e. $|a_n| < \left(\frac{\delta}{B}\right)^n$.

Now, let $N > M \geq L$. We have, for any $x \in E$, $|x| \leq B$ so

$$\begin{aligned}
\left|\begin{pmatrix} N^{th} \text{ partial} \\ \text{sum of } \sum a_n x^n \end{pmatrix} - \begin{pmatrix} M^{th} \text{ partial} \\ \text{sum of } \sum a_n x^n \end{pmatrix}\right| &= \left|\sum_{n=0}^{N} a_n x^n - \sum_{n=0}^{M} a_n x^n\right| \\
&= \left|\sum_{n=M+1}^{N} a_n x^n\right| \\
&\leq \sum_{n=M+1}^{N} |a_n|\,|x|^n \\
&< \sum_{n=M+1}^{N} \left(\frac{\delta}{B}\right)^n B^n \\
&\leq \sum_{n=K}^{\infty} \delta^n = \frac{\delta^K}{1-\delta} < \epsilon.
\end{aligned}$$

Therefore the partial sums of $\sum a_n x^n$ are uniformly Cauchy on E.
That means $\sum a_n x^n$ converges uniformly on E. This proves (2).

Finally, for statement (1), let $x \in \mathbb{R}$. x is contained in the compact interval $E = [-|x| - 1, |x| + 1]$, and $\sum a_n x^n$ converges uniformly on E, so it must converge pointwise on E (and in particular at x). □

Transcendental functions

Theorem 4.24 *Let $x \in \mathbb{R}$. Then, the following series all converge pointwise on $\mathbb{R}$, and converge uniformly on compact subsets of $\mathbb{R}$:*

- $\displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \ldots$
- $\displaystyle\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots$
- $\displaystyle\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots$

Remark: When $x = 0$, the $n = 0$ term of a series is problematic, but we can plug in $x = 0$ to the written out form of the series to see that...

PROOF For the first series, we start with this claim:

Claim: For any $a > 0$, there exists $N \in \mathbb{N}$ so that $n \geq N$ implies $n! > a^n$.

Proof of claim: Given $a > 0$, choose $N \geq a$ so that $\left(\frac{a+1}{a}\right)^N > \frac{(a+1)^a}{a!}$, which implies the following chain of inequalities:

$$\begin{aligned} \left(\frac{a+1}{a}\right)^{N-a} &> \frac{(a+1)^a}{a!} \cdot \left(\frac{a+1}{a}\right)^{-a} \\ \Rightarrow \frac{(a+1)^{N-a}}{a^{N-a}} &> \frac{a^a}{a!} \\ \Rightarrow (a+1)^{N-a} a! &> a^N. \end{aligned}$$

Now, for $n \geq N$,

$$\begin{aligned} n! &= n(n-1)(n-2)\cdots(a+2)(a+1)\, a(a-1)\cdots 3 \cdot 2 \cdot 1 \\ &\geq (a+1)(a+1)(a+1)\cdots(a+1)(a+1)\, a! \\ &= (a+1)^{N-a} a! \\ &> a^N \text{ (from above)}. \end{aligned}$$

Now, let $\epsilon > 0$.

Applying the claim with $a = \frac{1}{\epsilon}$, we can find N so that for all $n \geq N$,

$$\left| \sqrt[n]{|a_n|} - 0 \right| = \sqrt[n]{\frac{1}{n!}} \leq \sqrt[n]{\frac{1}{(1/\epsilon)^n}} = \epsilon.$$

This means $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} = 0$. The result then follows by Theorem 4.23.

The other two series are left as HW. □

Definition 4.25 *Given $x \in \mathbb{R}$, define*

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

This defines a function $\exp : \mathbb{R} \to \mathbb{R}$ *called the* **(natural) exponential function**. *We define the number e, called* **Euler's number**, *to be* $e = \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!}$.

What we know about $\exp$ at this point

- $\exp(x)$ is defined for every $x \in \mathbb{R}$, and the power series defining $\exp$ converges uniformly on any compact subset of $\mathbb{R}$.
- $\exp(0) = 1$ (just plug in $x = 0$ to the definition to see this).
- $\lim_{n\to\infty} \exp(n)$ DNE (this sequence is unbounded since $\exp(x) \geq 1+x$ when $x \geq 0$).

What we don't know about $\exp$ right now

- We don't know anything about the numerical value of e (other than $e > 1$).
- We don't know $e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$.
- We don't know $\exp(x) = e^x \ \forall\, x \in \mathbb{R}$.
- We don't know $\exp(x + y) = \exp(x)\exp(y) \ \forall\, x, y \in \mathbb{R}$ (and other exponent rules)
- We don't know $\exp$ is differentiable (or continuous, or integrable)
- We don't know $\exp$ is increasing
- We don't know $\lim_{n\to-\infty} \exp(n) = 0$
- We don't know how big $\exp(x)$ is compared to other functions that increase without bound as $x \to \infty$, like x or x^2 or x^x
- We don't know e is invertible, or anything else about logarithms

Definition 4.26 *Define functions* $\sin : \mathbb{R} \to \mathbb{R}$ *and* $\cos : \mathbb{R} \to \mathbb{R}$ *by*

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \qquad \text{and} \qquad \cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}.$$

These functions are respectively called **sine** *and* **cosine**.

What we know about $\sin$ **and** $\cos$ **at this point**

- $\sin x$ and $\cos x$ are defined for every $x \in \mathbb{R}$, and the power series defining these functions converge uniformly on any compact subset of $\mathbb{R}$.
- $\sin 0 = 0$
- $\cos 0 = 1$
- $\sin(-x) = -\sin x$ (just plug in $-x$ to the definition of $\sin$)
- $\cos(-x) = \cos x$ (just plug in $-x$ to the definition of $\cos$)

What we don't know about $\sin$ **and** $\cos$ **right now**

- We don't know any connection with our $\sin/\cos$ defined here and triangles or the unit circle
- We don't know $\cos^2 x + \sin^2 x = 1$ (or any other identities)
- We don't have any knowledge of values of $\sin x$ or $\cos x$ when $x \neq 0$
- We don't know $-1 \le \sin x \le 1$, $-1 \le \cos x \le 1$
- We don't know $\frac{d}{dx}(\sin x) = \cos x$ or $\frac{d}{dx}(\cos x) = -\sin x$ (we don't even know $\sin$ and/or $\cos$ are differentiable, integrable or even continuous)

4.5 Chapter 4 Summary

Definitions to know

Nouns

- The **partial sums** of $\sum_{n=1}^{\infty} a_n$ are the sequence $\{S_N\}_N$ where $S_N = a_1 + a_2 + \ldots + a_N$.
- The **floor** of $x \in \mathbb{R}$ is the largest integer less than or equal to x; this is denoted $\lfloor x \rfloor$.
- Fix $b \in \{2, 3, 4, \ldots\}$. A **base** b **representation** of $x \in \mathbb{R}$ is a sequence $\{x_n\} \subseteq \{0, 1, 2, \ldots, b-1\}$ so that

$$x = \lfloor x \rfloor . x_1 x_2 x_3 x_4 \cdots_{[b]} = \lfloor x \rfloor + \sum_{n=1}^{\infty} \frac{x_n}{b^n}.$$

  A **binary representation** means a base 2 representation.
  A **ternary representation** means a base 3 representation.
  A **decimal representation** means a base 10 representation.
- (★) A **power series** is an infinite series $\sum a_n x^n$, where $\{a_n\}$ is a sequence of numbers.

Adjectives that describe subsets of $\mathbb{R}$

- To say $\sum a_n$ **converges** to S (i.e. $\sum a_n = S$) means $S_N \to S$, where $\{S_N\}$ are the partial sums of $\sum a_n$.
- To say $\sum a_n$ **diverges** means $\sum a_n$ does not converge to any number $S \in \mathbb{R}$.

Theorems with names

Geometric series formula If $r \in (-1, 1)$, then $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$.

Finite geometric sum formula For any $r \in \mathbb{R}$, $\sum_{n=M}^{N} r^n = r^M \left(\frac{1 - r^{N-M+1}}{1-r} \right)$.

Uniqueness of base b representations Every $x \in \mathbb{R}$ has a base b representation.

x has a unique base b representation unless $x = \frac{a}{b^N}$ for some $a \in \mathbb{Z}$ and $N \in \mathbb{N}$, in which case x has exactly two base b representations that look like

$$x = \lfloor x \rfloor . x_1 x_2 x_3 \cdots x_{N-1} x_N (b-1)(b-1)(b-1) \cdots_{[b]}$$

and

$$x = \lfloor x \rfloor . x_1 x_2 x_3 \cdots x_{N-1} (x_N + 1) 00000000 \cdots_{[b]}.$$

(★) **Linearity of Convergence of Infinite Series** If $\sum a_n = S$ and $\sum b_n = T$, then $(a_n + b_n) = S + T$ and $\sum (r\, a_n) = rS$ for any $r \in \mathbb{R}$.

(★) **Triangle Inequality for Infinite Series** If $\sum |a_n|$ converges, so does $\sum a_n$.

(★) **Ratio Test** If $\lim\limits_{n\to\infty} \dfrac{|a_{n+1}|}{|a_n|} < 1$, then $\sum a_n$ converges.

If $\lim\limits_{n\to\infty} \dfrac{|a_{n+1}|}{|a_n|} > 1$, then $\sum a_n$ diverges.

Other theorems to remember

- The harmonic series $\sum \dfrac{1}{n}$ diverges.
- The p-series $\sum \dfrac{1}{n^p}$ converges if $p \geq 2$.

  (In fact, this series converges if $p > 1$ (HW).)
- (★) If $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} = 0$, then the power series $\sum a_n x^n$ converges pointwise on $\mathbb{R}$ and converges uniformly on compact subsets of $\mathbb{R}$.

(★) Series definitions of transcendental functions

- $\exp(x) = \sum\limits_{n=0}^{\infty} \dfrac{x^n}{n!}$

  $\exp 0 = 1$; $\exp x > 0 \; \forall \, x$; $\exp$ is increasing
- $\cos(x) = \sum\limits_{n=0}^{\infty} (-1)^n \dfrac{x^{2n}}{(2n)!}$

  $\cos 0 = 1$; $\cos(-x) = \cos x$
- $\sin(x) = \sum\limits_{n=0}^{\infty} (-1)^n \dfrac{x^{2n+1}}{(2n+1)!}$

  $\sin 0 = 0$; $\sin(-x) = -\sin x$

What we have learned about the Cantor function c

- $c(x)$ is defined by taking a base 3 representation of x, deleting any digits after the first 1, replacing any remaining 2s with 1s, and intepreting the resulting string as a binary representation of $c(x)$.
- $c : [0,1] \to [0,1]$ is well-defined, surjective and increasing.

4.6 Chapter 4 Homework

Exercises from Section 4.1

In Exercises 1-3, we finish the proof of the geometric series formula by verifying that if $r \in (-1, 1)$, then $r^n \to 0$. (In Chapter 2, we proved this is true when $r \in \left(-\frac{1}{2}, \frac{1}{2}\right)$, but we don't know this yet when $r \in \left(-1, -\frac{1}{2}\right] \cup \left[\frac{1}{2}, 1\right)$.)

1. Prove **Bernoulli's inequality**, which says that for any $\delta > 0$ and any $n \in \mathbb{N}$, $(1 + \delta)^n \geq 1 + \delta n$.

   Hints: Fix $\delta > 0$. To prove Bernoulli's inequality, suppose not. Then there is a smallest n so that $(1 + \delta)^n < 1 + \delta n$. Check that this n cannot be 0, which means that $n - 1 \in \mathbb{N}$. Since $n - 1$ is less than n, and n is the smallest n for which Bernoulli's inequaity is false, Bernoulli's inequality must be true for exponent $n - 1$, i.e. $(1 + \delta)^{n-1} \geq 1 + \delta(n - 1)$. We now have these four inequalities:

   $$n \geq 1 \qquad (1 + \delta)^{n-1} \geq 1 + \delta(n - 1) \qquad \delta > 0 \qquad (1 + \delta)^n < 1 + \delta n$$

   A contradiction can be derived from these inequalities.

2. Prove $s > 1$ implies $\{s^n : n \in \mathbb{N}\}$ is unbounded.

   Hints: Suppose not, then there is B so that $s^n \leq B$ for all n. Write $s = 1+\delta$; since $\delta > 0$, we can apply Bernoulli's inequality; this will lead to a contradiction (related to the Archimedean Property).

3. Prove $r^n \to 0$ when $r \in (-1, 1)$.

   Hints: If $r \in (-1, 1)$, we can apply Exercise 2 to show $\left\{\frac{1}{|r|^n} : n \in \mathbb{N}\right\}$ is unbounded. This will help you choose your N when you write an ϵ-proof of $r^n \to 0$.

4. a) Prove that the geometric series $\sum_{n=0}^{\infty} r^n$ diverges when $r = -1$.

   b) Prove that the geometric series $\sum_{n=0}^{\infty} r^n$ diverges when $r = 1$.

5. Prove that the geometric series $\sum_{n=0}^{\infty} r^n$ diverges when $|r| > 1$.

   Hint: Use Bernoulli's inequality (Exercise 1) to show that the partial sums of this series are unbounded.

6. Prove that for any $p > 1$, the p-series $\sum \frac{1}{n^p}$ converges.

Hints: The situation where $p \geq 2$ was handled earlier in the chapter. In this exercise, argue similar to the proof of Theorem 4.5 by first showing that $S_{2^k-1} \leq \dfrac{1}{1-\frac{1}{2^{p-1}}}$. Then apply the MCT.

7. Consider the series $\sum_{n=0} \infty \dfrac{1}{n+1} n+2$.

   a) Find an explicit formula for the N^{th} partial sum of this series.
   Hint: Rewrite the terms of this sequence using partial fractions, and then write the partial sum out. A lot of the terms will cancel.

   b) Use part (a) to show that the series converges and find its sum.

8. Determine whether the series $\displaystyle\sum_{n=1}^{\infty} \frac{n^{n+1}}{(n+1)^n}$ converges or diverges.

   Hint: Find an inequality relating $\dfrac{n^{n+1}}{(n+1)^n}$ and the n^{th} term of a series we studied in this chapter.

9. Prove the **Comparison Test** for infinite series, which says that if $0 \leq a_n \leq b_n$ for all n, then

   - $\sum b_n$ converges $\Rightarrow \sum a_n$ converges, and
   - $\sum a_n$ converges $\Rightarrow \sum b_n$ diverges.

   Hints: For the first statement, apply the MCT to the partial sums of $\sum a_n$. The second statement is the contrapositive of the first.

10. (★) Prove that the alternating harmonic series $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ converges.

    Hints: Let S_N be the N^{th} partial sum of this series. Use the MCT to show that the subsequences $\{S_{2k}\}$ and $\{S_{2k+1}\}$ converge, respectively, to $\underline{\lim}\, S_N$ and $\overline{\lim}\, S_N$. Then, show that $S_{2k+1} - S_{2k} \to 0$, which proves that $\underline{\lim}\, S_N = \overline{\lim}\, S_N$.

11. Prove the n^{th} **term test for divergence**, which says that if $\{a_n\}$ is a sequence of numbers which does <u>not</u> converge to 0, then $\sum a_n$ diverges.

    Hints: Suppose not, i.e. that $\sum a_n$ converges. In this situation ,what must be true about the subsequences $\{S_n\}$ and $\{S_{n+1}\}$? Consequently, what must be true about $\{a_n\}$?

Exercises from Section 4.2

12. Compute each quantity:

a) $\lfloor 2.25 \rfloor + \lfloor -2.25 \rfloor$

b) $\lfloor \sqrt{70} \rfloor$

c) $2\lfloor \pi \rfloor$

d) $4 + \left\lfloor \frac{13}{5} \right\rfloor$

13. a) Write down two real numbers x and y so that $\lfloor x + y \rfloor = \lfloor x \rfloor + \lfloor y \rfloor$.
    b) Write down two real numbers x and y so that $\lfloor x + y \rfloor \neq \lfloor x \rfloor + \lfloor y \rfloor$.

14. Let $b \in \{2, 3, 4, \ldots\}$. Prove that the set of numbers that do not have a unique base b representation is countable.

    Hint: Establish that this set is the countable union of countable sets.

15. Let E be the set of real numbers that have only the digits 3 and 8 in their decimal representations. Determine, with proof, whether or not E is countable.

    Hint: Look carefully at the proof that $\mathbb{R}$ is uncountable in Corollary 4.12, and how a similar argument might apply in this situation.

16. a) What numbers have a base 8 representation that starts $.12\cdots_{[8]}$?
    b) What number has base 4 representation $.123123123123123\cdots_{[4]}$?
    c) Find a base 3 representation of $\frac{3}{8}$.
    d) Find two different base 5 representations of $\frac{367}{625}$.

17. Let c be the Cantor function. For each x, compute $c(x)$, writing your answer as a rational number.

    a) $x = \frac{14}{27}$

    b) $x = \frac{19}{27}$

    c) $x = .020221020121\cdots_{[3]}$

    d) $x = \frac{5}{6}$

    e) $x = .022022022022022022\cdots_{[3]}$

    f) $x = \frac{1}{7}$

Exercises from Section 4.3

18. Prove Theorem 4.18, which says that if $\sum a_n = S$ and $\sum b_n = T$, then

    a) $\sum (a_n + b_n) = S + T$, and
    b) $\sum r\, a_n = rS$ for any constant $r \in \mathbb{R}$.

19. Prove that if $\sum a_n = S$ but $\sum b_n$ diverges, then $\sum (a_n + b_n)$ diverges.

20. Prove that if $\sum a_n$ diverges and $r \neq 0$ is a constant, then $\sum r\, a_n$ diverges.

21. a) Give an example of series $\sum a_n$ and $\sum b_n$ which both diverge, but $\sum(a_n + b_n)$ converges.

    b) Give an example of series $\sum a_n$ and $\sum b_n$ which both diverge, and $\sum(a_n + b_n)$ also diverges.

22. Suppose $\sum |a_n|$ converges and that $\{b_n\}$ is a bounded sequence of numbers. Prove $\sum a_n b_n$ converges.

23. Prove the second statement of the Ratio Test (Theorem 4.20), which says that if $\{a_n\}$ is a sequence of real numbers such that $\lim\limits_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} > 1$, then $\sum a_n$ diverges.

24. Consider the series $\sum\limits_{n=1}^{\infty} \frac{n^{2000}}{2^n}$. Prove that this series converges.

    Hint: Use the Ratio Test.

25. Prove (part of) the **Root Test**, which says that if $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} < 1$, then $\sum a_n$ converges.

    Hints: The proof of this is similar to the proof of the Ratio Test. Let $L = \lim\limits_{n\to\infty} \sqrt[n]{|a_n|}$ and then let $r = \frac{1}{2}(1+L)$. Prove that there is N so that $k \geq N$ implies $|a_k| < r^{k-N}|a_N|$; then, the rest of the proof is the same as the Ratio Test.

Exercises from Section 4.4

26. Prove that the series that defines $\sin x$ (given in Theorem 4.24 and Definition 4.26) converges for all $x \in \mathbb{R}$.

27. Prove that the series that defines $\cos x$ (given in Theorem 4.24 and Definition 4.26) converges for all $x \in \mathbb{R}$.

28. In this exercise we prove that e is an irrational number. Remember that our definition of e is that
$$e = \exp(1) = \sum_{n=1}^{\infty} \frac{1}{n!}.$$
    To prove e is irrational, carry out the following steps:

    a) Suppose not, i.e. e is rational; this means $e = \frac{p}{q}$ where $p, q > 0$ are natural numbers with no common factors. Let $z = q!\left(e - \sum\limits_{n=0}^{q} \frac{1}{n!}\right)$. Explain why $z \in \mathbb{Z}$.

    b) Explain why $z > 0$.

c) Show that whenever $n > q+1$, $\dfrac{q!}{n!} < (q+1)^{q-n}$.

d) Use part (c) to explain why $z < 1$.

Parts (a), (b) and (d) yield a contradiction, because there is no integer between 0 and 1. Thus e is irrational.

29. In this problem, we prove this generalization of Theorem 4.23, which goes like this: let $\sum_{n=0}^{\infty} a_n x^n$ be a power series and let

$$R = \left[\overline{\lim} \sqrt[n]{|a_n|}\right]^{-1};$$

R is called the **radius of convergence** of the power series.

- If $R = 0$, then $\sum a_n x^n$ diverges for any $x \neq 0$;
- if $R = \infty$, then $\sum a_n x^n$ converges for any $x \in \mathbb{R}$;
- if $0 < R < \infty$, then $\sum a_n x^n$ converges when $|x| < R$ and diverges when $|x| > R$.

To do this, carry out the following steps:

a) Show the series $\sum a_n x^n$ always converges when $x = 0$.

b) Suppose $R = 0$. Conclude that $\sqrt[n]{|a_n|}$ is unbounded. Explain why this means, for any $x \neq 0$, that $a_n x^n$ cannot converge to 0. Apply the n^{th} term test for divergence (Exercise 11) to finish the proof of the first bullet point above.

c) Suppose $R = \infty$. Let $x \in \mathbb{R} - \{0\}$. Use the definition of R to show that there exists N so that $n \geq N$ implies $\sqrt[n]{|a_n|} \leq \dfrac{1}{|x|}$. Rearrange this into an inequality about $|a_n x^n|$, and apply the Comparison Test and Triangle Inequality for infinite series to finish the proof of the first part of the third bullet point above.

d) Suppose $0 < R < \infty$. Let $x \in (-R, R)$, and let $r = \frac{1}{2}\left(\frac{|x|}{R} + 1\right)$. Use the definition of R to show that there exists N so that $n \geq N$ implies $\sqrt[n]{|a_n|} \leq \dfrac{r}{|x|}$. Rearrange this into an inequality about $|a_n x^n|$, and then proceed similar to part (c).

e) Prove the second part of the third bullet point above (what you did in the previous parts is something of a prototype here).

Chapter 5

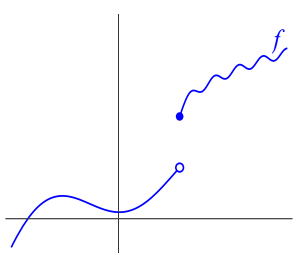
Continuity

5.1 Continuous functions

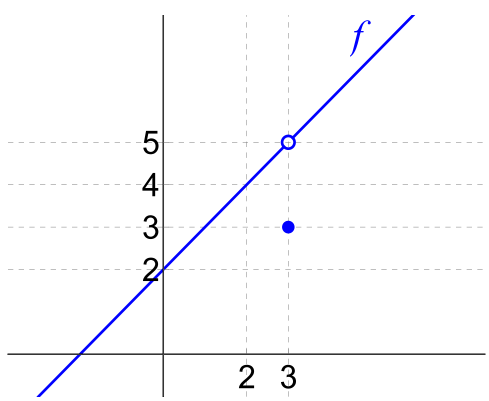
Definition 5.1 *A function $f : \mathbb{R} \to \mathbb{R}$ is called* **continuous (cts)** *if for every open set $U \subseteq \mathbb{R}$, the inverse image $f^{-1}(U)$ is also an open set.*

It is easy to use this definition to show that a function is **not** continuous.
All you need to do is find one open set U so that $f^{-1}(U)$ is not open.

EXAMPLE 1



$$f(x) = \begin{cases} x + 2 & x \neq 3 \\ 3 & x = 3 \end{cases}$$



EXAMPLE 2

Recall that the **Dirichlet function** is the indicator function of the rationals:

$$\mathbb{1}_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}.$$

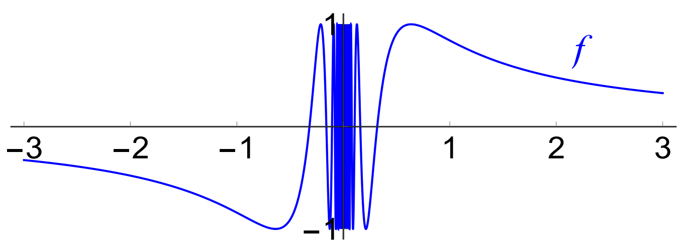
Determine, with proof, whether or not $\mathbb{1}_{\mathbb{Q}}$ is continuous.

EXAMPLE 3

Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} \sin \dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}.$$

Determine, with proof, whether or not f is continuous.



To show that a function *is* continuous, we often appeal to this theorem:

Theorem 5.2 *Let $f : \mathbb{R} \to \mathbb{R}$. If $\forall\, a, b \in \mathbb{R}$ with $a < b$, $f^{-1}(a, b)$ is an open set, then f is continuous.*

PROOF First, we claim that under the hypotheses of this theorem, $f^{-1}(a, \infty)$ is open. To show this, notice

$$(a, \infty) = \bigcup_{n=1}^{\infty} (a, a+n);$$

therefore

$$f^{-1}(a, \infty) = f^{-1}\left(\bigcup_{n=1}^{\infty}(a, a+n)\right) = \bigcup_{n=1}^{\infty} f^{-1}(a, a+n)$$

is the union of open sets, hence is open.

A similar argument (HW) shows that under the hypotheses of the theorem, $f^{-1}(-\infty, b)$ is open for any $b \in \mathbb{R}$.

Finally, let $U \subseteq \mathbb{R}$ be any open set. By ______________'s Theorem, we can write U as the disjoint union

$$U = \bigcup_j (a_j, b_j),$$

where it is possible that one of the a_j's is $-\infty$ and one of the b_j's is ∞. Now,

$$f^{-1}(U) = f^{-1}\left(\bigcup_j (a_j, b_j)\right) = \bigcup_j f^{-1}(a_j, b_j).$$

By hypothesis, this is the union of open sets, hence is open.

By definition, this makes f continuous. □

EXAMPLE 4

Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = x$. Prove that f is continuous.

EXAMPLE 5

Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = c$, where $c \in \mathbb{R}$ is a constant. Prove that f is continuous.

EXAMPLE 6

Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = x^3$. Prove that f is continuous.

Theorem 5.3 (Compositions of cts functions are cts) *Suppose $f, g : \mathbb{R} \to \mathbb{R}$. If f and g are continuous, then $f \circ g$ is continuous.*

PROOF Let $U \subseteq \mathbb{R}$ be open. Since f is continuous, we know

Since g is continuous, it follows that

Therefore

$$(f \circ g)^{-1}(U) = g^{-1}\left(f^{-1}(U)\right)$$

is open, making $f \circ g$ continuous by definition. □

5.2 Consequences of continuity

In this section, we study some important consequences of a function being continuous, related to topological ideas introduced in Chapter 3.

Preservation of compactness and existence of extrema

Lemma 5.4 *Let $f : A \to B$ be any function and let $E \subseteq A$. Then $f^{-1}(f(E)) \supseteq E$.*

PROOF This was HW from Chapter 1.

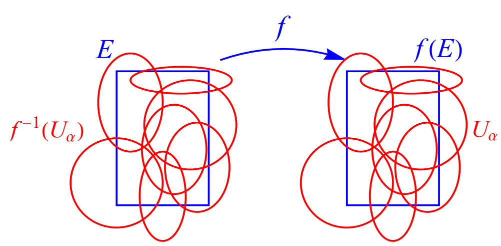
Theorem 5.5 (Preservation of compactness) *Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous. If $E \subseteq \mathbb{R}$ is compact, then $f(E)$ is also compact.*

PROOF Assume $E \subseteq \mathbb{R}$ is compact.

To prove $f(E)$ is compact, let ________ be an ________________ of $f(E)$.

Since f is continuous, each of the sets ________ are open, so $\{f^{-1}(U_\alpha)\}$ is an open cover of $f^{-1}(f(E))$.

Therefore $\{f^{-1}(U_\alpha)\}$ is also an open cover of E, since $f^{-1}(f(E)) \supseteq E$.



Since E is compact, there exists a ____________________ of $\{f^{-1}(U_\alpha)\}$, which we denote by

Claim:

Proof of claim: Let $x \in f(E)$.

That means ________________________________ .

Since $\{f^{-1}(U_j)\}_{j=1}^n$ covers E, $a \in f^{-1}(U_j)$ for some j.

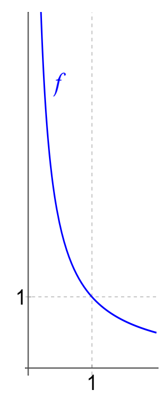
That means ________________ .

We've shown that every open cover of $f(E)$ has a finite subcover, so $f(E)$ is compact. □

Definition 5.6 *Let $f : \mathbb{R} \to \mathbb{R}$, and let $E \subseteq \mathbb{R}$.*
If $a \in E$ is such that $f(a) \geq f(x)$ for every $x \in E$, we say that $f(a)$ is the **absolute maximum value** *of f on E.*
If $a \in E$ is such that $f(a) \leq f(x)$ for every $x \in E$, we say that $f(a)$ is the **absolute minimum value** *of f on E.*

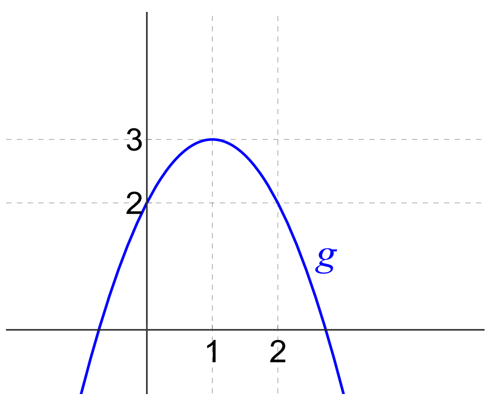
EXAMPLES

$f(x) = \dfrac{1}{x}$



Abs max of f on $(0, 1]$

$g(x) = 3 - (x - 1)^2$



Abs max of g on $\mathbb{R}$

Abs max of g on $(-\infty, 0]$

Abs max of g on $(1, 2)$

Theorem 5.7 (Max-Min Existence Theorem) *Let $E \subseteq \mathbb{R}$ be compact, and let $f : E \to \mathbb{R}$ be continuous.*
Then f has an absolute maximum value on E and an absolute minimum value on E.

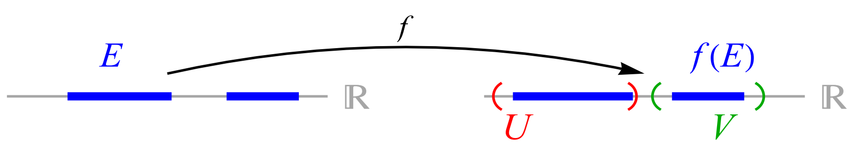
PROOF By preservation of compactness, $f(E)$ is compact.
Compact sets contain their maximum and minimum (Corollary 3.30). □

Preservation of connectedness and the IVT

Theorem 5.8 (Preservation of connectedness) *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous.*
If $E \subseteq \mathbb{R}$ is connected, then $f(E)$ is also connected.

PROOF Suppose not.
That means $f(E)$ has a disconnection, meaning a pair $\{U, V\}$ of sets such that

-
-
-
-



Claim: $\{f^{-1}(U), f^{-1}(V)\}$ is a disconnection of E.
Proof of claim:

- Since ______________________, $f^{-1}(U)$ and $f^{-1}(V)$ are both open.
- $f^{-1}(U)$ and $f^{-1}(V)$ are disjoint, since
$$f^{-1}(U) \cap f^{-1}(V) = f^{-1}(U \cap V) = f^{-1}(\varnothing) = \varnothing.$$
- $\{f^{-1}(U), f^{-1}(V)\}$ covers E, since
$$E \subseteq f^{-1}(f(E)) \subseteq f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V).$$
- $f^{-1}(U)$ hits E: let $y \in U \cap f(E)$; then $y = f(x)$ for $x \in f^{-1}(U) \cap E$.
The same logic shows $f^{-1}(V)$ hits E, proving the claim.

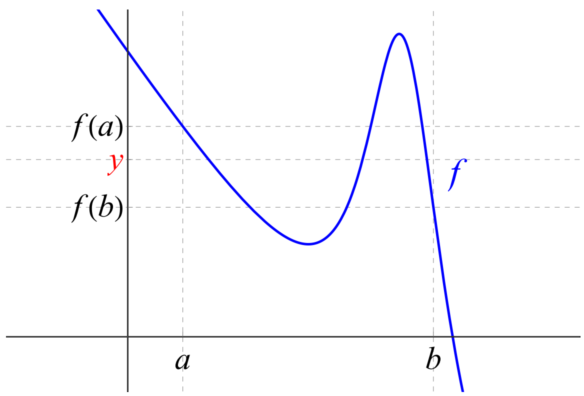
This contradicts E being connected.
So by contradiction, $f(E)$ must be connected. □

After five and a half chapters of work, we have finally done enough to prove one of the theorems we discussed at the very beginning of the course:

Theorem 5.9 (Intermediate Value Theorem (IVT)) *Let $f : [a,b] \to \mathbb{R}$ be continuous.*
For any y between $f(a)$ and $f(b)$, there exists $x \in (a,b)$ such that $f(x) = y$.

PROOF Let $E = [a,b]$.
E is connected, so by preservation of connectedness, $f(E)$ is also connected.

That makes $f(E)$ an ______________, so $f(E)$ has _________________ .
This property says that for any y between two numbers in $f(E)$ (such as $f(a)$ and $f(b)$), _____________ .
That means there is $x \in E$ such that $f(x) = y$. □

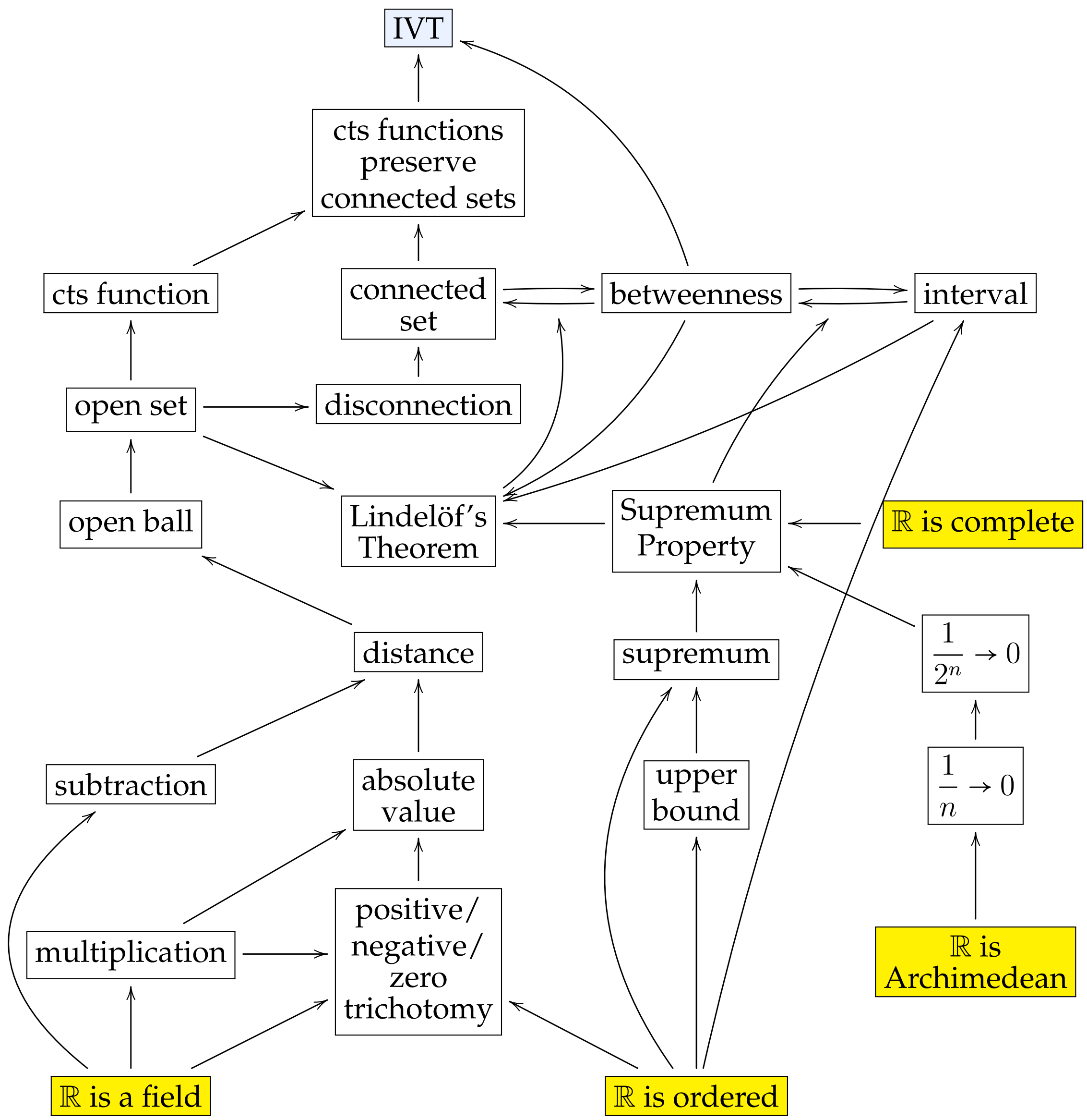


APPLICATION

Let $f(x) = x^5 + 2x^3 - x - 1$. Assuming that f is continuous (we'll prove it's cts later), prove that the equation $f(x) = 0$ has a solution between 0 and 1.

Ideas needed to prove the IVT

The IVT is a fairly easy result to understand (by MATH 430 standards), because of the associated picture. And its proof was pretty short (four sentences). However, if we look at what we had to prove to get to the IVT, we see this:



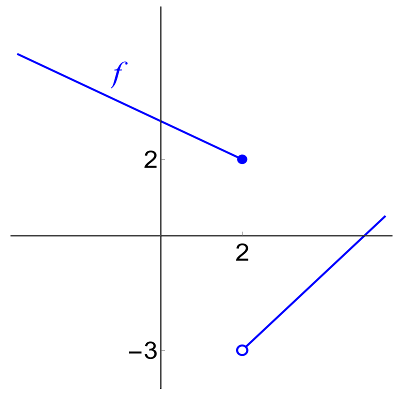
The point is that while the IVT looks simple, there's a lot going on behind the scenes. In particular, any proof of the IVT either directly or indirectly uses all the essential properties of $\mathbb{R}$.

5.3 Equivalent formulations of continuity

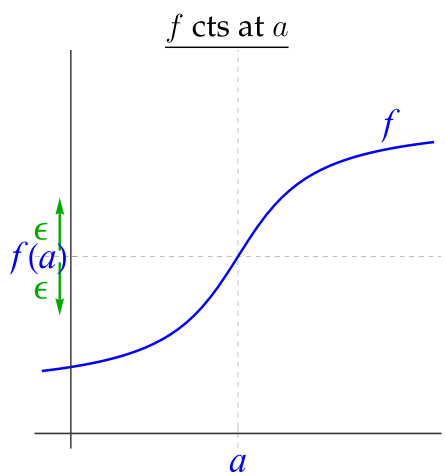
Continuity at a point

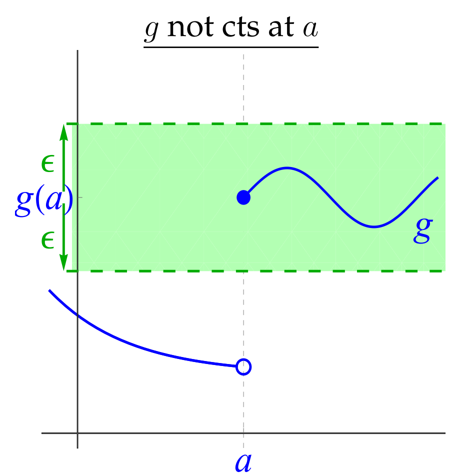
EXAMPLE 7

Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = \begin{cases} 3 - \frac{1}{2}x & \text{if } x \le 2 \\ x - 5 & \text{if } x > 2 \end{cases}$.



Definition 5.10 *Let $f : \mathbb{R} \to \mathbb{R}$, and let $a \in \mathbb{R}$. We say f is* **continuous at** a *if for every $\epsilon > 0$, there is $\delta > 0$ so that if $|x - a| < \delta$, then $|f(x) - f(a)| < \epsilon$.*
f is **continuous on** $E \subseteq \mathbb{R}$ *if, for every $a \in E$, f is continuous at a.*





EXAMPLE 8

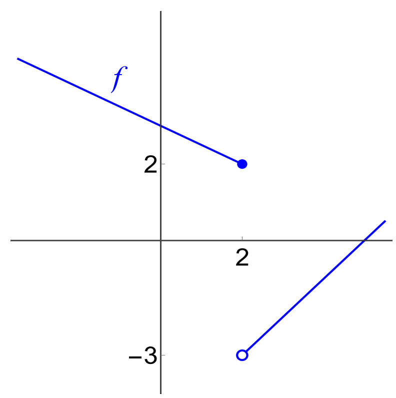
Prove that the function $f(x) = 5x + 2$ is continuous at $x = 3$.

Scratch work:

PROOF

EXAMPLE 9

Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = \begin{cases} 3 - \frac{1}{2}x & \text{if } x \leq 2 \\ x - 5 & \text{if } x > 2 \end{cases}$. Prove that f is not continuous at $x = 2$.



At this point, we have two (potentially competing) notions of continuity:

- continuity of a function (defined in terms of open sets), and
- continuity at each point (defined in terms of ϵs and δs).

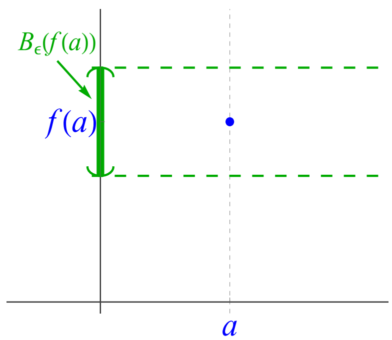
The next result reconciles these two ideas:

Theorem 5.11 *Let $f : \mathbb{R} \to \mathbb{R}$. f is continuous $\Leftrightarrow$ f is continuous at every $a \in \mathbb{R}$.*

PROOF ($\Rightarrow$) Suppose f is continuous (meaning that the inverse image of any open set is open).
Let $a \in \mathbb{R}$ and fix $\epsilon > 0$.
Since $B_\epsilon(f(a)) = (f(a) - \epsilon, f(a) + \epsilon)$ is open, its inverse image

$$U = f^{-1}(B_\epsilon(f(a)))$$

is an open set that contains a.



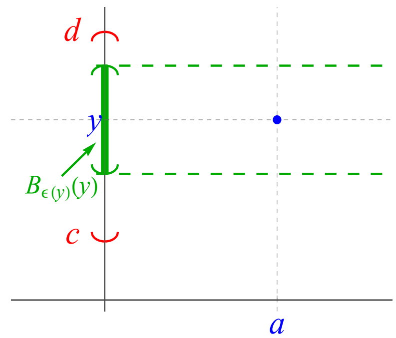
By definition of open set, there is $\delta > 0$ such that $B_\delta(a) \subseteq U$.
Now, for any $x \in \mathbb{R}$ such that $|x - a| < \delta$,

$$x \in U \Rightarrow f(x) \in B_\epsilon(f(a)) \Rightarrow |f(x) - f(a)| < \epsilon.$$

Thus f is continuous at a.

($\Leftarrow$) Suppose f is continuous at every real number.
Now consider an open interval $(c, d) \in \mathbb{R}$.
By openness, for each $y \in (c, d)$, there is $\epsilon(y) > 0$ such that $B_{\epsilon(y)}(y) \subseteq (c, d)$.
Furthermore, for every $a \in f^{-1}(y)$, f is cts at a, meaning $\exists\, \delta(a) > 0$ s.t.

$$\left.\begin{array}{c} |x - a| < \delta(a) \\ \text{(equivalently,} \\ x \in B_{\delta(a)}(a)) \end{array}\right\} \quad \text{implies} \quad \left\{\begin{array}{c} |f(x) - f(a)| = |f(x) - y| < \epsilon(y) \\ \text{(meaning } f(x) \in B_{\epsilon(y)}(y) \subseteq (c, d)) \end{array}\right.$$



We have just shown $x \in B_{\delta(a)}(a) \Rightarrow f(x) \subseteq (c,d)$, which in set language means

$$\bigcup_{a \in f^{-1}(c,d)} B_{\delta(a)}(a) \subseteq f^{-1}(c,d).$$

The reverse set inclusion also holds: if $x \in f^{-1}(c,d)$, then

$$x \in B_{\delta(x)}(x) \subseteq \bigcup_{a \in f^{-1}(c,d)} B_{\delta(a)}(a).$$

Therefore

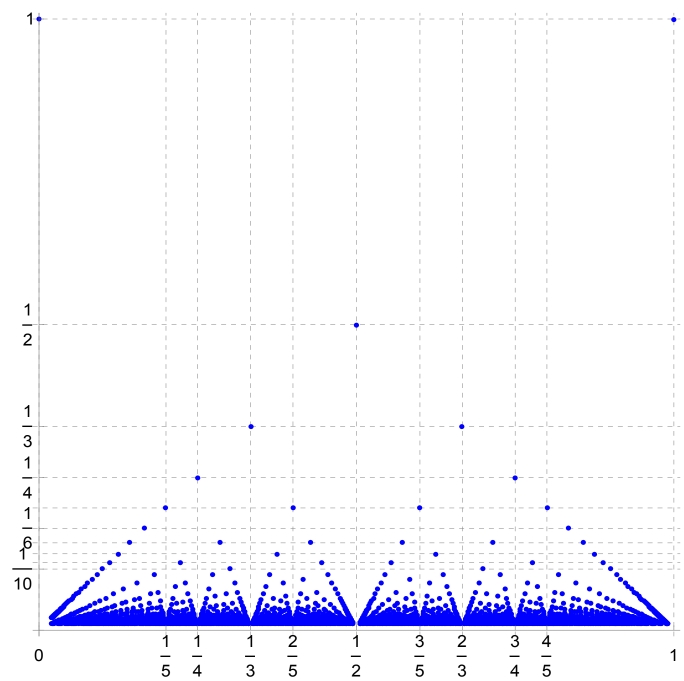
$$f^{-1}(c,d) = \bigcup_{a \in f^{-1}(c,d)} B_{\delta(a)}(a).$$

This means $f^{-1}(c,d)$ is a union of open balls, hence is an open set.
It follows that f is continuous. □

EXAMPLE 10

Let $\tau : \mathbb{R} \to \mathbb{R}$ be Thomae's function, defined by

$$\tau(x) = \begin{cases} \dfrac{1}{q} & x \in \mathbb{Q} \text{ and } x = \dfrac{p}{q} \text{ in lowest terms, with } q > 0 \\ 0 & x \notin \mathbb{Q} \end{cases}.$$



Determine the values a, if any, at which f is continuous.

Continuous functions preserve sequences

Theorem 5.12 (Preservation of convergent sequences) *Let $f : \mathbb{R} \to \mathbb{R}$. TFAE:*

1. *f is continuous at $a \in \mathbb{R}$.*
2. *For any sequence $\{x_n\}$ that converges to a, $f(x_n) \to f(a)$.*

PROOF (⇒) Suppose f is cts at x and suppose $x_n \to a$.
Our goal is to show $f(x_n) \to f(a)$.

Toward that end, ____________________ .

Since f is cts at a, there is ________ so that

Next, since $x_n \to a$, there is _____ so that

Thus, $\forall\, n \geq N$, we have $|x_n - a| < \delta$, which implies $|f(x_n) - f(a)| < \epsilon$.
This means $f(x_n) \to f(a)$.

(⇐) Suppose that for any sequence $\{x_n\}$ that converges to a, $f(x_n) \to f(a)$.
We prove this by contradiction: suppose f is not cts at a; therefore
$\exists\, \epsilon_0 > 0$ s.t. $\forall\, \delta > 0$, $\exists\, x$ with $|x - a| < \delta$ but $|f(x) - f(a)| \geq \epsilon_0$.
In particular, for every $n \in \mathbb{N}$, there is x_n with

$$|x_n - a| < \frac{1}{n}, \text{ but } \boxed{|f(x_n) - f(a)| \geq \epsilon_0}.$$

Since $|x_n - a| < \frac{1}{n}$, $x_n \to a$ by the ____________________________ .
By hypothesis, $f(x_n) \to f(a)$, so $\exists\, N$ s.t. for $n \geq N$, $\boxed{|f(x_n) - f(a)| < \epsilon_0}$.
The two boxed inequalities contradict one another.
Therefore f must be continuous at a. □

Corollary 5.13 *Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. For any convergent sequence $\{x_n\}$,*

$$f(\lim x_n) = \lim f(x_n).$$

EXAMPLE 11

Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = \begin{cases} 3 + 2x & \text{if } x \neq 1 \\ 4 & \text{if } x = 1 \end{cases}$. Prove that f is not continuous at $x = 1$.

Arithmetic with continuous functions

The preceding result (preservation of convergent sequences) gives us a nice way to show that constant multiples, sums, differences and products of continuous functions are continuous:

Corollary 5.14 *Suppose $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are continuous at $a \in \mathbb{R}$. Then:*

1. *For any constant $k \in \mathbb{R}$, kf is continuous at a;*
2. *$f + g$, $f - g$, and fg are continuous at a; and*
3. *if $g(a) \neq 0$, then $\dfrac{f}{g}$ is continuous at a;*

Furthermore, suppose $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are continuous. Then:

1. *For any constant $k \in \mathbb{R}$, kf is continuous;*
2. *$f + g$, $f - g$, and fg are continuous; and*
3. *if $g(x) \neq 0$ for all $x \in \mathbb{R}$, then $\dfrac{f}{g}$ is continuous.*

PROOF Let $\{x_n\}$ be any sequence that converges to a.

Since f and g are assumed continuous at a, $f(x_n) \to f(a)$ and $g(x_n) \to g(a)$.
From previous theorems about sequences (Ch. 2), we know

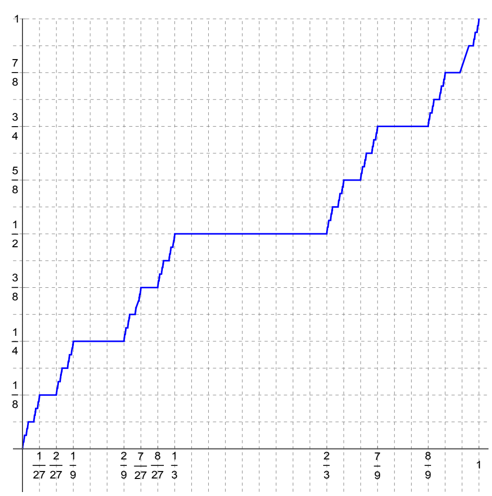
$$\begin{aligned}(kf)(x_n) &= k\,f(x_n) \to k\,f(a) = (kf)(a);\\ (f+g)(x_n) &= f(x_n) + g(x_n) \to f(a) + g(a) = (f+g)(a);\\ (f-g)(x_n) &= f(x_n) - g(x_n) \to f(a) - g(a) = (f-g)(a);\\ (fg)(x_n) &= f(x_n)g(x_n) \to f(a)g(a) = (fg)(a);\\ \left(\frac{f}{g}\right)(x_n) &= \frac{f(x_n)}{g(x_n)} \to \frac{f(a)}{g(a)} = \left(\frac{f}{g}\right)(a) \text{ so long as } g(a) \neq 0.\end{aligned}$$

So all these functions preserve convergent sequences, so they are all cts at a by Theorem 5.12.

Now for the second part of the corollary.
If f and g are cts, then they are cts at every $a \in \mathbb{R}$ by Theorem 5.11.
By the first part of this corollary, kf, $f+g$, $f-g$ and fg are cts $\forall\, a \in R$, and $\dfrac{f}{g}$ is continuous for all $a \in \mathbb{R}$ so long as $g(a) \neq 0$ for all a.
All of these functions are therefore continuous by Theorem 5.11. □

Continuity of monotone surjections

Recall the Cantor function, discussed in Chapter 3, whose graph is as follows:



QUESTION

At what points $a \in [0,1]$ is the Cantor function continuous?

We'll answer this by proving a theorem that applies to not just the Cantor function, but any function that is monotone and surjective.

Theorem 5.15 *Let $f : \mathbb{R} \to \mathbb{R}$. If f is monotone and surjective, then f is continuous.*

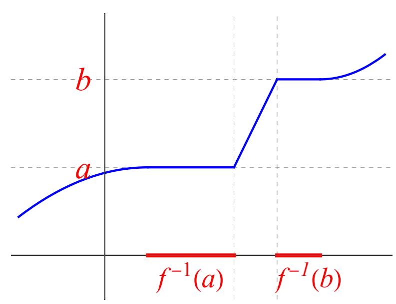
PROOF Suppose for now that f is increasing and surjective.
(We'll handle the situation when f is decreasing later.)
Let $a, b \in \mathbb{R}$ with $a < b$.

Claim: $f^{-1}(a,b) = (\sup f^{-1}(a), \inf f^{-1}(b))$.

If we prove this claim, then the inverse image of any open interval is open, meaning f is continuous, as wanted.

Proof of claim: This is a set equality argument:

($\subseteq$) Let $x \in f^{-1}(a,b)$. That means $f(x) \in (a,b)$, so $a < f(x) < b$.
Now let $w \in f^{-1}(a)$. If $w \geq x$, then $a = f(w) \geq f(x)$, a contradiction to f being increasing. Therefore $w < x$.
This makes x an upper bound of $f^{-1}(a)$, so $x \geq \sup f^{-1}(a)$.



To show $x \neq \sup f^{-1}(a)$, let $y = \frac{1}{2}(a + f(x))$. Note $a < y < f(x)$.
Since f is surjective, there is $c \in f^{-1}(y)$.
c is also an upper bound of $f^{-1}(a)$ (for the same reason x is).
At the same time, $c < x$ since $f(c) = y < f(x)$ and f is increasing.
So x is not the least upper bound of $f^{-1}(a)$, i.e. $x > \sup f^{-1}(a)$.

A similar argument shows $x < \inf f^{-1}(b)$ (this is left as HW).

($\supseteq$) Let $x \in (\sup\{f^{-1}(a)\}, \inf\{f^{-1}(b)\})$.
Then, for any $y \in f^{-1}(a)$ and $z \in f^{-1}(b)$, $y < x < z$, so since f is increasing,

$$a = f(y) \leq f(x) \leq f(z) = b.$$

If $a = f(x)$, then $x \in f^{-1}(a)$.
This means $\sup\{f^{-1}(a)\} \geq x$, a contradiction.
Similarly, if $f(x) = b$, then $x \in f^{-1}(b)$, so $\inf\{f^{-1}(b)\} \leq x$, also impossible.
Therefore $a < f(x) < b$, i.e. $x \in f^{-1}(a,b)$, as wanted.

This proves the claim, which shows f is continuous.

Finally, if f is decreasing and surjective, then $-f$ is increasing and surjective, hence $-f$ is continuous by the first part of this proof. That means $f = -(-f)$ is also continuous. □

Corollary 5.16 *The Cantor function $c : [0,1] \to [0,1]$ is continuous.*

PROOF We proved in Chapter 3 that c is surjective and increasing. □

5.4 Limits of functions

In Calculus 1, you learn about the concept of *limit* of a function. The concept of limit is the building block of the rest of the subject; indeed, **calculus is, to a large extent, the study of limits**.

However, you are often told what a limit is in a vague, imprecise way. The reason is that the actual definition of limit is technical and requires some understanding of advanced material. Of course, **YOU** are now an advanced student, so you can handle the legitimate, mathematically rigorous definition of limit:

Definition 5.17 *Let $f : \mathbb{R} \to \mathbb{R}$ and let $a \in \mathbb{R}$. We say a real number L is* **a limit of f as x approaches a**, *and write*

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

if given any $\epsilon > 0$, there is $\delta > 0$ such that

$$0 < |x - a| < \delta \text{ implies } |f(x) - L| < \epsilon.$$

REMARKS

1. In order for this definition to make sense, we don't need f to be defined everywhere, and we don't actually need f to be defined at a.

   The minimum requirement is that the domain of f includes all points in some open interval containing a, except for perhaps a itself.

2. By definition, nothing about what happens with f when $x = a$ has anything to do with whether or not $\lim_{x \to a} f(x)$ exists, or what its value is.

So, for instance, for any two functions $f, g : \mathbb{R} \to \mathbb{R}$, we immediately know (just based on the definition) that we can "factor and cancel" without changing the value of the limit, i.e. write things like

$$\lim_{x \to a} \frac{f(x)(x-a)}{g(x)(x-a)} = \lim_{x \to a} \frac{f(x)}{g(x)}$$

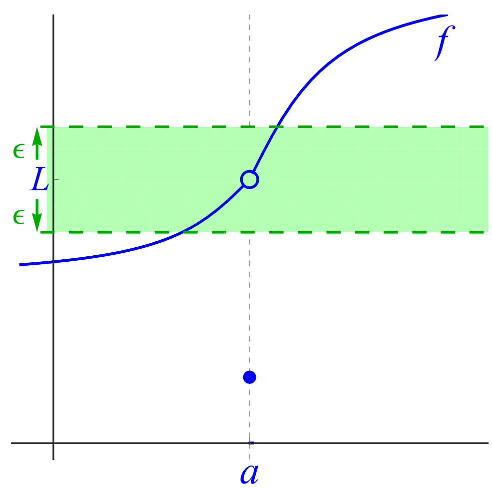
based solely on the definition of limit, since the red and blue fractions are identical except when $x = a$.

3. In this definition, a and L must be real numbers. In Calculus 1, you also learn about "infinite limits" and "limits at infinity", like

$$\lim_{x \to \pm\infty} f(x) = L \quad \text{and} \lim_{x \to a} f(x) = \pm\infty.$$

We (probably) won't discuss these. They have their own separate definitions with ϵ and δ (or N) in them.

A picture to explain the ϵ, δ definition of limit



Limits of functions and limits of sequences

When you are taught limits (or maybe when you teach limits someday) in Calculus 1, you are taught the informal nonsense that

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

means

"as x gets closer and closer to a, $f(x)$ gets closer and closer to L".

What does this "closer and closer" mean? Well, nothing really (which is why this idea is imprecise), but it sort of has something to do with convergence of sequences:

Theorem 5.18 (Limits preserve convergent sequences) *Let $f : \mathbb{R} \to \mathbb{R}$, and let $a \in \mathbb{R}$. TFAE:*

1. $\lim_{x\to a} f(x) = L$.
2. *Given any sequence $\{x_n\}$ of real numbers such that*
   a) $x_n \neq a$ *for every n, and*
   b) $x_n \to a$,

   it follows that $f(x_n) \to L$.

NOTE: Earlier we had a result that showed that *continuous* functions preserved convergent sequences when you evaluate the function at the limit of the sequence:

$$x_n \to a;\ f \text{ continuous} \quad \Rightarrow \quad f(x_n) \to f(a).$$

This theorem says that given *any* function (not necessarily a continuous one), non-constant convergent sequences are preserved when you evaluate the **limit of the function** at the limit of the sequence:

$$x_n \to a \text{ (where } x_n \neq a);\ f \text{ any function} \quad \Rightarrow \quad f(x_n) \to \lim_{x\to a} f(a).$$

This suggests that when f is continuous,

We'll prove that a little later.

PROOF $(1 \Rightarrow 2)$ Assume $\lim_{x \to a} f(x) = L$.

Let $\{x_n\}$ be any sequence with $x_n \neq a$ for all n, where $x_n \to a$.
To prove $f(x_n) \to L$, let $\epsilon > 0$.
Since $\lim_{x \to a} f(x) = L$, $\exists\, \delta > 0$ s.t.

$$\tag{5.1}$$

Since $x_n \to a$, $\exists\, N > 0$ s.t.

$$\tag{5.2}$$

So for $n \geq N$, we have $|x_n - a| < \delta$ by line (5.2) above.
Since $x_n \neq a$ for all n, we have $|f(x_n) - L| < \epsilon$ from line (5.1) above.
By definition, $f(x_n) \to L$.

$(2 \Rightarrow 1)$ Assume that for any sequence $\{x_n\}$ with $x_n \neq a$ for all n such that $x_n \to a$, $f(x_n) \to L$.

To prove $\lim_{x \to a} f(x) = L$, suppose not.
That means $\exists\, \epsilon_0 > 0$ such that $\forall\, \delta > 0$, there is x with

$$0 < |x - a| < \delta \text{ but } |f(x) - f(a)| \geq \epsilon_0.$$

In particular, for every $n \in \mathbb{N}$, there is $x_n \in \mathbb{R}$ with

$$0 < |x_n - a| < \frac{1}{n} \text{ but } |f(x_n) - L| \geq \epsilon_0.$$

This produces a sequence $\{x_n\}$; since $0 < |x_n - a|\ \forall\, n$, $x_n \neq a\ \forall\, n$.
Also, since $|x_n - a| < \frac{1}{n}$, it follows that $x_n \to a$.
But $|f(x_n) - L| \geq \epsilon_0$ for all n, so $f(x_n) \not\to L$, contradicting the hypothesis.
By contradiction, the result is true. □

EXAMPLE 12

Prove that $\lim_{x \to 0} \frac{1}{x}$ DNE.

One reason Theorem 5.18 is important is because it enables us to translate all the facts we proved earlier about convergence of sequences over to the setting of limits of functions:

Theorem 5.19 (Uniqueness of limit of a function) *Let $f : \mathbb{R} \to \mathbb{R}$ and let $a \in \mathbb{R}$. If $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} f(x) = M$, then $L = M$.*

PROOF Let $\{x_n\}$ be an arbitrary sequence with $x_n \neq a$ for all n, so that $x_n \to a$.
By Theorem 5.18, $f(x_n) \to L$ and $f(x_n) \to M$.
But limits of sequences are unique (Chapter 2), so $L = M$. □

Theorem 5.20 *Let $f : \mathbb{R} \to \mathbb{R}$. If $\lim_{x \to a} f(x)$ exists, then there is a $\delta > 0$ so that the set*

$$\{f(x) : x \in (a - \delta, a + \delta)\}$$

is bounded.

PROOF Suppose not, i.e. that for every $\delta > 0$, the set

$$\{f(x) : x \in (a - \delta, a + \delta)\}$$

is unbounded. That means that for every $n \in \mathbb{N}$, there is

$$x_n \in \left(a - \frac{1}{n}, a + \frac{1}{n}\right)$$

such that $f(x_n) > n$.
This produces a sequence $\{x_n\}$ with $|x_n - a| < \frac{1}{n}$, meaning $x_n \to a$.
So by Theorem 5.18, $f(x_n)$ converges to $\lim_{x \to a} f(x)$.
However, since $f(x_n) > n$, $\{f(x_n)\}$ is unbounded, hence diverges.
This is a contradiction. □

Theorem 5.21 (Main Limit Theorem (for functions)) *Let $f, g : \mathbb{R} \to \mathbb{R}$ be functions and let $a \in \mathbb{R}$ be such that*

$$\lim_{x \to a} f(x) = L \text{ and } \lim_{x \to a} g(x) = M.$$

Then:

1. *for any constant $c \in \mathbb{R}$, $\lim_{x \to a} [cf(x)] = cL$;*
2. $\lim_{x \to a} [f(x) + g(x)] = L + M$;
3. $\lim_{x \to a} [f(x) - g(x)] = L - M$;
4. $\lim_{x \to a} [f(x)g(x)] = LM$; *and*
5. *if $M \neq 0$, then $\lim_{x \to a} \dfrac{f(x)}{g(x)} = \dfrac{L}{M}$.*

PROOF Let $\{x_n\}$ be a sequence with $x_n \neq a$ for all n, such that $x_n \to a$.

By Theorem 5.18, $f(x_n) \to L$ and $g(x_n) \to M$.

By the Main Limit Theorem for sequences,

$$\begin{aligned}
(cf)(x_n) &= cf(x_n) \to cL; \\
(f+g)(x_n) &= f(x_n) + g(x_n) \to L + M; \\
(f-g)(x_n) &= f(x_n) - g(x_n) \to L - M; \\
(fg)(x_n) &= f(x_n)g(x_n) \to LM;
\end{aligned}$$

and if $M \neq 0$,

$$\left(\frac{f}{g}\right)(x_n) = \frac{f(x_n)}{g(x_n)} \to \frac{L}{M}.$$

Applying Theorem 5.18 again yields the desired facts. □

EXAMPLE 13

Prove that $\lim_{x \to 3} \frac{x^2 - 9}{x^2 + 4x - 21}$ exists, and find its value.

Solution of a Calculus 1 student:

A more rigorous version of that argument:

A direct argument, using the definition of limit:

Theorem 5.22 (Limits preserve soft inequalities) *Let $f, g : \mathbb{R} \to \mathbb{R}$ be functions and let $a \in \mathbb{R}$.*
Assume $f(x) \leq g(x)$ for all $x \in \mathbb{R} - \{a\}$.
Then, if these limits exist, we have

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

PROOF Let $L = \lim_{x\to a} f(x)$ and let $M = \lim_{x\to a} g(x)$.

Next, let $\{x_n\}$ be a sequence with $x_n \neq a$ for all n, such that $x_n \to a$.

By Theorem 5.18, $f(x_n) \to L$ and $g(x_n) \to M$.

But since $f(x_n) \leq g(x_n)$, it follows that $L \leq M$, since limits of sequences preserve soft inequalities. □

Theorem 5.23 (Squeeze Theorem (for functions)) *Let $f, g, h : \mathbb{R} \to \mathbb{R}$ be functions and let $a \in \mathbb{R}$. Assume $f(x) \leq g(x) \leq h(x)$ for all $x \in \mathbb{R} - \{a\}$. If*

$$\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = L,$$

then $\lim_{x\to a} g(x) = L$ as well.

PROOF let $\{x_n\}$ be a sequence with $x_n \neq a$ for all n, such that $x_n \to a$.

By Theorem 5.18, $f(x_n) \to L$ and $h(x_n) \to L$.

But since $f(x_n) \leq g(x_n) \leq h(x_n)$, $g(x_n) \to L$ by the Squeeze Theorem for sequences.

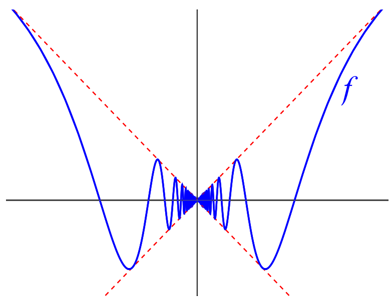
The result follows from Theorem 5.18. □

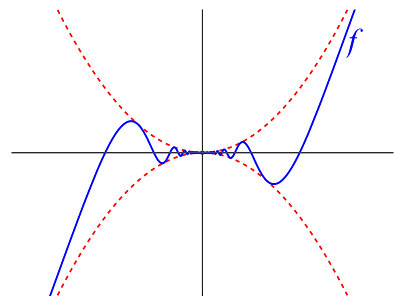
EXAMPLE 14

Let $m > 0$ and let $n \in \{1, 2, 3, \ldots\}$. Set

$$f(x) = \begin{cases} x^m \sin\left(\dfrac{1}{x^n}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}.$$

Determine if $\lim_{x \to 0} f(x)$ exists; if so, find its value.





Last, we'll verify that the way the concept of continuity is often presented in Calc 1 is valid:

Theorem 5.24 *Let* $f : \mathbb{R} \to \mathbb{R}$ *and let* $a \in \mathbb{R}$*. TFAE:*

1. f *is continuous at* a*.*
2. $\lim_{x \to a} f(x) = f(a)$*.*

PROOF (⇒) Suppose f is continuous at a.

To prove the limit statement, let $\epsilon > 0$.
By the definition of continuity at a, there is $\delta > 0$ such that

$$|x - a| < \delta \;\Rightarrow\; |f(x) - f(a)| < \epsilon.$$

Clearly, for this δ,

$$0 < |x - a| < \delta \;\Rightarrow\; |f(x) - f(a)| < \epsilon,$$

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

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

To prove f is continuous at a, let $\epsilon > 0$.
By the definition of limit, there is $\delta > 0$ such that

$$0 < |x - a| < \delta \;\Rightarrow\; |f(x) - f(a)| < \epsilon.$$

Now, if $|x - a| < \delta$, either

$$\begin{array}{cl} x = a, & \text{meaning } |f(x) - f(a)| = |f(a) - f(a)| = 0 < \epsilon \\ \text{or} & \\ 0 < |x - a| < \delta, & \text{meaning } |f(x) - f(a)| < \epsilon. \end{array}$$

Either way, we have

$$|x - a| < \delta \;\Rightarrow\; |f(x) - f(a)| < \epsilon,$$

meaning f is continuous at a as wanted. □

5.5 Sequences of continuous functions

The perils of interchanging limits

QUESTION

Is the limit of a sequence of continuous functions necessarily continuous?
More precisely, if $f_n : \mathbb{R} \to \mathbb{R}$ is cts for all n and $f_n \to f$, is f cts?

Let's dive into the question posed above using the characterization of continuous functions as those that preserve convergent sequences.

$$\begin{aligned} f \text{ continuous } &\Leftrightarrow \lim_{m\to\infty} f(x_m) = f(\lim_{m\to\infty} x_m) \text{ for any convergent sequence } \{x_m\} \\ &\Leftrightarrow \lim_{m\to\infty}\left[\lim_{n\to\infty} f_n(x_m)\right] = \lim_{n\to\infty} f_n\left(\lim_{m\to\infty} x_m\right). \\ &\qquad\qquad\qquad\qquad\qquad\qquad (\text{since } f_n \text{ is continuous}) \\ &\Leftrightarrow \lim_{m\to\infty}\left[\lim_{n\to\infty} f_n(x_m)\right] = \lim_{n\to\infty}\left[\lim_{m\to\infty} f_n(x_m)\right]. \end{aligned}$$

Notice that the question of whether or not the limit f is continuous boils down to a question about **iterated limits** (one limit inside another).
What we care about is whether one can *interchange limits*, i.e. do the limits $m \to \infty$ and $n \to \infty$ in either order and get the same answer.
Unfortunately, in general **iterated limits cannot be interchanged legally**:

EXAMPLE 14

Evaluate these iterated limits:

$$\lim_{x\to 0}\left[\lim_{y\to 0}\frac{x+y}{x-y}\right] \qquad\qquad \lim_{y\to 0}\left[\lim_{x\to 0}\frac{x+y}{x-y}\right]$$

Uniform limits of continuous functions

REVISED QUESTION

Is the **uniform** limit of a sequence of continuous functions necessarily cts?
More precisely, if $f_n : \mathbb{R} \to \mathbb{R}$ is cts for all n and $f_n \rightrightarrows f$, is f cts?
(In our example on the previous page, $f_n \not\rightrightarrows f$.)

Theorem 5.25 *Suppose $\{f_n\}$ is a sequence of continuous functions from $E \subseteq \mathbb{R}$ to $\mathbb{R}$. If $f_n \rightrightarrows f$ on E, then f is continuous on E.*

PROOF Suppose $f_n \rightrightarrows f$ on E.
Let $a \in E$; our goal is to show f is cts at $a \in E$.
To do this, let $\epsilon > 0$.
Since $f_n \rightrightarrows f$, $\exists\, N$ s.t. $n \geq N$ implies $|f_n(x) - f(x)| < \dfrac{\epsilon}{3}$ for all $x \in E$.
Since f_N is cts (at a), $\exists\, \delta > 0$ s.t. $|x - a| < \delta$ implies $|f_N(x) - f_N(a)| < \dfrac{\epsilon}{3}$.
Now

$$\begin{aligned}|f(x) - f(a)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(a)| + |f_N(a) - f(a)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &= \epsilon.\end{aligned}$$

This shows f is continuous at a.
Since $a \in E$ is arbitrary, f is continuous on E as wanted. □

Corollary 5.26 $\exp$, $\sin$ *and* $\cos$ *are continuous functions.*

PROOF Recall Theorem 4.24, which says that the power series defining $\exp$, $\sin$ and $\cos$ converge uniformly on compact subsets of $\mathbb{R}$.
That means, applying our latest result, that these functions are continuous on any compact subset of $\mathbb{R}$.
But every $x \in \mathbb{R}$ is contained in a compact subset $[x-1, x+1]$, so these functions must be continuous at every x. □

5.6 Chapter 5 Summary

DEFINITIONS TO KNOW

Nouns

- $f : \mathbb{R} \to \mathbb{R}$ has **absolute maximum value** $f(c)$ on $E \subseteq \mathbb{R}$ if $c \in E$ is so that $f(x) \leq f(c)$ for all $x \in E$.
  $f : \mathbb{R} \to \mathbb{R}$ has **absolute minimum value** $f(c)$ on $E \subseteq \mathbb{R}$ if $c \in E$ is so that $f(x) \geq f(c)$ for all $x \in E$.
- We say L is the **limit** of $f : \mathbb{R} \to \mathbb{R}$ as x approaches a, and write $\lim_{x \to a} f(x) = L$, if $\forall\, \epsilon > 0\ \exists\, \delta > 0$ s.t. $0 < |x - a| < \delta$ implies $|f(x) - L| < \epsilon$.

Adjectives that describe functions $f : \mathbb{R} \to \mathbb{R}$

- f is **continuous** if for every open $U \subseteq \mathbb{R}$, $f^{-1}(U)$ is also open.
- f is **continuous at** a if $\forall\, \epsilon > 0\ \exists\, \delta > 0$ s.t. $|x - a| < \delta$ implies $|f(x) - f(a)| < \epsilon$.

THEOREMS WITH NAMES

Preservation of compactness If f is continuous and E is compact, then $f(E)$ is compact.

Max-Min Existence Theorem If f is continuous and E is compact, then f achieves an absolute maximum value and absolute minimum value on E.

Preservation of connectedness If f is continuous and E is connected, then $f(E)$ is connected.

Intermediate Value Theorem (IVT) If f is continuous and $a < b$, then for every y between $f(a)$ and $f(b)$, there is $x \in (a, b)$ so that $f(x) = y$.

Continuous functions preserve convergent sequences f is continuous at a if and only if for every sequence $x_n \to a$, then $f(x_n) \to f(a)$.

In other words, f is cts $\Leftrightarrow \lim f(x_n) = f(\lim x_n)$ for every convergent sequence $\{x_n\}$.

Limits preserve convergent sequences $\lim_{x \to a} f(x) = L$ if and only if for any nonconstant $\{x_n\}$ with $x_n \to a$, $f(x_n) \to L$.

Main Limit Theorem Limits of functions are preserved under arithmetic.

Squeeze Theorem If $f(x) \leq g(x) \leq h(x)$ and $\lim_{x \to a} f(x) = \lim_{x \to a} h(x) = L$, then $\lim_{x \to a} g(x) = L$.

Other theorems to remember

- Sums, differences, products, quotients and compositions of continuous functions are continuous.
- $f : \mathbb{R} \to \mathbb{R}$ is continuous $\Leftrightarrow$ f is continuous at every $a \in \mathbb{R}$.
- Monotone surjections are continuous.
- Limits of functions preserve soft inequalities.
- $f : \mathbb{R} \to \mathbb{R}$ is continuous at a if and only if $\lim_{x \to a} f(x) = f(a)$.
- (★) The uniform limit of a sequence of continuous functions is continuous.

Facts about specific functions

- Dirichlet's function $\mathbb{1}_{\mathbb{Q}}$ is not continuous at any $x \in \mathbb{R}$.
- $f(x) = \begin{cases} \sin \frac{1}{x^n} & x \neq 0 \\ 0 & x = 0 \end{cases}$ is discontinuous at $x = 0$ for any n.
- $f(x) = \begin{cases} x^m \sin \frac{1}{x^n} & x \neq 0 \\ 0 & x = 0 \end{cases}$ is continuous if $m \geq 1$.
- Thomae's function τ is continuous at the irrationals, but discontinuous at the rationals.
- The Cantor function c is continuous.
- Constant functions are continuous.
- The identity function $f(x) = x$ is continuous.
- $|x|$ is continuous (HW).
- Polynomials are continuous.
- Rational functions are continuous, except at values which make their denominators zero.
- Root functions like $\sqrt{x}$ and $\sqrt[3]{x}$ are continuous (HW).
- (★) $\exp$, $\sin$ and $\cos$ are continuous.

PROOF TECHNIQUES

To prove that f is continuous, do one of these things:

1. Show that for any $a < b$, $f^{-1}(a, b)$ is open.
2. Show f is a sum/difference/product/composition of functions known to be continuous.
3. Show f is the quotient of functions known to be continuous, where the denominator is never zero.
4. Show f is a monotone surjection.
5. (★) Show f is a uniform limit of a sequence of functions known to be continuous.
6. Show f is continuous at every $a \in \mathbb{R}$ (see below).

To prove that f is continuous at a, do one of these things:

1. Show f is continuous (see above).
2. Show f is a sum/difference/product/composition of functions known to be continuous at a.
3. Show f is the quotient of functions known to be continuous at a, where the denominator is not zero at a.
4. Take an arbitrary sequence $x_n \to a$ and prove $f(x_n) \to f(a)$.
5. Use the definition: let $\epsilon > 0$; from scratch work come up with $\delta > 0$ so that $|x - a| < \delta$ implies $|f(x) - f(a)| < \epsilon$.
6. Show $\lim_{x \to a} f(x) = f(a)$ (see below).

To prove that f is not continuous, do one of these things:

1. Find a single open set U so that $f^{-1}(U)$ is not open.
2. Find a sequence $x_n \to x$ so that $f(x_n) \not\to f(x)$.
3. Show f is not continuous at some specific a (see below).

To prove that f is not continuous at a, do one of these things:

1. Find a sequence $x_n \to a$ so that $f(x_n) \not\to f(a)$.
2. Use the definition: find one $\epsilon_0 > 0$ so that $\forall\, \delta > 0$ there is x with $|x - a| < \delta$ implies $|f(x) - f(a)| \geq \epsilon_0$.
3. Show f is not bounded in any open interval containing a.

To prove $\lim_{x \to a} f(x) = L$, do one of these things:

1. Prove f is continuous, in which case $L = f(a)$.
2. Split the limit up using the Main Limit Theorem, factoring and canceling if necessary so that what remains is continuous.
3. Take an arbitrary nonconstant sequence $x_n \to a$ and prove $\lim f(x_n) = L$ using a method of Chapter 2.
4. Use the Squeeze Theorem.
5. Use the definition: let $\epsilon > 0$; from scratch work figure out $\delta > 0$ so that $0 < |x-a| < \delta$ implies $|f(x) - L| < \epsilon$.

To prove $\lim_{x\to a} f(x)$ **does not exist**, do one of these things:

1. Show f is unbounded in any open interval containing a.
2. Find a sequence $x_n \to a$ so that $\{f(x_n)\}$ diverges.
3. Find two sequences $x_n \to a$ and $y_n \to a$ so that $\lim x_n \neq \lim y_n$.
4. Use the definition: show that for any $L \in \mathbb{R}$, find one $\epsilon_0 > 0$ so that $\forall\, \delta > 0$ there is x with $0 < |x-a| < \delta$ but $|f(x) - L| > \epsilon_0$.

5.7 Chapter 5 Homework

Exercises from Section 5.1

1. Finish the proof of Theorem 5.2 by showing that if for all $a < b$, $f^{-1}(a,b)$ is open, then $f^{-1}(-\infty, b)$ is open for any $b \in \mathbb{R}$.

2. Prove, using the open set definition of continuity, that $f(x) = |x|$ is continuous.

3. Prove, using the open set definition of continuity, that $f(x) = x^2$ is continuous.

4. Prove, using the open set definition of continuity, that for any $n \in \mathbb{N}$, $f(x) = \sqrt[n]{x}$ is continuous.

   Hint: There are two cases depending on whether or not n is even or odd.

5. Prove, using the open set definition of continuity, that $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \begin{cases} 3 - x & x < 2 \\ 2x - 3 & x \geq 2 \end{cases}$ is continuous.

6. Prove, using the open set definition of continuity, that $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \begin{cases} 3 - x & x < 2 \\ 2x - 1 & x \geq 2 \end{cases}$ is not continuous.

Exercises from Section 5.2

7. Let $I = [a, b] \subseteq \mathbb{R}$ and let $f : I \to I$ be continuous. Prove that f must have at least one fixed point (a point $p \in X$ is a **fixed point** of a function $f : X \to X$ if $f(p) = p$).

   Hint: Apply the Intermediate Value Theorem to some appropriately chosen function.

8. Let $f : [0, 1] \to \mathbb{R}$ be continuous such that $f(0) = f(1)$. Prove that there exists $c \in \left[0, \frac{1}{2}\right]$ such that $f(c) = f\left(c + \frac{1}{2}\right)$. Explain why this result implies that at any time, there are two antipodal points on the earth's Equator that have the same temperature.

9. Prove or disprove: there is a non-constant, continuous function $f : \mathbb{R} \to \mathbb{R}$ such that $f(\mathbb{R}) \subseteq \mathbb{Q}$.

10. Prove that the equation $2x^4 - 11x^3 + 7x^2 - 15 = 0$ has at least two real solutions.

11. Let $f : \mathbb{R} \to \mathbb{R}$ be a polynomial of odd degree, i.e. $f(x) = a_0 + a_1 x + a_2 x + \ldots + a_n x^n$ where n is odd and $a_0, a_1, \ldots, a_n \in \mathbb{R}$, with $a_n \neq 0$.

    a) Suppose $a_n > 0$. Prove that $\exists\, b \in \mathbb{R}$ s.t. $f(b) > 0$.

       Hint: Factor out x^n from the formula for f to get x^n times the sum of a bunch of fractions. Choose $x = b$ where b is large enough so that the fractions add up to something sufficiently small.

    b) Suppose $a_n > 0$. Prove that $\exists\, a \in \mathbb{R}$ s.t. $f(a) < 0$.

       Hint: Using the factorization from part (a), now choose $x = a$ where a is sufficiently negative so that the fractions add up to something sufficiently small.

    c) Suppose $a_n > 0$. Prove that f has a root (i.e. that $\exists\, x \in \mathbb{R}$ s.t. $f(x) = 0$).

    d) Suppose $a_n < 0$. Prove that f has a root.

12. Give an example of a continuous function $f : \mathbb{R} \to \mathbb{R}$ and an open set E so that $f(E)$ is <u>not</u> open.

Exercises from Section 5.3

13. Finish the proof of Theorem 5.15 (monotone surjections are continuous) by showing, in the context of the proof of that theorem, that $x < \inf f^{-1}(b)$.

14. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous, monotone function and let $E \subseteq \mathbb{R}$ be bounded. Prove $f(\sup E) = \sup f(E)$.

15. Use the ϵ,δ-definition of continuity at a to prove that every linear function from $\mathbb{R}$ to $\mathbb{R}$ is continuous at every $a \in \mathbb{R}$.

16. Determine, with proof, whether or not the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \begin{cases} x & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$ is continuous at 0.

17. Consider the function introduced in Exercise 6, namely $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = \begin{cases} 3 - x & x < 2 \\ 2x - 1 & x \geq 2 \end{cases}$

    a) Prove f is not continuous at 2 using the ϵ,δ-definition.
    b) Prove f is not continuous at 2 by finding two sequences $\{x_n\}$ and $\{y_n\}$, both of which converge to 2, so that $\{f(x_n)\}$ and $\{f(y_n)\}$ do not have the same limit.

18. Prove or disprove: there exists a function $f : [0, 1] \to \mathbb{R}$ which is not continuous at any point, but for which the function $|f|$ is continuous on $[0, 1]$.

19. Give an example (with proof) of two functions $f, g : \mathbb{R} \to \mathbb{R}$ which are both discontinuous at $x = a$, but for which $f + g$ is continuous at a.

20. Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be continuous functions so that $f(x) = g(x)$ for all $x \in \mathbb{Q}$. Prove that $f = g$, i.e. $f(x) = g(x)$ for all $x \in \mathbb{R}$.

    Hint: Prove this by contradiction; the Density Theorem may be useful.

21. Let $f : \mathbb{R} \to \mathbb{R}$ be **additive**, meaning that $f(x + y) = f(x) + f(y)$ for all $x, y \in \mathbb{R}$. Prove that if f is additive and continuous, then $\exists\, r \in \mathbb{R}$ so that $f(x) = rx$.

    Hint: The preceding homework exercise may be helpful.

22. In this problem we prove that monotone functions from $\mathbb{R}$ to $\mathbb{R}$ can have only countably many points z at which they are discontinuous.

    To prove this, for now let $f : \mathbb{R} \to \mathbb{R}$ be increasing. Define $j : \mathbb{R} \to \mathbb{R}$ by setting

    $$j(z) = \inf\{f(x) : x > z\} - \sup\{f(x) : x < z\}.$$

    (j is for "jump", because it is intended to measure the size of any jump in the graph of f at z.)

a) Prove $j(z) \geq 0$ for all $z \in \mathbb{R}$.

b) Let $z \in \mathbb{R}$. Prove that f is continuous at z if and only if $j(z) = 0$.

c) Prove that the set of points at which f fails to be continuous is a countable set.

Hints: For each $n \in \mathbb{Z}$ and $k \in \mathbb{N}$, let

$$D_{n,k} = \left\{ z \in [n, n+1] : j(z) > \frac{f(n+1) - f(n)}{k} \right\}.$$

What is the cardinality of each $D_{n,k}$? How do the sets $D_{n,k}$ relate to the set of points where f fails to be continuous?

d) Use part (c) to prove that if f is decreasing, then f can have only countably many points z at which it is discontinuous.

Exercises from Section 5.4

23. Use the ϵ,δ-definition of limit to prove $\lim_{x \to 3} \left(x^2 - x + 1 \right) = 7$.

24. Give a rigorous argument that $\lim_{x \to 1} \frac{x^3 - 1}{x - 1} = 3$.

25. Prove $\lim_{x \to 0} \frac{x^2}{|x|} = 0$.

26. Prove $\lim_{x \to 0} \frac{|x|}{x}$ does not exist.

27. Prove $\lim_{x \to 0} \lfloor x \rfloor$ does not exist.

28. Prove the second statement of the Main Limit Theorem for functions (Theorem 5.21), which says that if $f, g : \mathbb{R} \to \mathbb{R}$ are such that $\lim_{x \to a} f(x) = L$ and $\lim_{x \to a} g(x) = M$, then $\lim_{x \to a} [f(x) + g(x)] = L + M$. But since we already proved this in the notes, there's a catch: your proof must use only the ϵ,δ-definition of limit and not refer to sequences that converge to a.

Exercises from Section 5.5

29. Let $\{f_n\}$ be the sequence of functions $[0, 1] \to \mathbb{R}$ defined by $f_n(x) = \frac{nx}{1 + nx}$. Use Theorem 5.25 to show that $\{f_n\}$ does <u>not</u> converge uniformly on $[0, 1]$.

30. a) Prove **Dini's Theorem**, which says that if E is compact and $\{f_n\}_n$ is a monotone sequence of continuous functions ("monotone" here means that for all $x \in X$, $\{f_n(x)\}$ is an increasing sequence of real numbers or

that for all $x \in X$, $\{f_n(x)\}$ is a decreasing sequence of real numbers), and if $f_n \to f$ where f is continuous on E, then $f_n \rightrightarrows f$ on E.

Hint: First assume WLOG that $\{f_n\}$ is increasing. Fix $\epsilon > 0$ and let $U_n \subseteq E$ be defined by $U_n = \{x \in X : f(x) - f_n(x) < \epsilon\}$. Show that the $\{U_n\}$ comprise an open cover of E.

b) Show by explicit example that Dini's Theorem may fail if the pointwise limit f is not continuous (but all the other hypotheses are satisfied).

Chapter 6

Differentiation

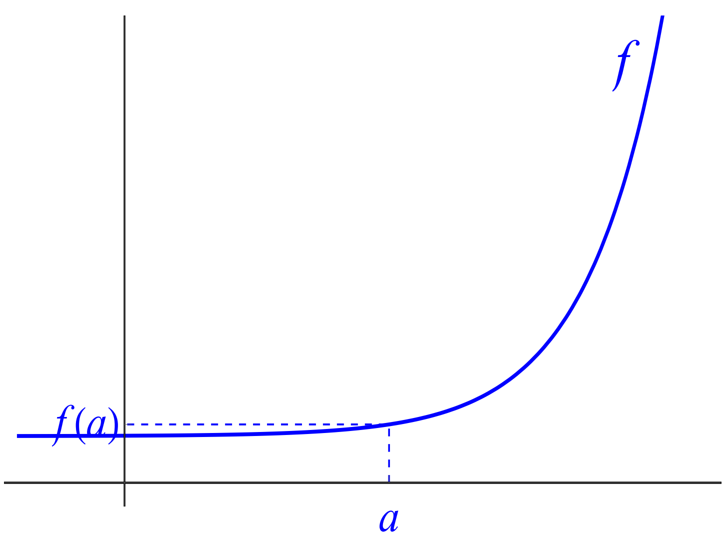
6.1 Definition of the derivative

A refresher on the limit definition of derivative

Definition 6.1 *A function* $f : \mathbb{R} \to \mathbb{R}$ *is called* **differentiable (diffble) at** $a \in \mathbb{R}$ *if there is a number* $f'(a) \in \mathbb{R}$ *such that*

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = f'(a).$$

In this case, f *is called* **differentiable at** a *and the number* $f'(a)$ *is called the* **derivative of** f **at** a.



Definition 6.2 *Let $E \subseteq \mathbb{R}$.*
A function $f : \mathbb{R} \to \mathbb{R}$ is called **differentiable on** *E if f is differentiable at every $a \in E$.*
f is called **differentiable** *if it is differentiable at every point in its domain, in which case the function f' which takes a to $f'(a)$ is called the* **derivative of** *f.*

Definition 6.3 *Let $f : \mathbb{R} \to \mathbb{R}$.*
The **zeroth derivative** *of f, denoted $f^{(0)}$, is f itself.*
The **first derivative of** *f, denoted $f^{(1)}$, is the derivative of f.*
For each $n \in \mathbb{N}$, define the n^{th} **derivative of** *f, denoted $f^{(n)}$, recursively by setting*

$$f^{(n)} = \left(f^{(n-1)}\right)'.$$

We also denote the n^{th} derivative of f with primes: $f^{(2)} = f''$; $f^{(3)} = f'''$; etc.
f is called **n-times differentiable** *or* **differentiable n times** *on $E \subseteq \mathbb{R}$ if $f^{(n)}(x)$ exists for every $x \in E$.*

Theorem 6.4 (Alternate definition of the derivative) *Suppose $f : \mathbb{R} \to \mathbb{R}$.*

- *If*
$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$
exists, then f is differentiable at a and the value of this limit is $f'(a)$.
- *If*
$$\lim_{h \to 0} \frac{f(a) - f(a-h)}{h}$$
exists, then f is differentiable at a and the value of this limit is $f'(a)$.

PROOF Let $L = \lim_{x \to a} \dfrac{f(x) - f(a)}{x - a}$.

To prove the first statement, let $\epsilon > 0$.

By the definition of limit, $\exists$ ______ such that ____________________ implies

________________________ .

Now, for h such that $0 < |h| < \delta$, $|(a + h) - a| \leq |h| < \delta$.

So, by letting $x = a + h$ we have $0 < |x - a| < \delta$ so

$$\left|\frac{f(x) - f(a)}{x - a}\right| = \left|\frac{f(a+h) - f(a)}{a + h - a}\right| = \left|\frac{f(a+h) - f(a)}{h} - L\right| < \epsilon.$$

By the definition of limit,

$$L = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

By uniqueness of limit and definition of the derivative, $L = f'(a)$.

The proof of the second statement is a HW problem. □

Theorem 6.5 (Constant Function Rule) *Let $c \in \mathbb{R}$ and let $f : \mathbb{R} \to \mathbb{R}$ be a constant function $f(x) = c$. Then f is differentiable, and $f'(x) = 0$.*

PROOF HW (use a limit definition of derivative)

Theorem 6.6 (Identity Function Rule) *Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = x$. Then f is differentiable, and $f'(x) = 1$.*

PROOF For all $x \in \mathbb{R}$,

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{x + h - x}{h} = \lim_{h \to 0} \frac{h}{h} = \lim_{h \to 0} 1 = 1.$$

This proves the theorem. □

Theorem 6.7 (Reciprocal Rule) *Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = \frac{1}{x}$. Then f is differentiable at every $x \neq 0$, and $f'(x) = \frac{-1}{x^2}$.*

PROOF HW (you must use a limit definition of derivative, not any other rule)

Theorem 6.8 (Differentiability implies continuity) *Suppose $f : \mathbb{R} \to \mathbb{R}$ is differentiable at $a \in \mathbb{R}$. Then f is continuous at a.*

PROOF Suppose f is differentiable at a. Then

$$\begin{aligned}
\lim_{x \to a} [f(x) - f(a)] &= \lim_{x \to a} \frac{f(x) - f(a)}{x - a}(x - a) \\
&= \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \cdot \lim_{x \to a}(x - a) \\
&= \qquad f'(a) \qquad \cdot \qquad 0 \\
&= 0.
\end{aligned}$$

We have shown $\lim_{x \to a} [f(x) - f(a)] = 0$. Therefore

$$\left[\lim_{x \to a} f(x)\right] - f(a) = 0,$$

i.e.

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

making f continuous at a. □

The meaning of differentiability

In Calculus 1 you are taught that the derivative $f'(a)$ measures

More formally, to say f is differentiable at a means that **locally** (meaning near a), f is very well-approximated by a linear function near a. (What does "very" mean?)

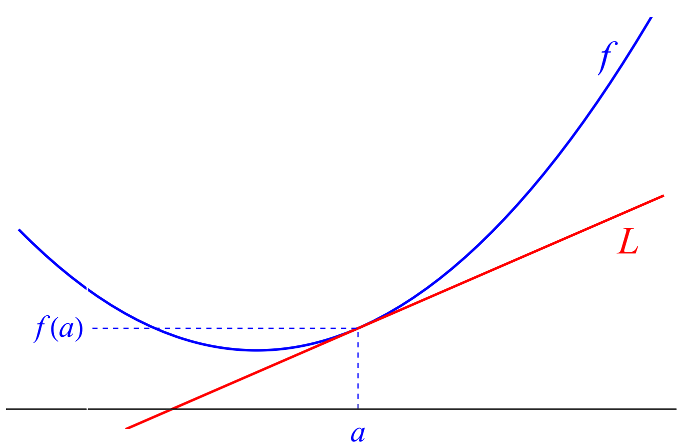
Theorem 6.9 *Let $f : \mathbb{R} \to \mathbb{R}$. TFAE:*

1. *f is differentiable at a.*
2. *There is a linear function $L(x) = mx + b$ so that for all $\epsilon > 0$, there is $\delta > 0$ such that*

$$|x - a| < \delta \text{ implies } |f(x) - L(x)| \le \epsilon |x - a|.$$

(This is what we mean by saying f is "very" well-approximated by L.)

A picture to explain:



PROOF We begin by proving (1) ⇒ (2). Assume f is diffble at a.

Set $L(x) = f(a) + f'(a)(x - a)$.

(This is the point-slope equation of the tangent line to f at a.)

L is linear, with slope $m = f'(a)$ and y-int $b = f(a) - f'(a)a$.

Notice $L(a) = f(a) + f'(a)(a - a) = f(a) + 0 = f(a)$.

Now, let $\epsilon > 0$. Since we are assuming f is diffble at a, $\exists\, \delta > 0$ s.t.

$$0 < |x - a| < \delta \text{ implies } \left|\frac{f(x) - f(a)}{x - a} - f'(a)\right| < \epsilon.$$

Now we prove (2) ⇒ (1). Observe that from (2),

$$|f(a) - L(a)| \leq \epsilon|a - a| = 0$$

so $f(a) - L(a) = 0$, i.e. $L(a) = f(a)$.

That means the linear function L passes through $(a, f(a))$.

Next, let m be the slope of L. By the point-slope formula,

$$L(x) = f(a) + m(x - a). \tag{6.1}$$

Now let $\epsilon > 0$. From what is given in Statement (2), $\exists\, \delta > 0$ so that

$$|x - a| < \delta \text{ implies } |f(x) - L(x)| \leq \frac{\epsilon}{2}|x - a|.$$

Let x be such that $0 < |x - a| < \delta$. Then:

$$\begin{aligned}
|f(x) - L(x)| &\leq \frac{\epsilon}{2}|x - a| \\
\Rightarrow |f(x) - L(x)| &< \epsilon|x - a| \\
\Rightarrow |f(x) - (f(a) + m(x - a))| &< \epsilon|x - a| \qquad \text{(from (6.1))} \\
\Rightarrow |f(x) - f(a) - m(x - a)| &< \epsilon|x - a| \\
\Rightarrow \frac{|f(x) - f(a) - m(x - a)|}{|x - a|} &< \epsilon \\
\Rightarrow \left|\frac{f(x) - f(a) - m(x - a)}{x - a}\right| &< \epsilon \\
\Rightarrow \left|\frac{f(x) - f(a)}{x - a} - m\right| &< \epsilon.
\end{aligned}$$

Thus $f'(a) = m$ by definition of derivative (so f is diffble at a). □

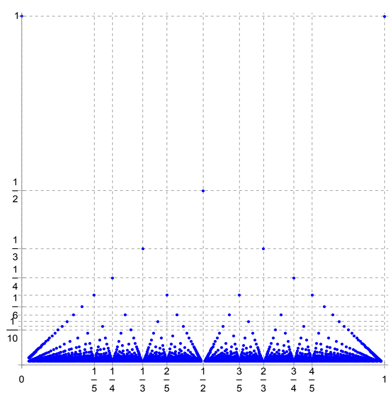
Examples

Example 1

The Dirichlet function $\mathbb{1}_{\mathbb{Q}}$ is not differentiable at any $x \in \mathbb{R}$, since it is not continuous at any x.

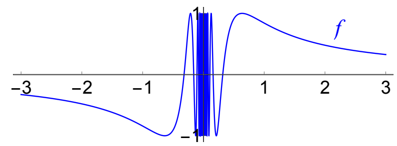
Example 2

Thomae's function τ is not differentiable at any rational number, since it is not continuous there. Is τ differentiable at any irrational numbers? If so, which ones, and what is τ' at those numbers?



EXAMPLE 3

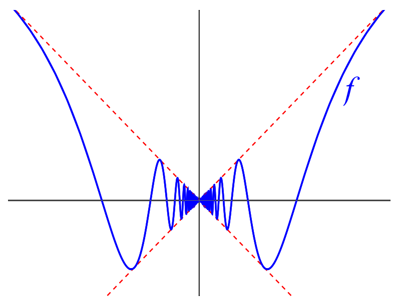
The function $f(x) = \begin{cases} \sin\dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$ is not differentiable at 0, since it is not continuous there.



EXAMPLE 4

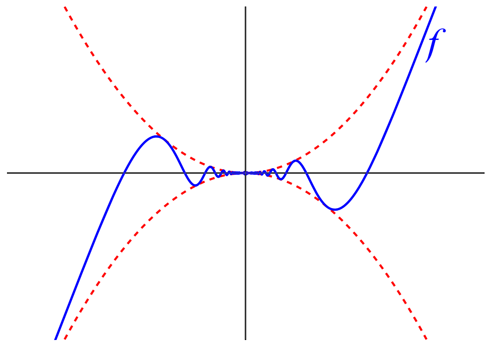
Determine whether the function $f(x) = \begin{cases} x\sin\dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$ is differentiable at 0.

Note: in the previous chapter, we proved this function is continuous at 0.



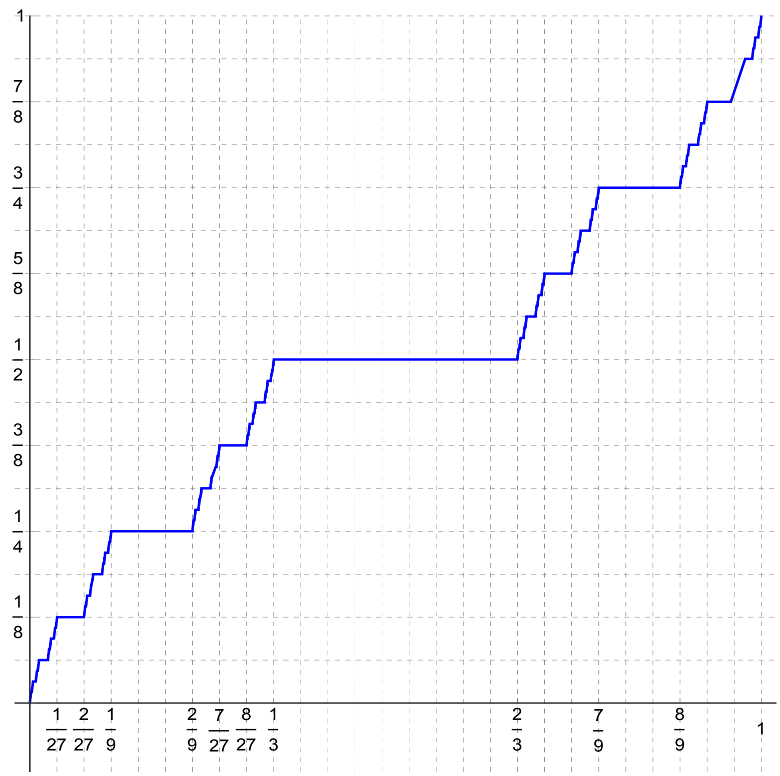
EXAMPLE 5

Determine whether the function $f(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$ is differentiable at 0.



6.2 Differentiability of the Cantor function

Let $c : [0,1] \to [0,1]$ be the Cantor function. At what numbers $a \in [0,1]$ is c differentiable?



Definition 6.10 *The* **(middle-thirds) Cantor set** *is the subset $\mathcal{C}$ of $[0,1]$ consisting of all real numbers in $[0,1]$ that have a base 3 expansion, none of whose digits is 1.*

Theorem 6.11 *Let $c : [0,1] \to [0,1]$ be the Cantor function.*

1. *If $a \notin \mathcal{C}$, then c is differentiable at a and $c'(a) = 0$.*
2. *If $a \in \mathcal{C}$, then c is not differentiable at a.*

PROOF We start with statement (1). Let $a \notin \mathcal{C}$ and let $\epsilon > 0$.

Since $a \notin \mathcal{C}$, then every ternary expansion of a has at least one 1 in it.
For such an expansion, let a_n be the first digit which is a 1. Then

$$a = .a_1a_2\cdots a_{n-1}1a_{n+1}a_{n+2}\cdots_{[3]}.$$

where none of $a_1, a_2, ...a_{n-1}$ are 1.
Given such an a, set

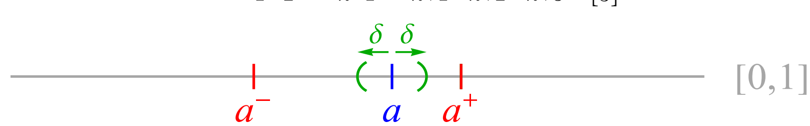
$$\begin{aligned} a^+ &= .a_1a_2\cdots a_{n-1}122222\cdots_{[3]} \\ &= .a_1a_2\cdots a_{n-1}200000\cdots_{[3]} \end{aligned}$$

and

$$\begin{aligned} a^- &= .a_1a_2\cdots a_{n-1}1000000\cdots_{[3]} \\ &= .a_1a_2\cdots a_{n-1}0222222\cdots_{[3]}; \end{aligned}$$

It's clear that $a^- \le a \le a^+$; since a^+ and a^- belong to $\mathcal{C}$ but a doesn't, it must be that $a^- < a < a^+$.
Now, let $\delta = \dfrac{1}{2}\min\{|a - a^-|, |a^+ - a|\}$. If $|x - a| < \delta$, then x has ternary expansion

$$x = .a_1a_2\cdots a_{n-1}1x_{n+1}x_{n+2}x_{n+3}\cdots_{[3]}.$$



Therefore

$$c(x) = c(a),$$

since after Step 1 of the Cantor function process x and a would produce the same sequence.

This implies that for $|h - 0| < \delta$,

$$\left|\frac{c(a+h) - c(a)}{h} - 0\right| = \frac{0}{|h|} = 0 < \epsilon,$$

meaning

$$\lim_{h \to 0} \frac{c(a+h) - c(a)}{h} = 0,$$

i.e. c is differentiable at a with $c'(a) = 0$. This proves (1).

Now for statement (2). Let $a \in \mathcal{C}$.

Since $a \in \mathcal{C}$, we can consider a ternary expansion of a with no 1s:

$$a = .a_1 a_2 a_3 \cdots_{[3]}.$$

Let $b_n = .a_1 a_2 a_3 \cdots a_{n-1} a_n b_{n,n+1} b_{n,n+2} \cdots_{[3]}$, where

$$b_{n,k} = \begin{cases} 2 & \text{if } a_k = 0 \\ 0 & \text{if } a_k = 2 \end{cases}.$$

b_n is also in $\mathcal{C}$, since it has no 1s in its ternary expansion.

Next, let $h_n = b_n - a$ and observe

$$\begin{aligned} |h_n| = |b_n - a| &= \left|\sum_{k=1}^{\infty} \frac{b_{n,k}}{3^k} - \sum_{n=1}^{\infty} \frac{a_n}{3^k}\right| \\ &\leq \sum_{k=1}^{\infty} \left|\frac{b_{k,n}}{3^n} - \frac{a_n}{3^k}\right| \\ &= \sum_{k=n+1}^{\infty} \frac{2}{3^k} \\ &= 2\frac{1}{3^{n+1}} \cdot \frac{1}{1 - \frac{1}{3}} \\ &= \frac{1}{3^n}. \end{aligned}$$

By the Squeeze Theorem, $h_n \to 0$.

However, we can also compute

$$|c(a + h_n) - c(a)| = |c(b_n) - c(a)| = \sum_{k=n+1}^{\infty} \frac{1}{2^k} = \frac{1}{2^n}.$$

This means

$$\left|\frac{c(a+h_n) - c(a)}{h_n}\right| \geq \frac{\frac{1}{2^n}}{\frac{1}{3^n}} = \left(\frac{3}{2}\right)^n \to \infty.$$

Since $h_n \to 0$ but $\left\{ \dfrac{c(a+h_n)-c(a)}{h_n} \right\}$ diverges, it cannot be the case that

$$\lim_{h\to 0} \frac{c(a+h)-c(a)}{h} \text{ exists.}$$

Therefore c is not differentiable at a, proving statement (2). □

More about the Cantor set

Earlier in the course we encountered the *floor function* $\lfloor x \rfloor$:

$$\lfloor x \rfloor = \sup\{n \in \mathbb{Z} : n \le x\}.$$

If you subtract the floor of x from x, you get the *fractional part* of x, denoted $\{x\}$ or by writing "$x \mod 1$":

$$x \mod 1 = \{x\} = x - \lfloor x \rfloor.$$

EXAMPLES

$$\begin{aligned}
\{5\} &= \\
\left\{\frac{7}{4}\right\} &= \\
\{-3.62\} &= \\
\{\pi\} &= \\
\{10\pi\} &= \\
\{10^4\pi\} = \{10000\pi\} &=
\end{aligned}$$

CONCEPT

When you multiply a number in base 10 by a power of 10 like 10^k, all that does is shift the decimal point by k places.

That means the floor of $10^k x$ is the same as the floor of x, except that the first k decimal digits are "erased".

The next lemma says that the same thing works in base b: if you multiply by b^k, the base b expansion of x is the same except the "no-longer-decimal" point is shifted by k places.

Lemma 6.12 *Let $b \in \{2, 3, 4, \ldots\}$ and suppose $x \in [0, 1]$ has base b expansion*

$$x = .x_1 x_2 x_3 \cdots_{[b]}.$$

Then:

1. *A base b expansion of bx is $bx = x_1 \,.\, x_2 x_3 x_4 \cdots_{[b]}$.*
2. *For every $k \in \{1, 2, 3, \ldots\}$, a base b expansion of $b^k x$ is*

$$b^k x = x_1 x_2 \cdots x_k \,.\, x_{k+1} x_{k+2} \cdots_{[b]}$$

and a base b expansion of $\{b^k\}$ is

$$\{b^k\} = .x_{k+1} x_{k+2} \cdots_{[b]}.$$

PROOF We only have to prove the first claim of statement (2), since statement (1) follows by setting $k = 1$, and the second part of statement (2) follows from the second by dropping the digits before the "decimal" point. Let x have the indicated base b expansion. Then

$$x = \sum_{n=0}^{\infty} \frac{x_n}{b^n},$$

so for any $k \in \{1, 2, 3, \ldots\}$,

$$\begin{aligned} b^k x &= b^k \sum_{n=1}^{\infty} \frac{x_n}{b^n} \\ &= b^k \frac{x_1}{b} + b^k \frac{x_2}{b^2} + \cdots + b^k \frac{x_k}{b^k} + b^k \frac{x_{k+1}}{b^{k+1}} + b^k \frac{x_{k+2}}{b^{k+2}} + \cdots \\ &= b^{k-1} x_1 + b^{k-2} x_2 + \cdots + x_k + \frac{x_{k+1}}{b} + \frac{x_{k+2}}{b^2} + \ldots \\ &= x_1 x_2 \cdots x_k \,.\, x_{k+1} x_{k+2} \cdots_{[b]}. \ \square \end{aligned}$$

Lemma 6.13 *Let $x \in \mathcal{C}$, the Cantor set. Then $\{3x\} \in \mathcal{C}$.*

PROOF Suppose $x \in \mathcal{C}$. Then $x = .x_1 x_2 x_3 \cdots_{[3]}$ where none of the x_j equal 1.

Then, by the previous lemma, $\{3x\} = .x_2 x_3 x_4 \cdots_{[3]}$.

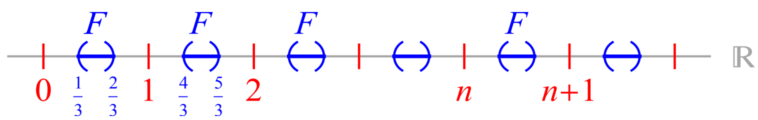
None of the digits in this expansion are 1, so $\{3x\} \in \mathcal{C}$ as wanted. □

Theorem 6.14 (Equivalent characterization of the Cantor set) *Define*

$$F = \bigcup_{n=0}^{\infty} \left(n + \frac{1}{3}, n + \frac{2}{3}\right) = \left(\frac{1}{3}, \frac{2}{3}\right) \cup \left(\frac{4}{3}, \frac{5}{3}\right) \cup \left(\frac{7}{3}, \frac{8}{3}\right) \cup \cdots.$$

Then, the Cantor set $\mathcal{C}$ is the complement of

$$E = \{x \in [0,1] : 3^k x \in F \text{ for some } k \in \mathbb{N}\}.$$



PROOF This is a set equality argument.

$(\mathcal{C} \subseteq E^C)$, i.e. $(E \subseteq \mathcal{C}^C)$:

Suppose $x \in E$. Then $3^k x \in F$ for some $k \in \mathbb{N}$.

So $\{3^k x\} \in \left(\frac{1}{3}, \frac{2}{3}\right)$, so any ternary expansion of $\{3^k x\}$ must start with 1:

$$\{3^k x\} = .1x_2x_3\cdots_{[3]}.$$

By the previous lemma, any ternary expansion of x must look like

$$x = .y_1y_2y_3\cdots y_k1x_2x_3\cdots_{[3]}$$

for suitable digits $y_1, y_2, \ldots, y_k$.

Since such an expansion of x has a digit 1 in it, $x \notin \mathcal{C}$.

This proves $E \subseteq \mathcal{C}^C$; the set inclusion $\mathcal{C} \subseteq E^C$ follows by contraposition.

$(E^C \subseteq \mathcal{C})$, i.e. $(\mathcal{C}^C \subseteq E)$: Suppose $x \notin \mathcal{C}$.

So for any ternary expansion $x = .x_1x_2x_3\cdots_{[3]}$, at least one digit is 1.

Let n be the smallest index such that $x_n = 1$. Now, from the preceding lemmas,

$$\begin{aligned}\{3^{n-1}x\} &= .x_nx_{n+1}x_{n+2}\cdots_{[3]} \\ &= .1x_{n+1}x_{n+2}\cdots_{[3]}\end{aligned}$$

has initial digit 1, so $\{3^{n-1}x\} \in \left[\frac{1}{3}, \frac{2}{3}\right]$.

- If $\{3^{n-1}x\} = \frac{1}{3}$, then $\{3^{n-1}x\} = .022222\cdots_{[3]}$ and therefore
$$x = .x_1x_2\cdots x_{n-1}02222\cdots_{[3]}$$
is a ternary expansion of x with no 1s in it, meaning $x \in \mathcal{C}$, a contradiction. (None of the $x_1, ..., x_{n-1}$ are 1, by the definition of n.)

- If $\lfloor 3^n x \rfloor = \frac{2}{3}$, then $3^n x = .20000\cdots_{[3]}$, and therefore
$$x = .x_1x_2\cdots x_{n-1}2000000\cdots_{[3]}$$
is a ternary expansion of x with no 1s in it, meaning $x \in \mathcal{C}$, a contradiction.

Therefore $\{3^n x\} \in \left(\frac{1}{3}, \frac{2}{3}\right)$, meaning $3^n x \in F$, so $x \in E$ as wanted.

This proves $\mathcal{C}^C \subseteq E$, so $E^C \subseteq \mathcal{C}$ by contraposition. □

Theorem 6.15 *The Cantor set $\mathcal{C}$ is closed (and therefore also compact, since it is clearly bounded by 0 and 1).*

Proof HW

Hints: Note the set F defined in Theorem 6.14 is open. Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = 3x$. This f is continuous, which tells you something about the sets $E_1 = f^{-1}(F)$, $E_2 = f^{-1}(E_1) = f^{-1}(f^{-1}(F))$, $E_3 = f^{-1}(E_2)$, etc. These sets E_j have something to do with the E in Theorem 6.14.

Theorem 6.16 *The Cantor set $\mathcal{C}$ is uncountable.*

Proof HW

Hint: Suppose $\mathcal{C}$ is countable. Write the ternary expansion of the elements you counted, and then construct a number in $\mathcal{C}$ that wasn't on your list, with a procedure similar to how we proved $[0, 1]$ wasn't countable in Chapter 3.

Theorem 6.17 *The Cantor set $\mathcal{C}$ is perfect, meaning that for every $x \in \mathcal{C}$ and every $\epsilon > 0$, there is $y \in (B_\epsilon(x) \cap \mathcal{C}) - \{x\}$.*

Proof HW

Hints: Since $\left(\frac{1}{3}\right)^n \to 0$, given any $\epsilon > 0$, we can choose n so that $\left(\frac{1}{3}\right)^n < \epsilon$. Now, take a ternary expansion $.x_1x_2x_3\cdots_{[3]}$ of $x \in \mathcal{C}$. Use this ternary expansion to cook up a $y \in \mathcal{C}$ which isn't x (because it has a digit different from x, and because it only has one ternary expansion) but is within distance ϵ of x. To ensure $|y - x| < \epsilon$, use the n chosen at the start of this hint.

Theorem 6.18 *The Cantor set $\mathcal{C}$ is totally disconnected, meaning that it does not contain any interval of positive length.*

PROOF HW

Hint: Let E and F be as in Theorem 6.14. Prove this by contradiction: suppose $\mathcal{C}$ contains an interval (a, b) with $a < b$. Explain why this interval (a, b) must contain an x with $3^k x \in F$ for some $k \in \mathbb{N}$.

6.3 Differentiation rules for elementary functions

Linearity, Product and Power Rules

Theorem 6.19 (Linearity of Differentiation) *Suppose $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are differentiable at a. Then:*

1. **Constant Multiple Rule:** *for any $r \in \mathbb{R}$, rf is diffble at a and $(rf)'(a) = rf'(a)$;*
2. **Sum Rule:** *$f + g$ is diffble at a, and $(f + g)'(a) = f'(a) + g'(a)$;*
3. **Difference Rule:** *$f - g$ is diffble at a, and $(f - g)'(a) = f'(a) - g'(a)$.*

PROOF HW

Theorem 6.20 (Product Rule) *Suppose $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are differentiable at a. Then fg is differentiable at a and*

$$(fg)'(a) = f'(a)g(a) + g'(a)f(a).$$

PROOF This is a direct calculation with the limit definition of derivative, using a gimmick of adding and subtracting an extra term in the limit, shown in red below:

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

$$\begin{aligned}
&= \lim_{h\to 0} \frac{f(a+h)g(a+h) - f(a)g(a+h)}{h} + \lim_{h\to 0} \frac{f(a)g(a+h) - f(a)g(a)}{h} \\
&= \lim_{h\to 0} g(a+h)\left[\frac{f(a+h)-f(a)}{h}\right] + \lim_{h\to 0} f(a)\left[\frac{g(a+h)-g(a)}{h}\right] \\
&= \left[\lim_{h\to 0} g(a+h)\right]\left[\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}\right] + f(a)\left[\lim_{h\to 0}\frac{g(a+h)-g(a)}{h}\right] \\
&= \qquad g(a) \qquad \left[\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}\right] + f(a)\left[\lim_{h\to 0}\frac{g(a+h)-g(a)}{h}\right]
\end{aligned}$$

Since f and g are assumed differentiable at a, the blue limits are $f'(a)$ and $g'(a)$, respectively. So the above limit becomes

$$g(a)f'(a) + f(a)g'(a),$$

which is the derivative of fg at a as wanted. □

Theorem 6.21 (Power Rule) *For any $n \in \{1,2,3,\ldots\}$, the function $f(x) = x^n$ is differentiable, and $f'(x) = nx^{n-1}$.*

PROOF # 1 Let's prove this by induction on n.

The base case ($n = 1$) is the Identity Function Rule, done earlier this chapter: (the derivative of $f(x) = x^1 = x$ is $1 = 1x^0 = 1x^{1-1}$).

For the induction step, we need to show:

if $[x^n]' = nx^{n-1}$, **then** $[x^{n+1}]' = (n+1)x^n$.

To verify this, assume the induction hypothesis $f(x) = x^{n+1} = x(x^n)$.
By the Product Rule,

$$\begin{aligned}
f'(x) &= [x]'(x^n) + [x^n]'x \\
&= 1(x^n) + (nx^{n-1})x \\
&= x^n + nx^n \\
&= (n+1)x^n.
\end{aligned}$$

By induction, we are done. □

PROOF # 2: This time, we will use the Binomial Theorem, which says $\forall\, x, h \in \mathbb{R}$,

$$(x+h)^n = \sum_{k=0}^{n} \binom{n}{k} x^k h^{n-k} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} x^k h^{n-k}.$$

(For a proof of this theorem, take MATH 414, or write your own proof.)

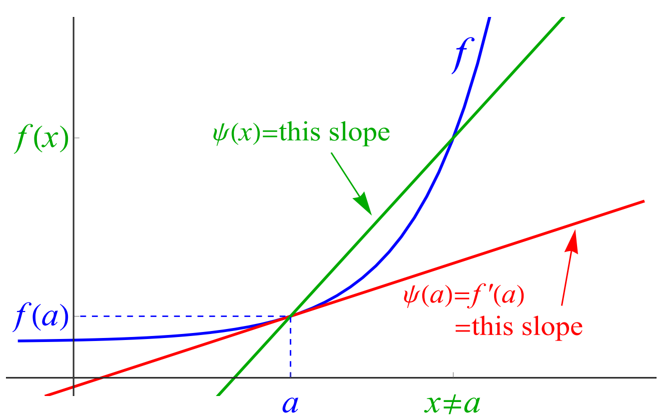
Assuming the Binomial Theorem, the Power Rule follows like this:

$$
\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \\
&= \lim_{h\to 0} \frac{(x+h)^n - x^n}{h} \\
&= \lim_{h\to 0} \frac{\sum\limits_{k=0}^{n} \binom{n}{k} x^k h^{n-k} - x^n}{h} \\
&= \lim_{h\to 0} \frac{\left[x^0 h^n + nx^1 h^{n-1} + \ldots + \binom{n}{n-2} x^{n-2} h^2 + nx^{n-1} h + x^n\right] - x^n}{h} \\
&= \lim_{h\to 0} \frac{h^n + nx^1 h^{n-1} + \ldots + \binom{n}{n-2} x^{n-2} h^2 + nx^{n-1} h}{h} \\
&= \lim_{h\to 0} \frac{h\left[h^{n-1} + nx^1 h^{n-2} + \ldots + \binom{n}{n-2} x^{n-2} h + nx^{n-1} 1\right]}{h} \\
&= \lim_{h\to 0} \left[h^{n-1} + nx^1 h^{n-2} + \ldots + \binom{n}{n-2} x^{n-1} h + nx^{n-1}\right] \\
&= 0 + 0 + 0 + \cdots + 0 + nx^{n-1} = nx^{n-1}. \square
\end{aligned}
$$

Chain Rule

Theorem 6.22 (Carathéodory's Theorem) *Let $f : \mathbb{R} \to \mathbb{R}$ and let $a \in \mathbb{R}$. The following are equivalent:*

1. *f is differentiable at a.*
2. *There exists $\psi : \mathbb{R} \to \mathbb{R}$ such that*
   a) *ψ is continuous at a;*
   b) *$f(x) - f(a) = \psi(x)(x-a)$ for all $x \in \mathbb{R}$; and*
   c) *$\psi(a) = f'(a)$.*



PROOF (1) $\Rightarrow$ (2): Assume f is differentiable at a. Now, define ψ by

$$\psi(x) = \begin{cases} \dfrac{f(x) - f(a)}{x - a} & \text{if } x \neq a \\ f'(a) & \text{if } x = a. \end{cases}$$

ψ obviously satisfies (b) and (c), and it satisfies (a) since by def'n of derivative,

$$\psi(a) = f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \psi(x).$$

(2) $\Leftarrow$ (1): Starting with (b), write $\psi(x) = \dfrac{f(x) - f(a)}{x - a}$ for all $x \neq a$. Since ψ is cts at a,

$$f'(a) = \psi(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

Thus $f'(a)$ exists by a definition of derivative. □

Theorem 6.23 (Chain Rule) *Suppose $g : \mathbb{R} \to \mathbb{R}$ is differentiable at a and $f : \mathbb{R} \to \mathbb{R}$ is differentiable at $g(a)$. Then $f \circ g$ is differentiable at a, and*

$$(f \circ g)'(a) = f'(g(a))g'(a).$$

PROOF By Carathéodory's Theorem, since g is differentiable at a, there is $\psi_g : \mathbb{R} \to \mathbb{R}$ such that

1a) ψ_g is cts at a;
1b) $g(x) - g(a) = \psi_g(x)(x - a)$; and
1c) $g'(a) = \psi_g(a)$.

Applying Carathéodory's Theorem again, since f is diffble at $g(a)$, there is $\psi_f : \mathbb{R} \to \mathbb{R}$ such that

2a) ψ_f is cts at $g(a)$;
2b) $f(y) - f(g(a)) = \psi_f(y)(y - g(a))$; and
2c) $f'(g(a)) = \psi_f(g(a))$.

Now,

$$\begin{aligned} (f \circ g)(x) - (f \circ g)(a) &= f(g(x)) - f(g(a)) \\ &= \psi_f(g(x))(\boxed{g(x) - g(a)}) && \text{(by (2b) above)} \\ &= \psi_f(g(x))\boxed{\psi_g(x)(x - a)}. && \text{(by (1b) above)} \end{aligned}$$

Let $\Psi(x) = \psi_f(g(x))\psi_g(x)$. From the previous page, we have

$$(f \circ g)(x) - (f \circ g)(a) = \Psi(x)(x - a).$$

Furthermore, Ψ, being made up of functions which are cts at a, is cts at a.
By Carathéodory's Theorem, $(f \circ g)$ is diffble at a and

$$(f \circ g)'(a) = \Psi(a) = \psi_f(g(a))\psi_g(a) = f'(g(a))g'(a). \square$$

Quotient Rule

Theorem 6.24 (Quotient Rule) *Suppose $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ are differentiable at a. Then, if $g(a) \neq 0$, $\frac{f}{g}$ is differentiable at a and*

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - g'(a)f(a)}{[g(a)]^2}.$$

PROOF We will prove the Quotient Rule by rewriting an arbitrary quotient as a product and a composition, allowing us to use the Product and Chain Rules.

Let $h(x) = \dfrac{1}{g(x)} = [g(x)]^{-1}$. By the Chain Rule and the Reciprocal Rule, h is differentiable at a and

$$h'(a) = -1[g(a)]^{-2}g'(a) = \frac{-g'(a)}{[g(a)]^2}.$$

That means that by the Product Rule,

$$\begin{aligned}\left(\frac{f}{g}\right)'(a) = (fh)'(a) &= f'(a)h(a) + h'(a)f(a) \\ &= f'(a)\frac{1}{g(a)} + \frac{-g'(a)}{[g(a)]^2}f(a).\end{aligned}$$

Add these fractions by finding a common denominator to get

$$\frac{f'(a)g(a)}{[g(a)]^2} - \frac{g'(a)f(a)}{[g(a)]^2} = \frac{f'(a)g(a) - g'(a)f(a)}{[g(a)]^2}. \square$$

EXAMPLE 5, CONTINUED

Earlier, we saw that the function $f(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$ is differentiable at 0, and $f'(0) = 0$.

1. Is f differentiable at x when $x \neq 0$? If so, what is $f'(x)$ for $x \neq 0$?
   (Let's assume here, without proof, that $\sin x$ is differentiable at all x and that $\dfrac{d}{dx}(\sin x) = \cos x$.)
2. Is the derivative f' continuous at 0?

6.4 Optimization

Recall the **Max-Min Existence Theorem**, which says:

> If $f : \mathbb{R} \to \mathbb{R}$ is ________________ on E, where E is ______________ ,
> then f has an absolute maximum and an absolute minimum on E.

In Calculus 1, you learn how to find the absolute maximum and/or absolute minimum of a function on a compact (i.e. closed and bounded) interval. To optimize function f on $[a, b]$ you:

The reason this method is logically sound is because of the following theorem:

Theorem 6.25 (Fermat's Theorem) *Let $f : [a, b] \to \mathbb{R}$ be differentiable.*

1. *If $c \in (a, b)$ is the location of the absolute maximum value of f on $[a, b]$ (i.e. c is such that $f(x) \le f(c)$ for all $x \in [a, b]$), then $f'(c) = 0$.*
2. *If $c \in (a, b)$ is the location of the absolute minimum value of f on $[a, b]$, (i.e. c is such that $f(x) \le f(c)$ for all $x \in [a, b]$), then $f'(c) = 0$.*

Note: This is not Fermat's Last Theorem, or Fermat's Little Theorem. Those are different things, that don't have to do with MATH 430. This is Fermat's Theorem.

PROOF We prove the first statement here.

Assume c is a location of the absolute maximum value of f on $[a, b]$.

By the limit definition of derivative,

$$f'(c) = \lim_{h\to 0} \frac{f(c+h) - f(c)}{h}. \tag{6.2}$$

Since c is the location of the maximum value of f, $f(c) \ge f(c+h)$ for all h.

This means the numerator in (6.2) is ≤ 0, so since limits preserve $\le$, $f'(c) \le 0$. However, we also know from the previous lemma that

$$f'(c) = \lim_{h\to 0} \frac{f(c) - f(c-h)}{h}. \tag{6.3}$$

And since c is the location of the maximum value of f, $f(c) \ge f(c-h)$ for all h. This means the numerator in (6.3) is ≥ 0, so since limits preserve $\ge$, $f'(c) \ge 0$. Since $f'(c) \le 0$ and $f'(c) \ge 0$, it follows that $f'(c) = 0$.

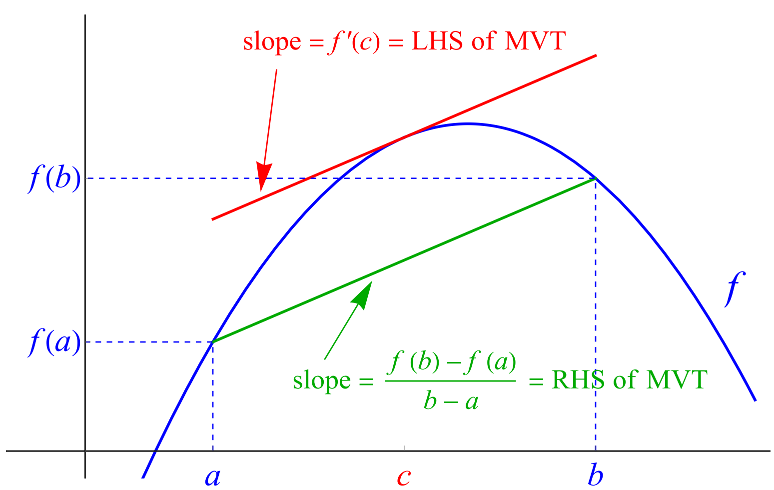
The proof of the second statement is HW. □

6.5 Mean Value Theorem

Theorem 6.26 (Mean Value Theorem (MVT)) *Let $f : \mathbb{R} \to \mathbb{R}$ and let $a < b$ be real numbers. If f is continuous on $[a, b]$ and differentiable on (a, b), then there is $c \in (a, b)$ such that*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Picture that makes the statement "obvious":



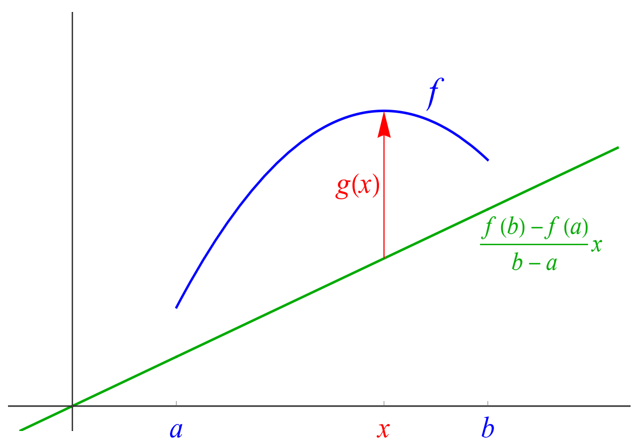
Physics explanation that makes the MVT "obvious":

Suppose $f(t)$ = the position of an object at time t. Then:

The MVT says that given any trip, your instantaneous velocity must equal your average velocity over the whole trip.

PROOF Let $g : [a, b] \to \mathbb{R}$ be defined by

$$g(x) = f(x) - \left(\frac{f(b) - f(a)}{b - a}\right) x. \tag{6.4}$$



Since f is differentiable on (a, b), g is also differentiable on (a, b) and by usual differentiation rules,

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}. \tag{6.5}$$

Furthermore, since f is continuous on $[a, b]$, so is g.

Claim: There exists $c \in (a, b)$ such that $g'(c) = 0$.

Proof of claim: By the Max-Min Existence Theorem, g has an absolute maximum value and an absolute minimum value, occurring at some point in $[a, b]$. There are two cases:

Case 1: Either the abs. max. or the abs. min. of g occurs at $c \in (a, b)$

(i.e. not at an endpoint of $[a, b]$).

In this situation, by Fermat's Theorem, $g'(c) = 0$.

Case 2: The abs. max. and abs. min. values of g occur only at a and b.

Here, observe that the following algebra shows $g(a) = g(b)$:

$$\begin{aligned}
g(b) &= f(b) - \left(\frac{f(b) - f(a)}{b - a}\right) b \qquad \text{(from (6.4))} \\
&= \frac{f(b)(b - a) - (f(b) - f(a))b}{b - a} \\
&= \frac{bf(b) - af(b) - bf(b) + bf(a)}{b - a} \\
&= \frac{bf(a) - af(b)}{b - a} \\
&= \frac{bf(a) - af(a) - af(b) + af(a)}{b - a} \\
&= \frac{f(a)(b - a) - (f(b) - f(a))a}{b - a} \\
&= f(a) - \left(\frac{f(b) - f(a)}{b - a}\right) a \\
&= g(a) \qquad \text{(from (6.4))}.
\end{aligned}$$

So if the endpoints are the locations of the abs. max./min. values of g, then those absolute max./min. values of g must coincide, making g constant on $[a, b]$.

That means $g'(c) = 0$ for every $c \in (a, b)$, proving the claim.

For whatever $c \in (a, b)$ that has $g'(c) = 0$, we have

$$\begin{aligned}
0 &= g'(c) \\
0 &= f'(c) - \frac{f(b) - f(a)}{b - a} \qquad \text{(from (6.5))} \\
\Rightarrow \frac{f(b) - f(a)}{b - a} &= f'(c).
\end{aligned}$$

This completes the proof of the MVT. □

Consequences of the MVT

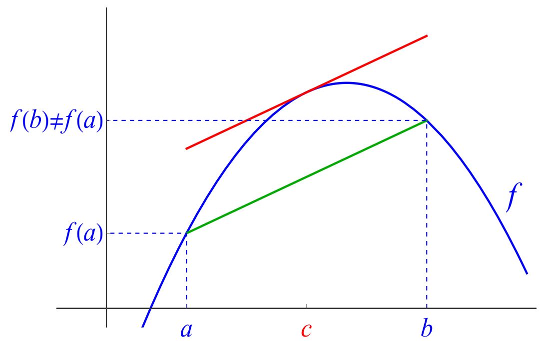
Theorem 6.27 (Zero Derivative Theorem (ZDT)) *Let $f : \mathbb{R} \to \mathbb{R}$ and let α, β be real numbers. If $f'(x) = 0$ for all $x \in (\alpha, \beta)$, then f is constant on (α, β).*

PROOF Suppose not, i.e. that f is nonconstant on (α, β).

That means there are two numbers $a < b$ in (α, β) such that $f(a) \neq f(b)$.

Apply the MVT to find $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$



This is a contradiction! Thus f is constant on (α, β). □

APPLICATION

Suppose $c : \mathbb{R} \to \mathbb{R}$ and $s : \mathbb{R} \to \mathbb{R}$ are two differentiable functions with the properties

$$c'(x) = -s(x); \quad s'(x) = c(x); \quad s(0) = 0; \quad c(0) = 1.$$

Prove that for all $x \in \mathbb{R}$, $[c(x)]^2 + [s(x)]^2 = 1$.

Definition 6.28 *Let $f : \mathbb{R} \to \mathbb{R}$. An* **antiderivative** *of f is another function $'f : \mathbb{R} \to \mathbb{R}$ (pronounced "f antiprime") such that*

$$('f)' = f.$$

EXAMPLE 7

Find an antiderivative of $f(x) = x^2$.

QUESTION

Are antiderivatives unique?

Theorem 6.29 (Antiderivative Theorem) *Let F and G be two antiderivatives of $f : \mathbb{R} \to \mathbb{R}$. Then there is a constant C such that*

$$F(x) = G(x) + C.$$

PROOF Suppose F and G are both antiderivatives of f.

That means $F'(x) = f(x)$ and $G'(x) = f(x)$.

Now, let $H(x) = F(x) - G(x)$. Then,

$$H'(x) = \dots$$

The rest of this proof is HW. □

Definition 6.30 *Let $f : \mathbb{R} \to \mathbb{R}$. The* **indefinite integral** *of f, denoted*

$$\int f(x)\,dx \quad \text{or just} \quad \int f,$$

is the set of all antiderivatives of f.

EXAMPLE 8

Find all antiderivatives of $f(x) = x^2$.

Darboux's Theorem

QUESTION

Let $f : \mathbb{R} \to \mathbb{R}$. Is it necessarily the case that f has an antiderivative?

Theorem 6.31 (Darboux's Theorem) *Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable, and let $a < b$ be real numbers. For any z between $f'(a)$ and $f'(b)$, there is $c \in [a, b]$ such that $f'(c) = z$.*

This looks a lot like the IVT, and follows immediately from the IVT if f' is continuous. But, as we have seen, f' may not be continuous (as an example: let $f(x) = x^2 \sin \dfrac{1}{x}$ for $x \neq 0$ and $f(0) = 0$)!

PROOF If $f'(a) = f'(b)$, the result is vacuously true as there is no z between $f'(a)$ and $f'(b)$.

Assume for now that $f'(a) < f'(b)$ and let $z \in (f'(a), f'(b))$.

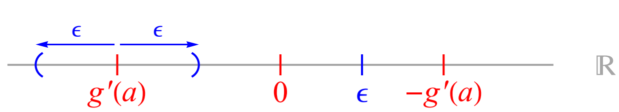
Let $g : \mathbb{R} \to \mathbb{R}$ be $g(x) = f(x) - zx$.

g is diffble since f is, and $g'(x) = f'(x) - z$.

As g is diffble on $[a, b]$, it is cts on $[a, b]$ so by the Max-Min Existence Theorem, g obtains its minimum value on $[a, b]$.

Claim 1: The minimum of g on $[a, b]$ is not achieved at a.

Proof of Claim 1: Observe that $g'(a) = f'(a) - z < 0$.



Let $\epsilon = \dfrac{-g'(a)}{2}$ to obtain a $\delta > 0$ s.t. $0 < x - a < \delta$ implies

$$\left|\frac{g(x) - g(a)}{x - a} - g'(a)\right| < \epsilon = \frac{-g'(a)}{2},$$

which implies $g(x) - g(a) < \dfrac{g'(a)}{2} < 0$ which implies $g(x) < g(a)$.
Thus the minimum of g on $[a, b]$ is not achieved at a, proving Claim 1.

Claim 2: The minimum of g on $[a, b]$ is not achieved at b.
Proof of Claim 2: This is similar to Claim 1. This time, $g'(b) = f'(b) - z > 0$.
Therefore, we can let $\epsilon = \dfrac{g'(a)}{2}$ to obtain $\delta > 0$ s.t. $0 < b - x < \delta$ implies

$$\left|\frac{g(x) - g(b)}{x - b} - g'(b)\right| < \epsilon = \frac{g'(a)}{2},$$

which implies $g(b) - g(x) > \dfrac{g'(a)}{2} > 0$ which implies $g(x) < g(b)$, proving Claim 2.

From Claims 1 and 2, g obtains a minimum value at some $c \in [a, b]$.
By Fermat's Theorem, we have $g'(c) = 0$, i.e. $f'(c) - z = 0$ i.e. $f'(c) = z$.
This proves the theorem in the situation where $f'(a) < f'(b)$.

If $f'(a) > f'(b)$, given $z \in (f'(b), f'(a))$ we can apply the previous case to $h(x) = -f(x)$ and $-z \in (h'(a), h'(b)) = (-f'(a), -f'(b))$ to obtain a c where $h'(c) = -z$, thus $f'(c) = z$. □

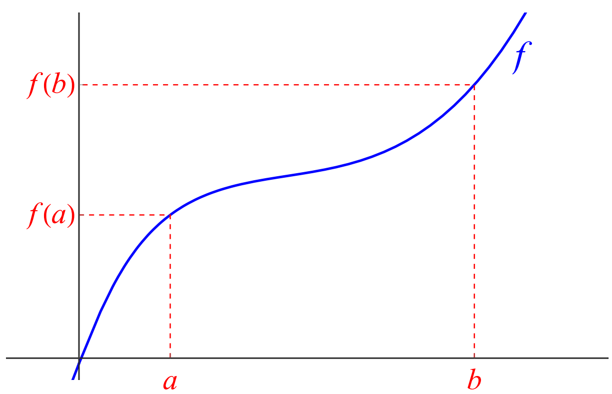
EXAMPLE 9

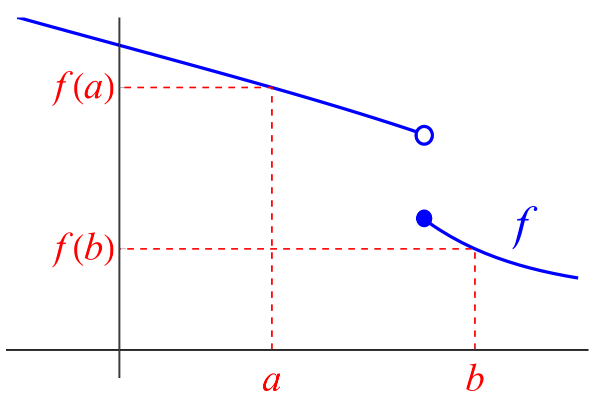
Prove that the Dirichlet function $\mathbb{1}_{\mathbb{Q}}$ has no antiderivative.

Classifying functions as increasing or decreasing

In Calculus 1, you learn that a differentiable function f is **increasing** if ______________ and **decreasing** if ______________ .

But what exactly do *increasing* and *decreasing* mean?





Definition 6.32 *Let $f : \mathbb{R} \to \mathbb{R}$, and let $a < b$.*

1. *f is called* **increasing on** (a, b) *if for every $x, y \in (a, b)$, $x \le y$ implies $f(x) \le f(y)$. (In other words, f preserves soft inequalities.)*
2. *f is called* **strictly increasing on** (a, b) *if for every $x, y \in (a, b)$, $x < y$ implies $f(x) < f(y)$. (In other words, f preserves hard inequalities.)*
3. *f is called* **decreasing on** (a, b) *if for every $x, y \in (a, b)$, $x \le y$ implies $f(x) \ge f(y)$. (In other words, f reverses soft inequalities.)*
4. *f is called* **strictly decreasing on** (a, b) *if for every $x, y \in (a, b)$, $x < y$ implies $f(x) > f(y)$. (In other words, f reverses hard inequalities.)*
5. *f is called* **monotone on** (a, b) *if either (f is increasing on (a, b)) or (f is decreasing on (a, b)).*

Theorem 6.33 (Monotonicity Test) *Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable.*

1. *If $f'(x) \ge 0$ for all $x \in (a, b)$, then f is increasing on (a, b).*
2. *If $f'(x) > 0$ for all $x \in (a, b)$, then f is strictly increasing on (a, b).*
3. *If $f'(x) \le 0$ for all $x \in (a, b)$, then f is decreasing on (a, b).*
4. *If $f'(x) < 0$ for all $x \in (a, b)$, then f is strictly decreasing on (a, b).*

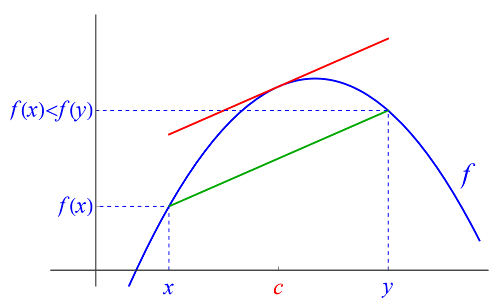
PROOF (1) and (2) are HW.

Hint: The proofs of (3) and (4) can be used as a prototype.

For the third statement, suppose $f'(x) \le 0$ for all $x \in (a, b)$.

Suppose not, then there are $x, y \in (a, b)$ with $x < y$ but $f(x) < f(y)$.

Apply the MVT to get $z \in (x, y)$ with $f'(z) = \dfrac{f(y) - f(x)}{y - x} = \dfrac{\text{positive}}{\text{positive}} > 0$.



This is a contradiction to $f' \le 0$ on (a, b), proving (3).

The proof of (4) is almost identical to that of (3): replace the red $\le$ with $<$, the green $<$ with $\le$ and the orange $>$ with $\ge$. □

APPLICATION

Let $f(x) = \dfrac{1}{x}$.

Using the MVT to prove inequalities

The concept in the proof of the preceding theorem can be used to prove lots of inequalities, like these:

EXAMPLE 10

Prove $\sqrt{x} + \dfrac{1}{\sqrt{x}} > 2$ for every $x \geq 1$.

EXAMPLE 11

Prove $\sqrt{1+x} \leq 1 + \dfrac{x}{2}$ for every $x \geq 0$.

6.6 L'Hôpital's Rule

Theorem 6.34 (Cauchy's Mean Value Theorem) *Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $c \in (a, b)$ such that*

$$f'(c)[g(b) - g(a)] = g'(c)[f(b) - f(a)].$$

Note: the equation in the conclusion of the Cauchy MVT can be rewritten as

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

PROOF *Case 1:* $g(b) = g(a)$.

Then, by the MVT applied to g, there is $c \in (a, b)$ such that

$$g'(c) = \frac{g(b) - g(a)}{b - a} = \frac{0}{b - a} = 0.$$

For this c, both sides of the equation in Cauchy's MVT are 0.

Case 2: $g(a) \neq g(b)$. Define $h : \mathbb{R} \to \mathbb{R}$ by

$$h(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)} g(x). \tag{6.6}$$

h is cts on $[a, b]$ and diffble on (a, b), since f and g are.

Observe

$$\begin{aligned} &h(b) - h(a) \\ &= \left[f(b) - \frac{f(b) - f(a)}{g(b) - g(a)} g(b) \right] - \left[f(a) - \frac{f(b) - f(a)}{g(b) - g(a)} g(a) \right] \qquad \text{(by (6.6))} \\ &= \frac{f(b)\,[g(b) - g(a)] - [f(b) - f(a)]\,g(b) - f(a)\,[g(b) - g(a)] + g(a)\,[f(b) - f(a)]}{g(b) - g(a)} \\ &= 0. \end{aligned}$$

So by the MVT applied to h on $[a, b]$, there is $c \in (a, b)$ such that

$$h'(c) = \frac{h(b) - h(a)}{b - a} = \frac{0}{b - a} = 0.$$

But

$$0 = h'(c) = f'(c) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(c) \qquad \text{(using the def'n of } h \text{ in (6.6))}$$

rearranges into

$$\begin{aligned} -f'(c) &= -\frac{f(b)-f(a)}{g(b)-g(a)}g'(c) \\ \Rightarrow \quad f'(c) &= \frac{f(b)-f(a)}{g(b)-g(a)}g'(c) \\ \Rightarrow \quad f'(c)[g(b)-g(a)] &= g'(c)[f(b)-f(a)]. \ \square \end{aligned}$$

Theorem 6.35 (L'Hôpital's Rule) *Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be differentiable functions. Suppose that $a \in \mathbb{R}$ is such that $f(a) = g(a) = 0$. If there exists $\eta > 0$ such that $g'(x) \neq 0$ for all $x \in (a-\eta, a+\eta) - \{a\}$, then*

$$\lim_{x\to a}\frac{f(x)}{g(x)} \overset{L}{=} \lim_{x\to a}\frac{f'(x)}{g'(x)}.$$

Note: The "L" above the = is just a notational device that tells the reader we are using L'Hôpital's Rule.

APPLICATION

Compute $\lim\limits_{x\to 3}\dfrac{x^2-9}{x^2+10x+21}$ without factoring and cancelling.

PROOF We start with the following preliminary fact:

Claim: Given η as in the theorem, $g(x) \neq 0$ for $x \in (a-\eta, a+\eta)$.

Proof of claim: Suppose not, i.e $\exists\, x \in (a-\eta, a+\eta) - \{a\}$ such that $g(x) = g(a) = 0$.

Then, apply the MVT to find c between a and x (hence not equal to a) s.t.

$$g'(c) = \frac{g(x)-g(a)}{x-a} = \frac{0-0}{x-a} = 0.$$

But we have a hypothesis that $g'(c) \neq 0$ for $c \in (a-\eta, a+\eta)$ except when $c = a$.

By contradiction, the claim is true.

Now for the proof of L'Hôpital's Rule. Let $\epsilon > 0$ and let $L = \lim_{x \to a} \frac{f'(x)}{g'(x)}$.
By definition of limit, there is $\delta > 0$ such that

$$0 < |x - a| < \delta \text{ implies } \left| \frac{f'(x)}{g'(x)} - L \right| < \frac{\epsilon}{2}. \tag{6.7}$$

Now, fix x such that $0 < |x - a| < \delta$.
Observe that by the earlier claim, together with the Main Limit Theorem,

$$\lim_{y \to a} f(y) = 0 \Rightarrow \lim_{y \to a} \frac{f(y)}{g(x)} = 0 \text{ and } \lim_{y \to a} g(y) = 0 \Rightarrow \lim_{y \to a} \frac{g(y)}{g(x)} = 0.$$

By applying the Main Limit Theorem again, we see that for any constant K,

$$\begin{aligned} K &= K(1 - 0) + 0 \\ &= K\left(1 - \lim_{y \to a} \frac{g(y)}{g(x)}\right) + \lim_{y \to a} \frac{f(y)}{g(x)} \\ &= \lim_{y \to a} \left[K\left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)} \right]. \end{aligned}$$

Thus, for each K there is $\delta_1 = \delta_1(K) > 0$ so that

$$0 < |y - a| < \delta_1(K) \text{ implies } \left| K\left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)} - K \right| < \frac{\epsilon}{2}. \tag{6.8}$$

Now, let y be between a and x; by Cauchy's MVT, $\exists\, c$ between x and y such that

$$\begin{aligned} f'(c)[g(x) - g(y)] &= g'(c)[f(x) - f(y)] \\ \Rightarrow \frac{f'(c)}{g'(c)} &= \frac{f(x) - f(y)}{g(x) - g(y)} \end{aligned}$$

Substituting into (6.7), we see that for $x, y \in (a - \delta, a + \delta)$ with y between a and x,

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

Rearrange this to get

$$\begin{aligned}
L - \frac{\epsilon}{2} &< \frac{f(x) - f(y)}{g(x) - g(y)} < L + \frac{\epsilon}{2} \\
\left(L - \frac{\epsilon}{2}\right) \frac{g(x) - g(y)}{g(x)} &< \frac{f(x) - f(y)}{g(x)} < \left(L + \frac{\epsilon}{2}\right) \frac{g(x) - g(y)}{g(x)} \\
\left(L - \frac{\epsilon}{2}\right)\left(1 - \frac{g(y)}{g(x)}\right) &< \frac{f(x)}{g(x)} - \frac{f(y)}{g(x)} < \left(L + \frac{\epsilon}{2}\right)\left(1 - \frac{g(y)}{g(x)}\right) \\
\left(L - \frac{\epsilon}{2}\right)\left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)} &< \frac{f(x)}{g(x)} < \left(L + \frac{\epsilon}{2}\right)\left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)} \qquad (5.9)
\end{aligned}$$

Now, let $\delta_2 = \min\left\{\delta, \delta_1\left(L - \frac{\epsilon}{2}\right), \delta_1\left(L + \frac{\epsilon}{2}\right)\right\}$.
This yields, from (6.8), whenever x is such that $0 < |x - a| < \delta_2$,

$$\left|\left(L - \frac{\epsilon}{2}\right)\left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)} - \left(L - \frac{\epsilon}{2}\right)\right| < \frac{\epsilon}{2}$$

and

$$\left|\left(L + \frac{\epsilon}{2}\right)\left(1 - \frac{g(y)}{g(x)}\right) + \frac{f(y)}{g(x)} - \left(L + \frac{\epsilon}{2}\right)\right| < \frac{\epsilon}{2}.$$

Substituting the previous two lines into (5.9), we get

$$\begin{aligned}
\left(L - \frac{\epsilon}{2}\right) - \frac{\epsilon}{2} &< \frac{f(x)}{g(x)} < \left(L + \frac{\epsilon}{2}\right) + \frac{\epsilon}{2} \\
L - \epsilon &< \frac{f(x)}{g(x)} < L + \epsilon
\end{aligned}$$

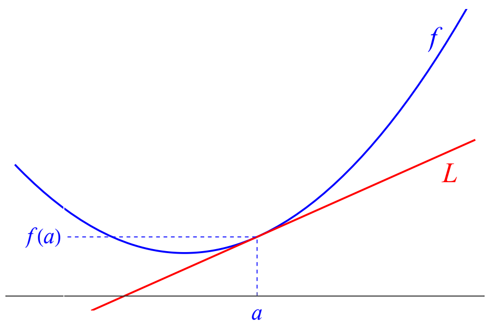
so $\left|\frac{f(x)}{g(x)} - L\right| < \epsilon$. By definition, we have proved

$$\lim_{x \to a} \frac{f(x)}{g(x)} = L = \lim_{x \to a} \frac{f'(x)}{g'(x)}. \quad \square$$

6.7 Taylor polynomials

RECALL

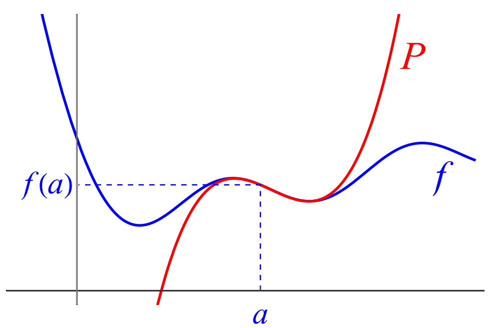
To say that a function f is differentiable at a means it can be very well-approximated by a line L (whose slope is $f'(a)$, i.e. $L'(a) = f'(a)$).
Since the line L has slope $f'(a)$ and passes through $(a, f(a))$, there is only one line that well-approximates f.



In this section, explore approximations of f (near $x = a$) by polynomials. Ostensibly this should lead to approximations that are harder to compute than the tangent line L, but that approximate f better than L does.

FIRST QUESTION

If I want a polynomial P to *well-approximate* f near $x = a$, what does that mean?



SECOND QUESTION

Given f and a, how many polynomials P (of degree $\leq n$) are there that do this?

Lemma 6.36 *Let $E \subseteq \mathbb{R}$ be open and suppose $f : E \to \mathbb{R}$ is differentiable n times at $a \in E$.*
Then, there is exactly one polynomial P_n of degree $\leq n$ so that

$$P_n^{(k)}(a) = f_n^{(k)}(a)$$

for all $k \in \{0, 1, 2, \dots, n\}$.

PROOF Suppose P_n is a polynomial of degree $\leq n$ with $P_n^{(k)}(a) = f^{(k)}(a)$ for all $k \in \{0, 1, \dots, n\}$.

Use algebra to write $P_n(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + \cdots + c_n(x-a)^n$. Then:

$$P_n(a) = c_0 + c_1(a-a) + c_2(a-a)^2 + c_3(a-a)^3 + \cdots + c_n(a-a)^n = c_0$$

$$P_n'(x) = 0 + c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \cdots + nc_n(x-a)^{n-1}$$
$$P_n'(a) = 0 + c_1 + 2c_2(a-a) + 3c_3(a-a)^2 + \cdots + nc_n(a-a)^{n-1} = c_1$$

$$P_n''(x) = 0 + 0 + 2c_2 + 3 \cdot 2c_3(x-a) + \cdots + n(n-1)c_n(x-a)^{n-2}$$
$$P_n''(a) = 0 + 0 + 2c_2 + 3 \cdot 2c_3(a-a) + \cdots + n(n-1)c_n(a-a)^{n-2} = 2c_2$$

$$P_n'''(x) = 0 + 0 + 0 + 3 \cdot 2 \cdot 1c_3 + \cdots + n(n-1)(n-2)c_n(x-a)^{n-3}$$
$$P_n'''(a) = 0 + 0 + 9 + 3!c_3 + \cdots + n(n-1)(n-2)c_n(a-a)^{n-3} = 3!c_3$$

Continuing in this fashion, we see $P_N^{(k)}(a) = k!c_k \ \forall\, k$.

To satisfy the lemma, we must have $f^{(k)}(a) = k!c_k \ \forall\, k$, i.e. $c_k = \dfrac{f^{(k)}(a)}{k!} \forall\, k$.

This forces P_N to have the form described in the lemma. □

Definition 6.37 *Let $E \subseteq \mathbb{R}$ be open, and suppose $f : E \to \mathbb{R}$ is differentiable n times at $a \in E$.*
The n^{th} **Taylor polynomial of f centered at** *a is the polynomial*

$$\begin{aligned} P_n(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \\ &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n. \end{aligned}$$

Concept: P_n is the only polynomial of degree $\leq n$ whose derivatives at a equal the derivatives of f at a up to the n^{th} derivative, so P_n should be the polynomial of degree $\leq n$ that best approximates f near a.

Technicality: When $x = a$, in the formula for Taylor polynomials, we get for the $k = 0$ term

$$\frac{f^{(k)}(0)}{0!}(x-a)^0 = 0^0.$$

This is technically indeterminate, but by convention in Taylor polynomials this 0^0 is **always** 1, i.e. the definition of P_n is really

$$P_n(x) = \begin{cases} \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k & x \neq a \\ f(a) & x = a \end{cases}.$$

and this ensures P_n is n-times differentiable at a.

Low order Taylor polynomials:

- The zeroth Taylor polymomial $P_0(x) = f(a)$ is a constant function;
- The first Taylor polynomial $P_1(x) = f(a) + f'(a)(x-a)$ is the tangent line to f at a.

Taylor's Theorem

RECALL

The MVT says that if f is differentiable 1 time on open interval E, then for any $x, a \in E$,

The equation of the MVT can be rearranged by solving for $f(x)$ and using the language of Taylor polynomials:

Corollary 6.38 (Restated MVT) *Let $E = (a, b) \subseteq \mathbb{R}$ and suppose $f : E \to \mathbb{R}$ is a differentiable function on E. Let $a \in E$ and let P_0 be the first Taylor polynomial of f centered at a.*

Then, $\forall\, x \in E$, $\exists\, c$ between a and x so that

$$f(x) = P_0(x) + \frac{f'(c)}{1!}(x-a)^1.$$

This restated version of the MVT generalizes to higher-order Taylor polynomials as follows:

Theorem 6.39 (Taylor's Theorem) *Let $E = (\alpha, \beta) \subseteq \mathbb{R}$ and suppose $f : E \to \mathbb{R}$ is a function that is differentiable $n+1$ times on E. Let $a \in E$ and let P_n be the n^{th} Taylor polynomial of f centered at a.*
Then, $\forall\, x \in E$, $\exists\, c$ between a and x so that

$$f(x) = P_n(x) + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

PROOF Fix $x \in E$ and define an auxiliary function $g : E \to \mathbb{R}$ by

$$g(t) = \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k + \frac{(x-t)^{n+1}}{(x-a)^{n+1}}\left[f(x) - P_n(x)\right].$$

g is built so that it has the following properties:

- g is differentiable on E;
- $g(a) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{(x-a)^{n+1}}{(x-a)^{n+1}}\left[f(x) - P_n(x)\right] = P_n(x) + f(x) - P_n(x) = f(x)$;
- $g(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(x-x)^k + \frac{(x-x)^{n+1}}{(x-a)^{n+1}}\left[f(x) - P_n(x)\right] = f(x)$.

So by the MVT, $\exists\, c$ between a and x so that

$$g'(c) = \frac{g(x) - g(a)}{x-a} = \frac{f(x) - f(x)}{x-a} = 0.$$

For this c,

$$
\begin{aligned}
0 &= g'(c) \\
&= \frac{d}{dt}[g(t)]_{t=c} \\
&= \frac{d}{dt}\left[\sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k + \frac{(x-t)^{n+1}}{(x-a)^{n+1}}[f(x) - P_n(x)]\right]_{t=c} \\
&= \left[\sum_{k=0}^{n} \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \sum_{k=1}^{n} \frac{f^{(k)}(t)}{k!}k(x-t)^{k-1} - \frac{(n+1)(x-t)^n}{(x-a)^{n+1}}[f(x) - P_n(x)]\right]_{t=c} \\
&= \sum_{k=0}^{n} \frac{f^{(k+1)}(c)}{k!}(x-c)^k - \sum_{k=1}^{n} \frac{f^{(k)}(c)}{(k-1)!}(x-c)^{k-1} - \frac{(n+1)(x-c)^n}{(x-a)^{n+1}}[f(x) - P_n(x)]
\end{aligned}
$$

(change indices on first series)

$$
\begin{aligned}
&= \sum_{k=1}^{n+1} \frac{f^{(k)}(c)}{(k-1)!}(x-c)^{k-1} - \sum_{k=1}^{n} \frac{f^{(k)}(c)}{(k-1)!}(x-c)^{k-1} - \frac{(n+1)(x-c)^n}{(x-a)^{n+1}}[f(x) - P_n(x)] \\
&= \frac{f^{(n+1)}(c)}{n!}(x-c)^n - \frac{(n+1)(x-c)^n}{(x-a)^{n+1}}[f(x) - P_n(x)].
\end{aligned}
$$

Divide through by $(x-c)^n$ to get

$$
0 = \frac{f^{(n+1)}(c)}{n!} - \frac{(n+1)}{(x-a)^{n+1}}[f(x) - P_n(x)].
$$

This rearranges into

$$
\begin{aligned}
\frac{f^{(n+1)}(c)}{n!} &= \frac{(n+1)}{(x-a)^{n+1}}[f(x) - P_n(x)] \\
\Rightarrow \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} &= f(x) - P_n(x) \\
\Rightarrow P_n(x) + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1} &= f(x)
\end{aligned}
$$

which is the formula we want. □

6.8 Interchanging limit and derivative

QUESTION 1

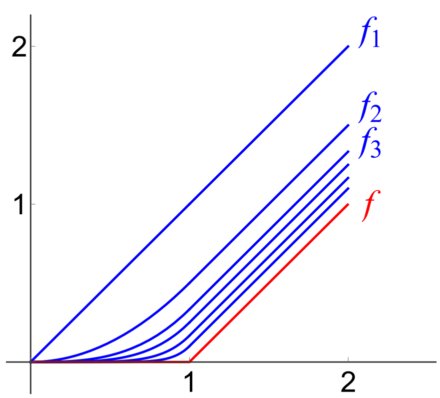
Let $E \subseteq \mathbb{R}$ be open, and $\{f_n\}$ a sequence of differentiable functions $E \to \mathbb{R}$.
If $f_n \to f$ on E, is f necessarily differentiable?

EXAMPLE $E = (-1, 1)$; $f_n(x) = \dfrac{nx}{1 + nx^2}$

QUESTION 2

Let $E \subseteq \mathbb{R}$ be open, and $\{f_n\}$ a sequence of differentiable functions $E \to \mathbb{R}$.
If $f_n \rightrightarrows f$ on E, is f necessarily differentiable?

EXAMPLE $E = (0, 2)$; $f_n(x) = \begin{cases} \dfrac{x^n}{n} & x < 1 \\ x + \dfrac{1}{n} - 1 & x \geq 1 \end{cases}$



QUESTION 3

Let $E \subseteq \mathbb{R}$ be open, and $\{f_n\}$ a sequence of differentiable functions $E \to \mathbb{R}$.
If $f_n \rightrightarrows f$ on E, and f is assumed differentiable, does $f' = \lim(f_n')$?
(In other words, is $(\lim f_n)' = \lim(f_n')$?)

EXAMPLE $E = \mathbb{R}$; $f_n(x) = \dfrac{1}{n} \sin nx$

All the "no" answers on the previous page should not surprise you, because they involve *interchange of limits*:

$$(\lim f_n)'(x) = \lim_{h\to 0}\left(\frac{\lim\limits_{n\to\infty} f_n(x+h) - \lim\limits_{n\to\infty} f_n(x)}{h}\right) = \lim_{h\to 0}\lim_{n\to\infty}\frac{f_n(x+h)-f_n(x)}{h}$$

$$\lim\,(f_n')(x) = \lim_{n\to\infty}\lim_{h\to 0}\frac{f_n(x+h)-f_n(x)}{h}.$$

QUESTION 4

Can you ever interchange limit (of a sequence of functions) and derivative legally?

Theorem 6.40 (Interchange of Limit and Derivative) *Let $E = (\alpha, \beta) \subseteq \mathbb{R}$, and let $\{f_n\}$ be a sequence of differentiable functions $E \to \mathbb{R}$. If*

1. $f_n' \rightrightarrows g$ *on E, and*
2. $\exists\, a \in E$ *s.t.* $\{f_n(a)\}$ *converges,*

then $\exists\, f : E \to \mathbb{R}$ s.t. $f_n \rightrightarrows f$ and $f' = g$.

Note: We are not assuming $f_n \to f$ here; in statement (2) we only assume that there is a single value a so that the sequence $\{f_n(a)\}$ of *numbers* converges.
Rather, the assumption about convergence made is that the *sequence of derivatives* $\{f_n'\}$ converges uniformly on E.

PROOF We are going to start by showing $\{f_n\}$ is uniformly Cauchy.
Toward that end, let $\epsilon > 0$.
Since $\{f_n(a)\}$ converges, $\{f_n(a)\}$ is Cauchy, so $\exists N_1$ s.t.

$$m, n \geq N_1 \Rightarrow |f_m(a) - f_n(a)| < \frac{\epsilon}{2}.$$

Since $f_n' \rightrightarrows g$ on E, $\{f_n'\}$ is uniformly Cauchy, so $\exists N_2$ s.t.

$$m, n \geq N_2 \Rightarrow |f_m'(y) - f_n'(y)| < \frac{\epsilon}{2(\beta - \alpha)} \quad \forall\, y \in E.$$

Now, let $N = \max\{N_1, N_2\}$ and assume $m, n \geq N$.
By the MVT applied to the function $f_m - f_n$ with endpoints a and x, we know $\exists\, y \in E$ s.t.

$$\begin{aligned}
\frac{(f_m - f_n)(x) - (f_m - f_n)(a)}{x - a} &= (f_m - f_n)'(y) \\
\Rightarrow (f_m - f_n)(x) - (f_m - f_n)(a) &= [f_m'(y) - f_n'(y)]\,(x - a) \\
\Rightarrow (f_m - f_n)(x) &= (f_m - f_n)(a) + [f_m'(y) - f_n'(y)]\,(x - a) \\
\Rightarrow f_m(x) - f_n(x) &= f_m(a) - f_n(a) + [f_m'(y) - f_n'(y)]\,(x - a) \\
\Rightarrow |f_m(x) - f_n(x)| &= |f_m(a) - f_n(a) + [f_m'(y) - f_n'(y)]\,(x - a)| \\
&\leq |f_m(a) - f_n(a)| + |f_m'(y) - f_n'(y)|\,|x - a| \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2(\beta - \alpha)}(x - a) \\
&\leq \frac{\epsilon}{2} + \frac{\epsilon}{2(\beta - \alpha)}(\beta - \alpha) \\
&= \epsilon.
\end{aligned}$$

This shows $\{f_n\}$ is uniformly Cauchy on E, so by completeness $\exists\, f : E \to \mathbb{R}$ s.t. $f_n \rightrightarrows f$ on E.

It remains to show $f' = g$. To do this, choose z in E; we will show $f'(z) = g(z)$.
Let $\epsilon > 0$.
Since $\{f_n'\}$ is uniformly Cauchy, $\exists\, M_1$ s.t.

$$m, n \geq M_1 \Rightarrow |f_m'(y) - f_n'(y)| < \frac{\epsilon}{3} \quad \forall\, y \in E.$$

Since $f_n' \rightrightarrows g$, $\exists\, M_2$ s.t.

$$n \geq M_2 \Rightarrow |f_n'(x) - g(x)| < \frac{\epsilon}{3} \quad \forall\, x \in E.$$

Let $M = \max\{M_1, M_2\}$. Since f_M is differentiable at z, $\exists\, \delta > 0$ s.t.

$$0 < |x - z| < \delta \Rightarrow \left| \frac{f_M(x) - f_M(z)}{x - z} - f_M'(z) \right| < \frac{\epsilon}{3}.$$

Suppose x is such that $0 < |x - z| < \delta$, and let $n \geq M$.

Apply the MVT to $f_n - f_M$ with endpoints x and z to get $y \in E$ s.t.

$$\begin{aligned}
\frac{(f_n - f_M)(x) - (f_n - f_M)(z)}{x - z} &= (f_n - f_M)'(y) \\
\Rightarrow \frac{f_n(x) - f_M(x) - f_n(z) + f_M(z)}{x - z} &= f_n'(y) - f_M'(y) \\
\Rightarrow \frac{f_n(x) - f_n(z)}{x - z} - \frac{f_M(x) - f_M(z)}{x - z} &= f_n'(y) - f_M'(y) \\
\Rightarrow \left| \frac{f_n(x) - f_n(z)}{x - z} - \frac{f_M(x) - f_M(z)}{x - z} \right| &= |f_n'(y) - f_M'(y)| \\
\Rightarrow \left| \frac{f_n(x) - f_n(z)}{x - z} - \frac{f_M(x) - f_M(z)}{x - z} \right| &< \frac{\epsilon}{3} \\
\Rightarrow \lim_{n \to \infty} \left| \frac{f_n(x) - f_n(z)}{x - z} - \frac{f_M(x) - f_M(z)}{x - z} \right| &\leq \lim_{n \to \infty} \frac{\epsilon}{3} \\
\Rightarrow \left| \frac{f(x) - f(z)}{x - z} - \frac{f_M(x) - f_M(z)}{x - z} \right| &\leq \frac{\epsilon}{3}.
\end{aligned}$$

Finally,

$$\begin{aligned}
\left| \frac{f(x) - f(z)}{x - z} - g(z) \right| &\leq \left| \frac{f(x) - f(z)}{x - z} - \frac{f_M(x) - f_M(z)}{x - z} \right| \\
&\quad + \left| \frac{f_M(x) - f_M(z)}{x - z} - f_M'(z) \right| \\
&\quad + |f_M'(z) - g(z)| \\
&< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.
\end{aligned}$$

This shows $\lim_{x \to z} \frac{f(x) - f(z)}{x - z} = g(z)$, i.e. $f'(z) = g(z)$. □

Derivatives of transcendental functions

Theorem 6.40 can be used to verify that the derivatives of exp, sin and cos are what they are supposed to be. Recall that

$$\exp x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad \sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \qquad \cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!};$$

these series converge for all $x \in \mathbb{R}$ and converge uniformly on any compact subset of $\mathbb{R}$, so these functions are all continuous.

Theorem 6.41 $\exp : \mathbb{R} \to \mathbb{R}$ *is differentiable, and* $\exp' = \exp$.

PROOF Fix $x \in \mathbb{R}$.

Let S_N be the N^{th} partial sum of the series that defines $\exp$:

$$S_N = \sum_{n=0}^{N} \frac{x^n}{n!}.$$

Since S_N is a polynomial, it is differentiable and

$$\begin{aligned} S_N' &= \frac{d}{dx}\left[\sum_{n=0}^{N} \frac{x^n}{n!}\right] \\ &= \frac{d}{dx}\left[1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^N}{N!}\right] \\ &= 0 + 1 + \frac{2x}{2!} + \frac{3x^2}{3!} + \cdots + \frac{Nx^{N-1}}{N!} \\ &= 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^{N-1}}{(N-1)!} \\ &= S_{N-1}. \end{aligned}$$

Now, fix $x \in \mathbb{R}$.

From Chapter 5, $S_N' = S_{N-1} \rightrightarrows \exp$ on the compact set $E = [-|x| - 1, |x| + 1]$.

We also know that $S_N(0) \to 1$.

So by Theorem 6.40 (with $g = \exp$, $f_n' = S_N'$ and $a = 0$), $\exists\, f : E \to \mathbb{R}$ s.t. $S_N \rightrightarrows f$ on E and $f' = \exp$.

By uniqueness of limits, $f = \exp$ (since $S_N \to \exp$ on $\mathbb{R}$).

Therefore $\exp' = \exp$ as wanted. □

Theorem 6.42 $\sin : \mathbb{R} \to \mathbb{R}$ *is differentiable, and* $\sin' = \cos$.
$\cos : \mathbb{R} \to \mathbb{R}$ *is differentiable, and* $\cos' = -\sin$.

PROOF HW

Hints: Let $S_N = \sum_{n=0}^{N} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$ and $C_N = \sum_{n=0}^{N} (-1)^n \frac{x^{2n}}{(2n)!}$ denote the partial sums of the power series that define sine and cosine, respectively. Show that $S_N' = C_{\text{something}}$ and show that $C_N' = -S_{\text{something}}$, and then use logic similar to the proof that $\exp' = \exp$.

6.9 Taylor series

In Section 6.7 we studied how to approximate n-times differentiable functions f by polynomials P_n. Ostensibly these approximations got better as n got larger.

Suppose now that f is n-times differentiable for every n. This suggests that if we let $n \to \infty$, maybe the Taylor polynomials P_n become a better and better approximation of f, i.e. they converge to f.

This would give a representation of f as a convergent power series.

EXAMPLE

Let $f(x) = \exp(x)$ and let $a = 0$.

Then, since $\exp^{(k)}(x) = \exp x$ for all k, $\exp^{(k)}(a) = \exp(0) = 1$.

That makes the n^{th} Taylor polynomial of f centered at 0

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}(x-0)^k = \sum_{k=0}^{n} \frac{1}{k!}x^k = \sum_{k=0}^{n} \frac{x^k}{k!}.$$

As $n \to \infty$, $P_n \to \sum_{k=0}^{\infty} \frac{x^k}{k!} = \exp(x)$.

So we recover the original function $\exp$ as a limit of its Taylor polynomials centered at 0.

EXAMPLE

Let $f(x) = \sin(x)$ and let $a = 0$.

From the previous section, we see that the derivatives of $f(x) = \sin x$ at 0 are

$$\begin{aligned}
\sin^{(0)}(0) &= \sin 0 = 0 \\
\sin^{(1)}(0) &= \sin'(0) = \cos 0 = 1 \\
\sin^{(2)}(0) &= \sin''(0) = -\sin 0 = 0 \\
\sin^{(3)}(0) &= \sin'''(0) = -\cos 0 = -1 \\
&\vdots \qquad \vdots \\
\sin^{(k)}(0) &= \begin{cases} 1 & \text{if } k-1 \text{ is a multiple of } 4 \\ -1 & \text{if } k-3 \text{ is a multiple of } 4 \\ 0 & \text{if } k \text{ is even} \end{cases}
\end{aligned}$$

Thus the Taylor series of $\sin$ centered at 0 is

$$\begin{aligned}
\sum_{k=0}^{\infty} \frac{\sin^{(k)}(0)}{k!}x^k &= 0 + \frac{1}{1!}x + 0 + \frac{-1}{3!}x^3 + 0 + \frac{1}{5!}x^5 \cdots \\
&= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \\
&= \sin x.
\end{aligned}$$

As with $\exp$, we recover $\sin$ as a limit of its Taylor polynomials centered at 0. (The same thing works with $\cos$, which you will check in the HW.)

QUESTION

Does the same thing happen if we start with a different a and/or different f?

Answer: Sometimes. In this section we investigate this further.

Definition 6.43 *Let $E \subseteq \mathbb{R}$ be open and suppose that $f : E \to \mathbb{R}$ is infinitely differentiable at $a \in E$ (meaning f is n-times differentiable at a for every $n \in \mathbb{N}$).*
The **Taylor series of f centered at** a *is*

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

In other words, the Taylor series of f centered at a is the infinite series whose n^{th} partial sum is P_n, the n^{th} Taylor polynomial of f centered at a.

For any infinitely differentiable f, we can write down its Taylor series. Here's a test you can use to show that series converges to f:

Theorem 6.44 (Convergence of Taylor series) *Let $E = (\alpha, \beta) \subseteq \mathbb{R}$ and suppose $f : E \to R$ is infinitely differentiable on E. Let P_n be the n^{th} Taylor polynomial of f centered at $a \in E$. If, for every $c \in E$,*

$$\frac{f^{(n+1)}(c)}{(n+1)!}(\beta-\alpha)^{n+1} \to 0,$$

then $P_n \to f$ pointwise on E.

PROOF Let $x \in E$. By Taylor's Theorem, $\exists\, c$ between a and x so that

$$f(x) = P_n(x) + \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1},$$

meaning

$$\begin{aligned}
|f(x) - P_n(x)| &= \left|\frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}\right| \\
&= \frac{|f^{(n+1)}(c)|}{(n+1)!}|x-a|^{n+1} \\
&\le \frac{|f^{(n+1)}(c)|}{(n+1)!}(\beta-\alpha)^{n+1} \quad \text{(since } x, a \in (\alpha,\beta)\text{)}.
\end{aligned}$$

By the Squeeze Theorem, $P_n(x) \to f(x)$, i.e. $P_n \to f$ pointwise on E. □

APPLICATION

Consider $f(x) = \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.

Recall the n^{th}**-term test for Divergence** (Exercise 11 of the Chapter 4 HW), which says that if $a_n \not\to 0$, then $\sum a_n$ diverges.

For each $x \in \mathbb{R}$, we know $\sum \frac{x^n}{n!}$ converges (to $\exp(x)$), so by the contrapositive of the n^{th}-term test, $\frac{x^n}{n!} \to 0 \; \forall \, x \in \mathbb{R}$.

We just prove $f^{(n)}(x) = \exp(x)$ for all n, so for any $\alpha < \beta$ and any $c \in (\alpha, \beta)$,

$$\frac{f^{(n+1)}(c)}{(n+1)!}(\beta - \alpha)^{n+1} = \frac{e^c}{(n+1)!}(\beta - \alpha)^{n+1} = e^c \frac{(\beta - \alpha)^{n+1}}{(n+1)!} \to e^c(0) = 0.$$

By Theorem 6.44, any Taylor series of $\exp$ centered at any $a \in \mathbb{R}$ converges to $\exp x$:

$$\sum_{n=0}^{\infty} \frac{e^a}{n!}(x - a)^n = \exp x.$$

QUESTION

If you take *any* infinitely differentiable $f : E \to \mathbb{R}$, does its Taylor series centered at a necessarily converge to f?

Unfortunately, **NO**: consider $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

One can compute that $f^{(n)}(0) = 0$ for all n (HW), so the Taylor series of f is the constant 0. This converges to 0 for all $x \in \mathbb{R}$, but $0 \neq f(x)$ for any $x \neq 0$.

6.10 Properties of transcendental functions

More on the exponential function

Recall that we defined $\exp : \mathbb{R} \to \mathbb{R}$ as a power series: $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$. At this point, here's what we've proven about $\exp$:

- $\exp(0) = 1$
- $\exp(1) = e$ (this is the definition of e)
- for $x \geq 0$, $\exp(x) \geq 1 + x$, so $\{\exp(n)\}_n$ is unbounded as $n \to \infty$.
- $\exp$ is differentiable (hence continuous) and $\exp' = \exp$.

In this section, we rigorously prove other familiar facts about $\exp$.

Exponent rules

First, we can use the fact that $\exp' = \exp$ to recover the exponent rules:

Lemma 6.45 (Exponent rules) *For every $x, y \in \mathbb{R}$ and every $n \in \mathbb{Z}$, we have*

$$\exp(x+y) = \exp(x)\ \exp(y) \qquad \exp(x-y) = \frac{\exp(x)}{\exp(y)} \qquad \exp(nx) = [\exp(x)]^n .$$

PROOF For the first rule, fix y and let $g : \mathbb{R} \to \mathbb{R}$ be $g(x) = \frac{\exp(x+y)}{\exp(x)\exp(y)}$.

Differentiate g using the Quotient Rule and the fact $\exp' = \exp$ to get

$$g'(x) = \frac{\exp(x+y)\exp(x)\exp(y) - \exp(x)\exp(y)\exp(x+y)}{[\exp(x)\exp(y)]^2} = 0.$$

By the Zero Derivative Theorem, g must be constant.

Observe $g(0) = \frac{\exp(0+y)}{\exp(0)\exp(y)} = \frac{\exp(y)}{1\exp(y)} = 1$.

So since g is constant, $g(x) = 1$ for every x, i.e. $\frac{\exp(x+y)}{\exp(x)\exp(y)} = 1$.

This rearranges into $\exp(x+y) = \exp(x)\ \exp(y)$ as wanted.

For the second rule, fix y and let $h : \mathbb{R} \to \mathbb{R}$ be $h(x) = \frac{\exp(x-y)\exp(y)}{\exp(x)}$.

Differentiate h using the Quotient Rule and the fact $\exp' = \exp$ to get

$$h'(x) = \frac{\exp(x-y)\exp(y)\exp(x) - \exp(x)\exp(y)\exp(x-y)}{[\exp(x)]^2} = 0.$$

By the Zero Derivative Theorem, h must be constant.

But $h(0) = \dfrac{\exp(0-y)\exp(y)}{\exp(0)} = \exp(-y)\exp(y) = \exp(-y+y) = \exp(0) = 1.$

Since h is constant, $h(x) = 1$ for every x, i.e. $\dfrac{\exp(x-y)\exp(y)}{\exp(x)} = 1.$

This rearranges into $\exp(x-y) = \dfrac{\exp(x)}{\exp(y)}$ as wanted.

Finally, for the last rule, there are three situations: if $n = 0$, then

$$\exp(nx) = \exp(0x) = \exp(0) = 1 = [\exp(x)]^0 = [\exp(x)]^n.$$

If $n \geq 1$, then applying the first rule we get

$$\begin{aligned}\exp(nx) &= \exp(x + x + x + \cdots + x)\\ &= \exp(x)\exp(x)\exp(x)\cdots\exp(x) \qquad \text{(by the first rule)}\\ &= [\exp(x)]^n.\end{aligned}$$

If $n < 0$, then $-n \geq 1$ so by the second rule and the previous case,

$$\exp(nx) = \exp(0-(-nx)) = \frac{\exp(0)}{\exp(-nx)} = \frac{1}{\exp(-nx)} = \frac{1}{[\exp(x)]^{-n}} = [\exp(x)]^n$$

as wanted.

Connecting $\exp$ with e^x

Recall that we defined the number e to be $e = \exp(1)$. We are now in position to show that the function $\exp$ (defined with a power series) coincides with the function e^x.

This begs the question of what exactly e^x means. When $x \in \mathbb{Q}$, this isn't a problem, because e^x is defined algebraically:

- when $n \in \mathbb{N}$, $e^n = e \cdot e \cdot e \cdots e$.
- when $n = 0$, $e^n = e^0 = 1$.
- when $n \in \mathbb{Z}$ is negative, $-n$ is positive and then $e^n = \dfrac{1}{e^n}$.
- when $n = \dfrac{p}{q} \in Q$, $e^n = e^{p/q} = \sqrt[q]{e^p} = \left(\sqrt[q]{e}\right)^p$, and these roots are guaranteed to exist by work we did in Chapter 2.

But what does e^π mean? More generally, if $x \notin \mathbb{Q}$, what is e^x? If you use algebra alone, **such an expression isn't defined yet**. So for now, put aside what happens when x is irrational and let's worry about rational x:

Lemma 6.46 *For every $x \in \mathbb{Q}$, $e^x = \exp(x)$.*

PROOF First, let $q \in \{1, 2, 3, \dots\}$. By exponent rules proved just earlier,

$$e = \exp(1) = \exp\left(q \cdot \frac{1}{q}\right) = \left[\exp\left(\frac{1}{q}\right)\right]^q$$

so by taking the q^{th} root of both sides, we get $\sqrt[q]{e} = \exp\left(\frac{1}{q}\right)$.

Now, let $x \in \mathbb{Q}$. We can write $x = \dfrac{p}{q}$ with $p, q \in \mathbb{Z}$ and $q > 0$, so

$$\begin{aligned} e^x &= e^{p/q} \\ &= \left[\sqrt[q]{e}\right]^p \\ &= \left[\exp\left(\frac{1}{q}\right)\right]^p \\ &= \exp\left(p \cdot \frac{1}{q}\right) && \text{(by an exponent rule proved earlier)} \\ &= \exp\left(\frac{p}{q}\right). \square \end{aligned}$$

Here's how we handle irrational exponents: we have to define what e^x is when x is irrational. Since we just proved $e^x = \exp(x)$ when $x \in \mathbb{Q}$, it makes sense to simply do this:

Definition 6.47 *If $x \in \mathbb{R} - \mathbb{Q}$, we define e^x to be $e^x = \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.*

To summarize:

- $\exp(x)$ is **defined** for all x as a power series $\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$.
- when $x \in \mathbb{Q}$, e^x is **defined** algebraically with powers and roots, and it turns out that $e^x = \exp(x)$.
- when $x \notin \mathbb{Q}$, e^x is **defined** to be the power series $\exp(x)$.

The point is that we now know $\exp(x) = e^x$ for all $x \in \mathbb{R}$, and that our understanding of exponential functions coming from algebra can be applied to $\exp$.

$\exp$ is a strictly increasing bijection

To finish our discussion of exponential, we'll show that it is strictly increasing and is a bijection from $\mathbb{R}$ to $(0, \infty)$:

Lemma 6.48 *For all* $x \in \mathbb{R}$, $\exp(x) > 0$.

PROOF First, suppose $x \geq 0$.

In this case, since all the terms of the series defining $\exp$ are positive,

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \geq \sum_{n=0}^{1} \frac{x^n}{n!} = 1 + x \geq 1.$$

Now, suppose $x < 0$. Then $-x > 0$ so $\exp(-x) > 0$. But that means

$$\exp(x) = \exp(-(-x)) = \frac{1}{\exp(-x)} = \frac{1}{\text{positive \#}} > 0.$$

In either situation, $\exp(x) > 0$ as wanted. □

Corollary 6.49 $\exp : \mathbb{R} \to \mathbb{R}$ *is strictly increasing.*

PROOF We know $\exp'(x) = \exp(x) > 0$ for all x; apply the Monotonicity Test. □

Lemma 6.50 *The range of* $\exp$ *is* $(0, \infty)$.

PROOF $\exp$ is continuous and $\mathbb{R} = (-\infty, \infty)$ is connected, so by preservation of connectedness the range $\exp(\mathbb{R})$ must be an interval.

By Lemma 6.48, $\exp(\mathbb{R}) \subseteq (0, \infty)$.

Since for $x \geq 0$, $e^x \geq 1 + x$, $\{\exp(n)\}$ is unbounded. That means $\sup(\exp(\mathbb{R})) = \infty$, which means $\exp(\mathbb{R})$ must have the form $[a, \infty)$ or (a, ∞) for some $a \geq 0$.

Now, for any $\epsilon > 0$, choose $x \in \mathbb{R}$ so that $\exp(x) > \frac{1}{\epsilon}$. Then $0 < \exp(-x) < \epsilon$, so ϵ is not a lower bound of $\exp(\mathbb{R})$.

From Lemma 6.48, 0 is a lower bound of $\exp(\mathbb{R})$ and $0 \notin \exp(\mathbb{R})$, so $0 = \inf \exp(\mathbb{R})$ and therefore $\exp(\mathbb{R}) = (0, \infty)$ as wanted. □

Corollary 6.51 $\exp : \mathbb{R} \to (0, \infty)$ *is a bijection.*

PROOF By Corollary 6.49, $\exp$ is strictly increasing, hence injective. By Lemma 6.50, $\exp$ is a surjection onto $(0, \infty)$. □

More on sine and cosine

Here's what we've already shown about sine and cosine:

- $\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$; $\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$.

  These power series converge for every $x \in \mathbb{R}$ and converge uniformly on any compact subset of $\mathbb{R}$.
- $\sin : \mathbb{R} \to \mathbb{R}$ and $\cos : \mathbb{R} \to \mathbb{R}$ are differentiable functions with $\sin' = \cos$ and $\cos' = -\sin$.
- $\sin 0 = 0$ and $\cos 0 = 1$.
- $\sin(-x) = -\sin x$ and $\cos(-x) = \cos x$.

In this section, our goal is to recover familiar properties of sine and cosine.

Pythagorean identity

Lemma 6.52 (Pythagorean identity) *For any $x \in \mathbb{R}$, $\cos^2 x + \sin^2 x = 1$. In other words, for any $x \in \mathbb{R}$ the point $(\cos x, \sin x)$ is on the unit circle in $\mathbb{R}^2$.*

PROOF Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = \cos^2 x + \sin^2 x$.

Differentiate f using the Chain Rule and the facts $\sin' = \cos$, $\cos' = -\sin$ to get

$$f'(x) = 2\cos x(-\sin x) + 2\sin x(\cos x) = 0.$$

By the Zero Derivative Theorem, f is constant.

But $f(0) = \cos^2 0 + \sin^2 0 = 1^2 + 0^2 = 1$, so $f(x) = \cos^2 x + \sin^2 x = 1$ for all x. □

Corollary 6.53 *For any $x \in \mathbb{R}$, $-1 \le \cos x \le 1$ and $-1 \le \sin x \le 1$.*

PROOF Suppose not, then either $\cos^2 x > 1$ or $\sin^2 x > 1$.

That forces $\cos^2 x + \sin^2 x > 1$, contradicting the Pythagorean identity. □

Taylor series and convergence

Here, we check that $\sin$ and $\cos$ are equal to their Taylor series centered at any value of a:

Lemma 6.54 *Let $a \in \mathbb{R}$ and let $P_N(x)$ denote the N^{th} Taylor polynomial for $\sin x$, centered at a. Then $P_N(x) \to \sin x$ on $\mathbb{R}$.*

PROOF Let $\alpha < \beta$ and let $a, c \in \mathbb{R}$.

For any n, $\sin^{(n+1)}$ is one of $\pm\sin$ or $\pm\cos$, and all these functions are bounded by -1 and 1. So

$$\frac{\sin^{(n+1)}(c)}{(n+1)!}(\beta-\alpha)^{n+1} \leq \frac{(\beta-\alpha)^{n+1}}{(n+1)!}.$$

Observe $\sum_{n=0}^{\infty} \frac{(\beta-\alpha)^n}{n!} = \exp(\beta-\alpha)$.

So by the n^{th}-term test for Divergence, $\frac{(\beta-\alpha)^{n+1}}{(n+1)!} \to 0$.

By Theorem 6.44, $P_N(x) \to \sin x$. □

Lemma 6.55 *Let $a \in \mathbb{R}$ and let $P_N(x)$ denote the N^{th} Taylor polynomial for $\cos x$, centered at a. Then $P_N(x) \to \cos x$ on $\mathbb{R}$.*

PROOF HW (this is very similar to the previous proof).

Periodicity

Theorem 6.56 (Periodicity of sine and cosine) *There exists a real number $\pi > 0$ so that for all $x \in \mathbb{R}$,*

$$\sin(x+2\pi) = \sin x \qquad \textit{and} \;\; \cos(x+2\pi) = \cos x.$$

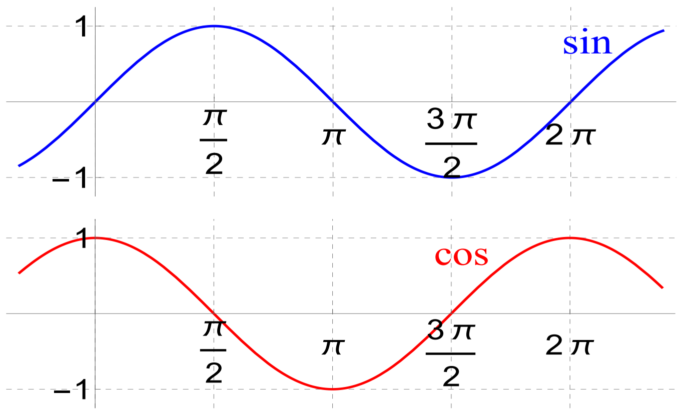
Furthermore, $\sin$ *and* $\cos$ *have the following values:*

x	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
$\sin x$	0	1	0	-1	0
$\cos x$	1	0	-1	0	1

Even further still,

- $\sin$ *strictly increases on* $\left(0, \frac{\pi}{2}\right)$ *and* $\left(\frac{3\pi}{2}, 2\pi\right)$ *and decreases on* $\left(\frac{\pi}{2}, \frac{3\pi}{2}\right)$*, and*
- $\cos$ *strictly decreases on* $(0, \pi)$ *and increases on* $(\pi, 2\pi)$*.*

Concept: this theorem tells us what the graphs of $\cos$ and $\sin$ look like.



PROOF *Claim # 1:* $\exists t_0 \in (0,2)$ so that $\cos t_0 = 0$, but $\cos x > 0$ for $x \in (0, t_0)$.

Proof of Claim # 1: Apply Taylor's Theorem to cos with $n = 3$, $a = 0$ and $x = 2$.

This gives us a $c \in (0,2)$ so that

$$\begin{aligned}\cos 2 &= P_2(2) + \frac{\cos^{(4)}(c)}{4!}(2-0)^4 \\ &= \sum_{k=0}^{2} \frac{\cos^{(k)}(0)}{k!}(2-0)^k + \frac{16}{24}\cos c \\ &= \frac{\cos 0}{0!} + \frac{\cos'(0)}{1!}(2-0) + \frac{\cos''(0)}{2!}(2-0)^2 + \frac{2}{3}\cos c \\ &= \frac{1}{1} + \frac{-\sin 0}{1}(2) + \frac{-\cos 0}{2}(2)^2 + \frac{2}{3}\cos c \\ &= 1 + 0 - \frac{1}{2}(4) + \frac{2}{3}\cos c \\ &= 1 - 2 + \frac{2}{3}\cos c.\end{aligned}$$

Therefore, since $\cos c \le 1$, $\cos 2 \le 1 - 2 + \frac{2}{3} = -\frac{1}{3} < 0$.

Now, we know cos is continuous, $\cos 0 = 1 > 0$ and $\cos 2 < 0$.

By the IVT, $\exists x \in (0,2)$ so that $\cos x = 0$.

Let t_0 be the smallest $x > 0$ so that $\cos t_0 = 0$. $\boxed{\text{Let } \pi = 2t_0}$ so that $t_0 = \frac{\pi}{2}$.

By continuity, $\cos x > 0$ for $x \in \left[0, \frac{\pi}{2}\right)$. For if not, by the IVT there would be $x \in (0, t_0)$ with $\cos x = 0$, violating the definition of t_0 as the smallest positive x with $\cos x = 0$. This proves Claim # 1.

Claim # 2: $\sin\dfrac{\pi}{2} = 1$.

Proof of Claim # 2: By the Pythagorean identity, $\cos^2\dfrac{\pi}{2} + \sin^2\dfrac{\pi}{2} = 1$, so $0^2 + \sin^2\dfrac{\pi}{2} = 1$, so $\sin t_0 \in \{-1, 1\}$.

But for $x \in \left[0, \dfrac{\pi}{2}\right)$, $\sin'(x) = \cos x > 0$ so $\sin$ is strictly increasing on $\left[0, \dfrac{\pi}{2}\right)$, so $\sin\dfrac{\pi}{2} > \sin 0 = 0$ meaning $\sin\dfrac{\pi}{2} = 1$ (as opposed to -1).

Note that in Claim # 2 we proved $\sin$ is strictly increasing on $\left(0, \dfrac{\pi}{2}\right)$.

That means $\sin x > \sin 0 = 0$ for all $x \in \left(0, \dfrac{\pi}{2}\right)$, so $\cos' = -\sin x < 0$ for all $x \in \left(0, \dfrac{\pi}{2}\right)$, and this means $\cos$ is strictly decreasing on $\left(0, \dfrac{\pi}{2}\right)$.

Claim # 3: For all $x \in \mathbb{R}$, $\sin x = \cos\left(x - \dfrac{\pi}{2}\right)$.

Proof of Claim # 3: Since $\cos\dfrac{\pi}{2} = 0$ and $\sin\dfrac{\pi}{2} = 1$, we now know all the derivatives of $\sin$ at $\dfrac{\pi}{2}$:

$$\sin^{(0)}\left(\frac{\pi}{2}\right) = \sin\frac{\pi}{2} = 1$$
$$\sin^{(1)}\left(\frac{\pi}{2}\right) = \sin'\left(\frac{\pi}{2}\right) = \cos\frac{\pi}{2} = 0$$
$$\sin^{(2)}\left(\frac{\pi}{2}\right) = \sin''\left(\frac{\pi}{2}\right) = -\sin\frac{\pi}{2} = -1$$
$$\sin^{(3)}\left(\frac{\pi}{2}\right) = \sin'''\left(\frac{\pi}{2}\right) = -\cos\frac{\pi}{2} = 0 \qquad \vdots$$
$$\sin^{(n)}\left(\frac{\pi}{2}\right) = \begin{cases} 1 & \text{if } n \text{ is a multiple of } 4 \\ -1 & \text{if } n \text{ is even but not a multiple of } 4 \\ 0 & \text{if } n \text{ is odd} \end{cases}$$

From the discussion on Taylor series of $\sin$ and $\cos$, we know $\sin x$ equals its

Taylor series centered at $\frac{\pi}{2}$, i.e.

$$\begin{aligned}\sin x &= \sum_{n=0}^{\infty} \frac{f^{(n)}\left(\frac{\pi}{2}\right)}{n!}\left(x-\frac{\pi}{2}\right)^n \\ &= 1 - \frac{1}{2!}\left(x-\frac{\pi}{2}\right)^2 + \frac{1}{4!}\left(x-\frac{\pi}{2}\right)^4 - \frac{1}{6!}\left(x-\frac{\pi}{2}\right)^t + \dots \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\left(x-\frac{\pi}{2}\right)^{2n} \\ &= \cos\left(x-\frac{\pi}{2}\right) \qquad \text{(by the series definition of } \cos\text{)}.\end{aligned}$$

This proves Claim # 3.

Claim # 4: For all $x \in \mathbb{R}$, $\cos x = -\sin\left(x - \frac{\pi}{2}\right)$.

Proof of Claim # 4: Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = \cos x + \sin\left(x - \frac{\pi}{2}\right)$.

f is differentiable and (applying Claim # 3),

$$f'(x) = -\sin x + \cos\left(x - \frac{\pi}{2}\right) = -\sin x + \sin x = 0.$$

So by the ZDT, f is constant.

When $x = \frac{\pi}{2}$, $f(x) = \cos t_0 + \sin(t_0 - t_0) = 0$.

Since f is constant, $f(x) = 0 \;\forall\, x$, i.e. Claim # 4 holds.

From Claims 3 and 4, we can compute

$$\begin{aligned}\sin\pi &= \cos\left(\pi - \frac{\pi}{2}\right) = \cos\frac{\pi}{2} = 0 \\ \cos\pi &= -\sin\left(\pi - \frac{\pi}{2}\right) = -\sin\frac{\pi}{2} = -1 \\ \sin\frac{3\pi}{2} &= \cos\left(\frac{3\pi}{2} - \frac{\pi}{2}\right) = \cos\pi = -1 \\ \cos\frac{3\pi}{2} &= -\sin\left(\frac{3\pi}{2} - \frac{\pi}{2}\right) = -\sin\pi = -0 = 0 \\ \sin 2\pi &= \cos\left(2\pi - \frac{\pi}{2}\right) = \cos\frac{3\pi}{2} = 0 \\ \cos 2\pi &= -\sin\left(2\pi - \frac{\pi}{2}\right) = -\sin\frac{3\pi}{2} = -(-1) = 1\end{aligned}$$

so the table of values in the theorem is complete.

Also, Claim 3 tells us that the graph of $\sin$ is the graph of $\cos$ shifted right by $\frac{\pi}{2}$ units. So since $\cos$ is strictly decreasing on $\left(0, \frac{\pi}{2}\right)$, $\sin$ is strictly decreasing on $\left(\frac{\pi}{2}, \pi\right)$. But since $\sin \pi = 0$, $\sin x > 0$ for $x \in \left(\frac{\pi}{2}, \pi\right)$, making $\cos$ strictly decreasing on $\left(\frac{\pi}{2}, \pi\right)$ since its derivative is $-\sin x$. This in turn makes $\sin$ strictly decreasing on $\left(\pi, \frac{3\pi}{2}\right)$, making $\sin x < 0$ on $\left(\pi, \frac{3\pi}{2}\right)$, making $\cos$ strictly increasing on $\left(\pi, \frac{3\pi}{2}\right)$, making $\sin$ strictly increasing on $\left(\frac{3\pi}{2}, 2\pi\right)$, making $\sin x < \sin 2\pi = 0$ for $x \in \left(\frac{3\pi}{2}, 2\pi\right)$, making $\cos$ strictly increasing on $\left(\frac{3\pi}{2}, 2\pi\right)$. This takes care of all the statements about the increasing/decreasing nature of $\sin$ and $\cos$ in the theorem.

Claim # 5: $\sin(x - 2\pi) = \sin x$ for all $x \in \mathbb{R}$.

Proof of Claim # 5: This is a direct calculation:

$$\begin{aligned}
\sin x &= \cos\left(x - \frac{\pi}{2}\right) && \text{(by Claim \# 3)}\\
&= -\sin\left(\left(x - \frac{\pi}{2}\right) - \frac{\pi}{2}\right) && \text{(by Claim \# 4)}\\
&= -\sin(x - \pi)\\
&= -\cos\left((x - \pi) - \frac{\pi}{2}\right) && \text{(by Claim \# 3)}\\
&= -\cos\left(x - \frac{3\pi}{2}\right)\\
&= -\left(-\sin\left(\left(x - \frac{3\pi}{2}\right) - \frac{\pi}{2}\right)\right) && \text{(by Claim \# 4)}\\
&= \sin(x - 2\pi).
\end{aligned}$$

From Claim # 5, we see that if we let $y = x + 2\pi$, then

$$\sin(x + 2\pi) = \sin y = \sin(y - 2\pi) = \sin x.$$

Thus $\sin$ is periodic with period 2π, as wanted.

Claim # 6: $\cos(x + 2\pi) = \cos x$.

Proof of Claim # 6: For this, apply Claim # 4 and the fact $\sin$ is periodic:

$$\cos(x + 2\pi) = -\sin\left((x + 2\pi) - \frac{\pi}{2}\right) = -\sin\left(\left(x - \frac{\pi}{2}\right) + 2\pi\right) = -\sin\left(x - \frac{\pi}{2}\right) = \cos x.$$

This, at last, finishes the proof of this theorem. □

Corollary 6.57 $\sin : \mathbb{R} \to [-1,1]$ *and* $\cos : \mathbb{R} \to [-1,1]$ *are surjective.*

PROOF From the previous theorem, $\sin\frac{\pi}{2} = 1$ and $\sin\frac{3\pi}{2} = -1$.

As $\sin$ is cts, by the IVT $\forall\, y \in (-1,1)\, \exists\, x \in \left(\frac{\pi}{2}, \frac{3\pi}{2}\right)$ so that $\sin x = y$.

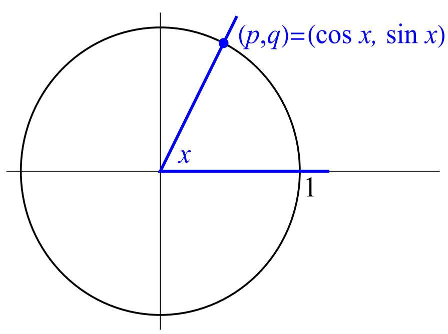
Turning to cosine, from the previous theorem, $\cos 0 = 1$ and $\cos\pi = -1$.

As $\cos$ is cts, by the IVT $\forall\, y \in (-1,1)\, \exists\, x \in (0,\pi)$ so that $\cos x = y$. □

Angle

Corollary 6.58 *For every point* $(p,q) \in \mathbb{R}^2$ *which lies on the unit circle,* $\exists\, x \in \mathbb{R}$ *so that* $\cos x = p$ *and* $\sin x = q$.
Furthermore, any two values of x *satisfying the conclusions of this corollary must differ by a multiple of* 2π.

Concept: we can define any x in this corollary to be the radian measure of the angle formed by the line segment from $(0,0)$ to $(1,0)$ and the line segment from $(0,0)$ to (p,q). This makes $\cos x$ and $\sin x$ the coordinates of the point on the unit circle at angle x, from which the rest of trigonometry (SOHCAHTOA, Laws of Sines and Cosines, Heron's formula, other trig identities, etc.) can be deduced.



There's a little more we need to check, namely that the arc length along the circle from $(1,0)$ to (p,q) is x; we'll have to postpone that until after we've done integrals.

PROOF Let (p,q) be a point on the unit circle.

If $p = 1$, then $q = 0$ so from the periodicity theorem, we can choose x to be any multiple of 2π, but (also by the periodicity theorem) there are no $x \in (0,2\pi)$ with $\cos x = 1$, proving this corollary.

If $p = -1$, then $q = 0$ so from the periodicity theorem, we can choose x to be π plus any multiple of 2π, but (also by the periodicity theorem) there are no $x \in [0,\pi) \bigcup (\pi, 2\pi]$ with $\cos x = -1$, proving this corollary.

When $0 \le p < 1$, since $\cos$ is surjective, $\exists\, x \in (0,2\pi)$ so that $\cos x = p$. From the periodicity theorem, either $x \in \left(0, \dfrac{\pi}{2}\right)$ or $x \in \left(\dfrac{3\pi}{2}, 2\pi\right)$. But since $\cos$ is strictly monotone on these intervals, there is only one x in each interval that works. For the $x \in \left(0, \dfrac{\pi}{2}\right)$, $\sin x > 0$ but for the $x \in \left(\dfrac{3\pi}{2}, 2\pi\right)$, $\sin x < 0$ so there is only one $x \in (0, 2\pi)$ for which $\cos x = p$ and $\sin x = q$ (for the other one, $\sin x = -q$). By periodicity, adding any multiple of 2π to this x produces another valid x, proving this corollary.

Finally, when $-1 < p < 0$, since $\cos$ is surjective, $\exists\, x \in (0, 2\pi)$ so that $\cos x = p$. From the periodicity theorem, either $x \in \left(\dfrac{\pi}{2}, \pi\right)$ or $x \in \left(\pi, \dfrac{3\pi}{2}\right)$, and since $\cos$ is strictly monotone on these intervals, there is only one x in each interval that works. For the $x \in \left(\dfrac{\pi}{2}, \pi\right)$, $\sin x > 0$ but for the $x \in \left(\pi, \dfrac{3\pi}{2}\right)$, $\sin x < 0$ so there is only one $x \in (0, 2\pi)$ for which $\cos x = p$ and $\sin x = q$ (for the other one, $\sin x = -q$). By periodicity, adding any multiple of 2π to this x produces another valid x, proving this corollary. □

6.11 Inverse functions and natural logarithm

RECALL

If a function $f : A \to B$ is bijective, then it is invertible, meaning that there is a function $f^{-1} : B \to A$ so that $f^{-1} \circ f(x) = x$ for all $x \in A$ and $f \circ f^{-1}(x) = x$ for all $x \in B$.

In this section we run through some results telling us about the continuity and differentiability of the inverse of a continuous or differentiable bijection.

Theorem 6.59 (Continuous Inverse Theorem) *Let $E \subseteq \mathbb{R}$ be an open interval, and suppose $f : E \to \mathbb{R}$ is a continuous bijection from E onto $f(E)$ Then $f^{-1} : f(E) \to E$ is continuous.*

Observe: In this setting $f(E)$ must be an interval since continuous functions preserve connectedness.

PROOF We start with this claim:

Claim: If $f : E \to f(E)$ is a continuous bijection, then either f is strictly increasing or f is strictly decreasing.

Proof of claim: Suppose not. Then, $\exists\, x, y, z \in E$ with $x < y < z$ and either

(a) $f(y) < f(x)$ and $f(y) < f(z)$, or

(b) $f(y) > f(x)$ and $f(y) > f(z)$.

In situation (a), there are two possibilities.
The first is that $f(x) < f(z)$, i.e. $f(x) < f(z) < f(y)$.
Since f is cts, by the IVT $\exists\, c \in (x, y)$ s.t. $f(c) = f(z)$.
But $c \neq z$, contradicting f being injective.
The second possibility is $f(z) < f(x) < f(y)$.
By the IVT, $\exists\, c \in (y, z)$ with $f(c) = f(x)$.
But $c \neq x$, contradicting f being injective.
In situation (b), there are also two possibilities.
The first is $f(x) < f(z)$, i.e. $f(y) < f(x) < f(z)$.
By the IVT, $\exists\, c \in (y, z)$ with $f(c) = f(x)$.
But $c \neq x$, contradicting f being injective.
The second case of situation (b) is $f(x) > f(z)$, i.e. $f(y) < f(z) < f(x)$.
By the IVT, $\exists\, c \in (y, x)$ with $f(c) = f(z)$.
But $c \neq z$, contradicting f being injective.
This proves the claim.

Now, to show f^{-1} is continuous, it suffices to show that for any $(a, b) \subseteq E$, $(f^{-1})^{-1}(a, b)$ is open, i.e. $f(a, b)$ is open.
Assume for now that f is strictly increasing.
We will prove $f(a, b) = (f(a), f(b))$ by a a set inclusion argument.
($\subseteq$) Suppose $y \in f(a, b)$. Then $\exists\, x \in (a, b)$ s.t. $f(x) = y$.
Since $x \in (a, b)$, $a < x < b$.
Since f is strictly increasing, $f(a) < f(x) < b$, i.e. $y \in (f(a), f(b))$.
($\supseteq$) Suppose $y \in (f(a), f(b))$.
That means $f(a) < y < f(b)$.
Since f is surjective, $\exists\, x \in E$ s.t. $f(x) = y$.
It must be that $a < x < b$, for if not, since f is strictly increasing we would have $f(a) \geq f(x) = y$ or $y = f(x) \geq f(b)$, both of which contradict $f(a) < y < f(b)$.

If f is strictly decreasing, then $f(a, b) = (f(b), f(a))$ by a similar argument (HW). □

Theorem 6.60 (Inverse Function Theorem) *Let $E \subseteq \mathbb{R}$ be an open interval and let $f : E \to \mathbb{R}$ be a bijection from E to $f(E)$. Let $a \in E$.*
If f is differentiable at $a \in E$ and $f'(a) \neq 0$, then f^{-1} is differentiable at $f(a)$, and

$$(f^{-1})'(f(a)) = \frac{1}{f'(a)}.$$

PROOF Apply Carathéodory's Theorem to f to obtain a function $\psi_f : \mathbb{R} \to \mathbb{R}$ s.t.

(a) ψ_f is continuous at a;
(b) $f(x) - f(a) = \psi_f(x)(x-a)$ for all $x \in \mathbb{R}$; and
(c) $\psi_f(a) = f'(a)$.

Claim: $\exists\, \delta > 0$ s.t. $\psi_f(x) \neq 0$ for all $x \in (a-\delta, a+\delta)$.

Proof of claim: Suppose not, then $\forall\, n\, \exists x_n \in \left(a - \frac{1}{n}, a + \frac{1}{n}\right)$ s.t. $\psi_f(x) = 0$.

Since $|x_n - a| < \frac{1}{n}$, by the Squeeze Theorem $x_n \to a$.
But ψ_f is cts, so $0 = \lim \psi_f(x_n) = \psi_f(\lim x_n) = \psi_f(a)$.
This is a contradiction, proving the claim.

Now, let $x \in (a-\delta, a+\delta) \cap E$ and let $y = f(x)$. Observe

$$\begin{aligned} y - f(a) = f(f^{-1}(y)) - f(a) &= \psi_f(f^{-1}(y))(f^{-1}(y) - a) \\ &= \psi_f(f^{-1}(y))(f^{-1}(y) - f^{-1}(f(a))). \end{aligned}$$

Since $\psi_f(f^{-1}(y)) = \psi_f(x) \neq 0$, we can divide through to get

$$\frac{1}{\psi_f(f^{-1}(y))}(y - f(a)) = f^{-1}(y) - f^{-1}(f(a))$$

Let $\psi_{f^{-1}}(y) = \frac{1}{\psi_f(f^{-1}(y))}$. Observe:

(a) $\psi_{f^{-1}}$ is the composition of cts functions at a, hence is cts at a; and
(b) $\psi_{f^{-1}}(y)(y - f(a)) = f^{-1}(y) - f^{-1}(f(a))$.

By Caratheodory's Theorem, f^{-1} is differentiable at $f(a)$ and

$$(f^{-1})'(f(a)) = \psi_{f^{-1}}(f(a)) = \frac{1}{f'(f^{-1}(f(a))} = \frac{1}{f'(a)}. \square$$

Natural logarithm

Definition 6.61 *The* **natural logarithm** *function* $\log : (0, \infty) \to \mathbb{R}$ *is the inverse of* $\exp$*:*

$$\log x = y \textit{ means } \exp y = x.$$

Remark on the notation: You are probably used to seeing the natural logarithm written "ln". Mathematicians don't write "ln" for logs. Usually, we write $\log$ for the natural logarithm.

Theorem 6.62 $\log : (0, \infty)$ *is differentiable and* $\log'(x) = \dfrac{1}{x}$.

PROOF Let $x \in (0, \infty)$ and let $y = \log x$. That means $\exp y = x$.

Apply the Inverse Function Theorem to get

$$\log'(x) = \frac{1}{\exp'(y)} = \frac{1}{\exp y} = \frac{1}{x}. \; \square$$

Log rules

Theorem 6.63 (Logarithm rules) *Let* $x, y \in (0, \infty)$ *and* $n \in \mathbb{N}$*. Then*

$$\log(xy) = \log x + \log y \qquad \log\left(\frac{x}{y}\right) = \log x - \log y \qquad \log(x^n) = n \log x.$$

PROOF Let $x, y \in \mathbb{R}$ and let $a = \log x$ and $b = \log y$. Thus $\exp a = x$ and $\exp b = y$.

For the first log rule, apply an exponent rule proven earlier to get

$$xy = \exp(a)\exp(b) = \exp(a + b) = \exp(\log x + \log y).$$

Take $\log$ of both sides to get the first log rule.

The other two rules are HW. $\square$

Exponentials and logarithms in other bases

Once you have natural logarithms, you can define exponentials and logarithms in other bases as follows:

Definition 6.64 *For any* $x \in \mathbb{R}$ *and any* $b > 0$*, define* $b^x = \exp(x \log b)$.
For any $x \in \mathbb{R}$ *and any* $b > 0$*, define* $\log_b x = \dfrac{\log x}{\log b}$.

From these definitions you can derive all the usual properties of exponentials and logs (some of these are in the HW).

6.12 Chapter 6 Summary

Definitions to know

Nouns

- The **derivative of f at a** is $f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$.
- The **(middle-thirds) Cantor set** is the set $\mathcal{C} \subseteq [0,1]$ consisting of numbers in $[0,1]$ that have a ternary expansion with no 1s.
- An **antiderivative** of f is another function $'f$ so that $('f)' = f$.
- The **indefinite integral** of f is the set of all its antiderivatives.
- (★) The n^{th} **Taylor polynomial of f centered at a** is

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

Adjectives that describe functions $f : \mathbb{R} \to \mathbb{R}$

- f is called **differentiable at** $a \in \mathbb{R}$ if $f'(a)$ exists (see above).
  f is called **differentiable on** E if it is differentiable at every $a \in E$.
  f is called **differentiable** if it is differentiable at every point in its domain.
- f is called **increasing** on (a,b) if $\forall\, x,y \in (a,b)$, $x \le y$ implies $f(x) \le f(y)$.
  f is called **strictly increasing** on (a,b) if $\forall\, x,y \in (a,b)$, $x < y$ implies $f(x) < f(y)$.
  f is called **decreasing** on (a,b) if $\forall\, x,y \in (a,b)$, $x \le y$ implies $f(x) \ge f(y)$.
  f is called **strictly decreasing** on (a,b) if $\forall\, x,y \in (a,b)$, $x < y$ implies $f(x) > f(y)$.
  f is called **monotone** on (a,b) if f is increasing on (a,b) or f is decreasing on (a,b).

Theorems with names

Differentiability implies continuity If f is differentiable at a, then f is continuous at a.

Carathéodory's Theorem f is differentiable at $a \Leftrightarrow \exists\, \psi : \mathbb{R} \to \mathbb{R}$ which is continuous at a, satisfies $f(x) - f(a) = \psi(x)(x-a)\ \forall\, x \in \mathbb{R}$, and has $\psi(a) = f'(a)$.

Fermat's Theorem If c is the location of an absolute extremum of differentiable function f, then $f'(c) = 0$.

Mean Value Theorem (MVT) Let f be continuous on $[a,b]$ and differentiable on (a,b). Then $\exists\, c \in (a,b)$ s.t. $f'(c) = \dfrac{f(b)-f(a)}{b-a}$.

Zero Derivative Theorem (ZDT) If $f'(x) = 0$ for all $x \in (a,b)$, then f is constant on (a,b).

Antiderivative Theorem Any two antiderivatives of the same function differ by at most a constant.

Darboux's Theorem If f is differentiable, then for any z between $f'(a)$ and $f'(b)$, there is $c \in [a,b]$ s.t. $f'(c) = z$.

Monotonicity Test If $f'(x) \geq 0$ on (a,b), then f is increasing on (a,b).

If $f'(x) > 0$ on (a,b), then f is strictly increasing on (a,b).

If $f'(x) \leq 0$ on (a,b), then f is decreasing on (a,b).

If $f'(x) < 0$ on (a,b), then f is strictly decreasing on (a,b).

Cauchy's Mean Value Theorem Let f, g be continuous on $[a,b]$ and differentiable on (a,b). Then, $\exists\, c \in (a,b)$ s.t. $f'(c)[g(b)-g(a)] = g'(c)[f(b)-f(a)]$.

L'Hôpital's Rule Let f, g be differentiable. If $f(a) = g(a) = 0$ and $\exists\, \eta > 0$ s.t. $g'(x) \neq 0$ for $x \in (a-\eta, a+\eta) - \{a\}$, then $\lim_{x\to a} \dfrac{f(x)}{g(x)} \overset{L}{=} \lim_{xtoa} \dfrac{f'(x)}{g'(x)}$.

(★) Taylor's Theorem If f is $n+1$-times differentiable on (α,β), then for any $a, x \in (\alpha,\beta)$, $\exists\, z$ between a and x so that $f(x) = P_n(x) + \dfrac{f^{(n+1)}(z)}{(n+1)!}(x-a)^{n+1}$.

(★) Interchange of Limit and Derivative If $f_n' \rightrightarrows g$ on $E = (\alpha,\beta)$ and $\exists\, a \in E$ s.t. $\{f_n(a)\}$ converges, then $\exists\, f$ s.t. $f_n \rightrightarrows f$ and $f' = g$.

Basic derivative rules

Constant Function Rule: $c' = 0$ for any constant $c \in \mathbb{R}$.

Identity Function Rule: $x' = 1$.

Power Rule: $\forall\, n \in \mathbb{N}$, $(x^n)' = nx^{n-1}$.

Reciprocal Rule: $\left(\dfrac{1}{x}\right)' = \dfrac{-1}{x^2}$.

Constant Multiple Rule: $(r\,f)' = f'$ for any $r \in \mathbb{R}$.

The rules below work under the assumption that f and g are differentiable:

Sum Rule: $(f+g)' = f' + g'$.

Difference Rule: $(f - g)' = f' - g'$.

Product Rule: $(fg)' = f'g + g'f$.

Quotient Rule: $\left(\frac{f}{g}\right)' = \frac{f'g - g'f}{g^2}$.

Chain Rule: $(f \circ g)' = (f' \circ g)\, g'$.

Other theorems to remember

- Alternate definitions of the derivative $f'(a)$ include

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} \qquad \text{and} \qquad \lim_{h \to 0} \frac{f(a) - f(a-h)}{h}.$$

- f is differentiable at a if and only if f is very well-approximated by a linear function, i.e. $\exists$ linear function L so that $\forall\, \epsilon > 0\, \exists\, \delta > 0$ s.t. $|x - a| < \delta$ implies $|f(x) - L(x)| \le \epsilon|x - a|$.
- If $x = .d_1d_2d_3d_4\cdots_{[b]}$, then $bx = d_1.d_2d_3d_4\cdots_{[b]}$.
- The Cantor set $\mathcal{C}$ is compact, uncountable, perfect and totally disconnected.
- (★) If f is n-times differentiable at a, then there is exactly one polynomial P_n of degree $\le n$ (namely, the n^{th} Taylor polynomial of f centered at a) which satisfies $P_n^{(k)}(a) = f^{(k)}(a)$ for all $k \in \{0, 1, 2, \dots, n\}$.

Facts about specific functions

- The **fractional part of** $x \in \mathbb{R}$ **is** $\{x\} = x - \lfloor x \rfloor$.
- Dirichlet's function $\mathbb{1}_{\mathbb{Q}}$ and Thomae's function τ are nowhere differentiable.
- For any $n \in \mathbb{N}$, $f(x) = \begin{cases} \sin\frac{1}{x^n} & x \neq 0 \\ 0 & x = 0 \end{cases}$ and $f(x) = \begin{cases} x\sin\frac{1}{x^n} & x \neq 0 \\ 0 & x = 0 \end{cases}$ are not differentiable at 0.
- For any $n \in \mathbb{N}$, $f(x) = \begin{cases} x^m\sin\frac{1}{x^n} & x \neq 0 \\ 0 & x = 0 \end{cases}$ is differentiable at 0 if $m > 1$.
- While $f(x) = \begin{cases} x^2\sin\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$ is differentiable at every $x \in \mathbb{R}$ (including $x = 0$), its derivative is not continuous at 0.
- The Cantor function is differentiable at every $a \notin \mathcal{C}$, and not differentiable at any $a \in \mathcal{C}$.

- (★) exp, sin and cos are differentiable, with respective derivatives exp, cos and $-\sin$.
- (★) $e^x = \exp(x)$ for every $x \in \mathbb{R}$ (this is a theorem for $x \in \mathbb{Q}$ and a definition for $x \in \mathbb{R} - \mathbb{Q}$).
- (★) exp is a strictly increasing bijection from $\mathbb{R}$ to $(0, \infty)$.
- (★) Exponent rules hold: $\exp(x+y) = \exp(x)\exp(y)$; $\exp(nx) = [\exp(x)]^n$; $\exp(x-y) = \dfrac{\exp(x)}{\exp(y)}$.
- (★) $\cos^2 x + \sin^2 x = 1$, meaning the point $(\cos x, \sin x)$ is on the unit circle in $\mathbb{R}^2$ for every x; conversely, for every point (p, q) on the unit circle, there is number x, unique up to multiples of 2π, so that $p = \cos x$ and $q = \sin x$.
- (★) $-1 \le \cos x \le 1$; $-1 \le \sin x \le 1$; $\cos(x + 2\pi) = \cos x$; $\sin(x + 2\pi) = \sin x$
- (★) $\log : (0, \infty) \to \mathbb{R}$ is the inverse of exp; log is a strictly increasing, differentiable bijection; the usual log rules hold; $\log'(x) = \dfrac{1}{x}$
- (★) $f(x) = \begin{cases} \exp\left(\dfrac{-1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$ is a function whose Taylor series centered at 0 converges for all x, but does not converge to f when $x \neq 0$.

Proof techniques

To prove that f is differentiable (or to compute f'), do one of these things:

1. Show that f is a sum/difference/product/composition of functions already known to be differentiable.
2. Use a definition and show that the limit that defines f' exists.
3. (★) Apply the interchange of limit and derivative (especially if f is defined as a power series).

To prove an inequality using the MVT, set one side equal to a constant, call the other side $f(x)$ and show that $f'(x)$ is either $\ge$ or $\le$ a constant. Then suppose not, and use the MVT to derive a contradiction.

To prove that f is not the derivative of another function, one option is to show that f does not satisfy the conclusion of Darboux's Theorem.

6.13 Chapter 6 Homework

Exercises from Section 6.1

1. Prove the second statement of Theorem 6.4, which says that if $\lim_{h\to 0} \frac{f(a)-f(a-h)}{h}$ exists, then $f:\mathbb{R}\to\mathbb{R}$ is differentiable at a and $\lim_{h\to 0} \frac{f(a)-f(a-h)}{h} = f'(a)$.

2. Prove the **Constant Function Rule** (Theorem 6.5), which says that if $f:\mathbb{R}\to\mathbb{R}$ is constant, then f is differentiable and $f'(x)=0$.

3. Prove the **Reciprocal Rule** (Theorem 6.7), which says that if $f:\mathbb{R}-\{0\}\to\mathbb{R}$ is $f(x)=\frac{1}{x}$, then f is differentiable at every $x\neq 0$ and $f'(x)=\frac{-1}{x^2}$.

4. Let $h:\mathbb{R}\to\mathbb{R}$ be defined by $h(x)=\begin{cases} x^2 & \text{if } x\in\mathbb{Q} \\ 0 & \text{else} \end{cases}$

   a) Show $h'(0)=0$.

   b) Show that if $x\neq 0$, then h is not differentiable at x. (One way to do this is to show that h is not continuous at x.)

5. Let $f:\mathbb{R}\to\mathbb{R}$ be $f(x)=x|x|$. Determine the numbers at which f is differentiable, and compute $f'(x)$ wherever it exists.

6. Let $f:\mathbb{R}\to\mathbb{R}$ be $f(x)=\sqrt[3]{x}$. Determine the numbers at which f is differentiable, and compute $f'(x)$ wherever it exists.

Exercises from Section 6.2

7. Prove that the Cantor set $\mathcal{C}$ is closed.

8. Prove that the Cantor set $\mathcal{C}$ is uncountable.

9. Prove that the Cantor set $\mathcal{C}$ is perfect, meaning that for every $x\in\mathcal{C}$ and every $\epsilon>0$, there is $y\in(B_\epsilon(x)\cap\mathcal{C})-\{x\}$.

10. Prove that the Cantor set $\mathcal{C}$ is totally disconnected, meaning that $\mathcal{C}$ does not contain any interval of positive length.

Exercises from Section 6.3

11. Prove the **Constant Multiple Rule**, which says that if $f:\mathbb{R}\to\mathbb{R}$ is differentiable at $a\in\mathbb{R}$, then for any constant $rin\mathbb{R}$ the function rf is differentiable at a and $(rf)'(a)=rf'(a)$.

12. Prove the **Sum Rule**, which says that if $f, g : \mathbb{R} \to \mathbb{R}$ are differentiable at $a \in \mathbb{R}$, then $f + g$ is differentiable at a and $(f + g)'(a) = f'(a) + g'(a)$.

13. Prove the **Difference Rule**, which says that if $f, g : \mathbb{R} \to \mathbb{R}$ are differentiable at $a \in \mathbb{R}$, then $f - g$ is differentiable at a and $(f - g)'(a) = f'(a) + g'(a)$.

Exercises from Section 6.4

14. Prove the second statement of Fermat's Theorem (Theorem 6.25), which says that if $f : [a, b] \to \mathbb{R}$ is differentiable and $c \in (a, b)$ is the location of the absolute minimum value of f on $[a, b]$, then $f'(c) = 0$.

Exercises from Section 6.5

15. Prove the Antiderivative Theorem (Theorem 6.29; the proof was started in the notes).

16. Prove the first two statements of the Monotonocity Test (Theorem 6.33), which say that if $f : \mathbb{R} \to \mathbb{R}$ is differentiable, then

    a) if $f'(x) \geq 0$ for $x \in (a, b)$, then f is increasing on (a, b); and
    b) if $f'(x) > 0$ for $x \in (a, b)$, then f is strictly increasing on (a, b).

17. Prove **Rolle's Theorem**, which says that if $f : \mathbb{R} \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) where $a < b$, then $\exists\, c \in (a, b)$ so that $f'(c) = 0$.

18. Give a proof of the Mean Value Theorem that assumes Rolle's Theorem.

    Hint: Apply Rolle's Theorem to the auxiliary function $g(x) = f(a) + \dfrac{f(b) - f(a)}{b - a}(x - a)$.

19. Let $g : \mathbb{R} \to \mathbb{R}$ be defined by

$$g(x) = \begin{cases} x + x^2 \sin\left(\dfrac{1}{x}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

    Prove that $g'(0) = 1$ (an example from the notes may be helpful), but prove that even though $g'(0) > 0$ the function g is not increasing on any open interval containing 0.

20. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. For each $x \in X$, consider the recursively defined sequence $\{a_n\}$ defined by setting $a_0 = x$ and $a_n = f(a_{n-1})$ for all $n > 0$. $\{a_n\}$ is called the **forward orbit of x under f**.

Suppose f is differentiable. Let $c \in \mathbb{R}$ be a fixed point of f (this means $f(c) = c$). Suppose f' is continuous at c and that $|f'(c)| < 1$. Prove that there exists an open set E containing c such that for all $x \in E$, $a_n \to c$ (where $\{a_n\}$ is the forward orbit of x under f).

Exercises from Section 6.7

21. In the notes, I mentioned that given $a \in \mathbb{R}$, any polynomial p of degree n can be written as

$$p(x) = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_n(x-a)^n = \sum_{k=0}^{n} a_k(x-a)^k$$

for suitably chosen constants $a_0, ..., a_n$. To illustrate why this is true, let's do an example: write the polynomial $f(x) = x^4 - x$ in the form $\sum_{k=0}^{4} a_k(x-2)^k$ for constants $a_0, a_1, ..., a_4$.

Hint: This is basically an algebra problem.

22. Prove that for any cubic polynomial f, $P_3 = f$.

Remark: It is true that for any polynomial f of degree n, $P_n = f$, but we'll only prove this when $n = 3$. To do this, write down a general form of a cubic polynomial and work out its third Taylor polynomial; you can show it simplifies to what you started with.

23. In Calculus I you learn something called the **Second Derivative Test**, which says the following:

Let $E \subseteq \mathbb{R}$ be open and let $f : E \to \mathbb{R}$ be a twice-differentiable function. If $c \in E$ is such that $f'(c) = 0$ and f'' is continuous at c, then:

- if $f''(c) > 0$, then c is a local minimum of f;
- if $f''(c) < 0$, then c is a local maximum of f.

Prove the Second Derivative Test.

Hints: Start by assuming that $f''(c) > 0$. First, use the fact that f'' is continuous at c to find a $\delta > 0$ such that $f'' > 0$ on $(c-\delta, c+\delta)$. Next, let $x \in (c-\delta, c+\delta)$ and use Taylor's Theorem with $n = 1$ to show that $f(x) \geq f(c)$ (i.e. that c is a local minimum).

Then, assume that $f''(c) < 0$. Apply the previous case to $-f$ to get show that c is a local minimum of $-f$, which implies c is a local maximum of f.

24. A function $f : \mathbb{R} \to \mathbb{R}$ is called **convex** (a.k.a. **concave up**) if given any two points on the graph of f, no part of the line segment connecting those two points lies below the graph of f.

a) Give a precise definition of what it means for f to be convex. Your definition should encapsulate the idea described above.

b) Prove that if f is differentiable, then f is convex if and only if $f(x) \geq L(x)$ for any function L which is a tangent line of f.

Hint: Taylor's theorem may be useful.

25. a) Prove that if $f : \mathbb{R} \to \mathbb{R}$ is twice-differentiable, then for every $a \in \mathbb{R}$,

$$f''(a) = \lim_{h \to 0} \frac{f(a+h) - 2f(a) + f(a-h)}{h^2}.$$

b) Prove that if f is twice-differentiable, then f is convex if and only if $f''(x) \geq 0$ for all x (in other words, for twice-differentiable functions, "convex" is a synonym of "concave up").

Exercises from Section 6.8

26. Let $f_n : (0,2) \to \mathbb{R}$ be $f_n(x) = \begin{cases} \dfrac{x^n}{n} & x < 1 \\ x + \dfrac{1}{n} - 1 & x \geq 1 \end{cases}$ be the sequence of functions from Question 2 of Section 6.8 in the notes. Prove $f_n \rightrightarrows f$ on $(0,2)$, where $f(x) = \begin{cases} 0 & x < 1 \\ x - 1 & x \geq 1 \end{cases}$.

27. Prove $\sin : \mathbb{R} \to \mathbb{R}$ and $\cos : \mathbb{R} \to \mathbb{R}$ are differentiable functions, with $\sin' = \cos$ and $\cos' = -\sin$.

28. Prove that for the function $f(x) = \begin{cases} \exp\left(\dfrac{-1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$, $f^{(n)}(x) = 0$ for all $n \geq 0$.

Exercises from Section 6.10

29. Prove Lemma 6.55, which says that if $P_N(x)$ is the N^{th} Taylor polynomial for $\cos x$ centered at a, then $P_N(x) \to \cos x$ on $\mathbb{R}$.

30. Use the MVT to prove that $|\sin x - \sin y| \leq |x - y|$ for all $x, y \in \mathbb{R}$.

Exercises from Section 6.11

31. Finish the proof of the Continuous Inverse Theorem (Theorem **??**) by showing that if f is continuous and strictly decreasing, then $f(a,b) = (f(b), f(a))$.

32. Prove that for any $x, y \in (0, \infty)$, $\log\left(\dfrac{x}{y}\right) = \log x - \log y$.

33. Prove that for any $x \in (0, \infty)$ and any $n \in \mathbb{N}$, $\log(x^n) = n \log x$.

34. Let $b > 0$. Prove that the functions $f : \mathbb{R} \to (0, \infty)$ and $g : (0, \infty) \to \mathbb{R}$ defined by $f(x) = b^x$ and $g(x) = \log_b x$ are inverses.

35. Let $b > 0$. Prove $\log_b(xy) = \log_b x + \log_b y$ for any $x, y \in (0, \infty)$.

36. Let $b > 0$. Prove the function $f(x) = b^x$ is differentiable; compute and simplify its derivative.

37. Prove $\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e$.

    Hint: Rewrite $\left(1 + \frac{1}{n}\right)^n$ using the definition of exponential function, then rearrange this rewritten form so that you can use L'Hôpital's Rule.

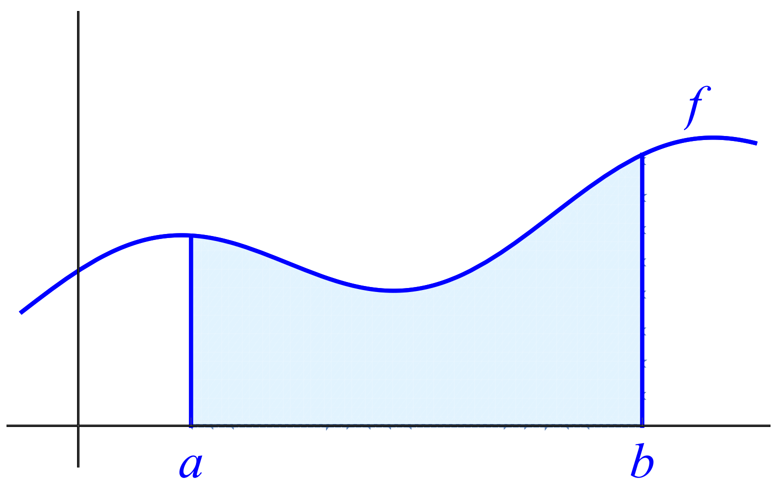
Chapter 7

Riemann integration

7.1 Definition of the Riemann integral

Motivation

In Calculus 1, integrals are developed to compute [ ]:



To do this, you approximate the area under f by computing the area of some rectangles, as shown above. We need appropriate notation for this, so we can derive the theory of integration rigorously.

Definition 7.1 *Let $[a, b] \subseteq \mathbb{R}$ be a closed, bounded interval with $a < b$.*

A **partition** *$\mathcal{P}$ of $[a, b]$ is a finite list of numbers $\{x_0, x_1, \ldots, x_n\}$ where*

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b.$$

The **size** *of the partition is n (even though the partition has $n + 1$ numbers in it).*

Given a partition $\mathcal{P} = \{x_0, \ldots, x_n\}$ of $[a, b]$, for each $k \in \{1, 2, \ldots, n\}$ we define the k^{th} **subinterval** *of $\mathcal{P}$ to be $[x_{k-1}, x_k]$.*

We define the **width** *of the k^{th} subinterval to be $\Delta x_k = x_k - x_{k-1}$.*

The **norm** *of $\mathcal{P}$, denoted $||\mathcal{P}||$, is the largest width of any subinterval, i.e.*

$$||\mathcal{P}|| = \max\{\Delta x_k : 1 \leq k \leq n\}.$$

EXAMPLE 1

Let $\mathcal{P} = \{0, 2, 5, 9, 10\}$. (This is a partition of $[0, 10]$.)

Keep in mind: For any partition $\mathcal{P}$ of $[a, b]$, $\sum_{k=1}^{n} \Delta x_k =$

Definition 7.2 *Let $\mathcal{P} = \{x_0, ..., x_n\}$ be a partition of $[a, b]$.*

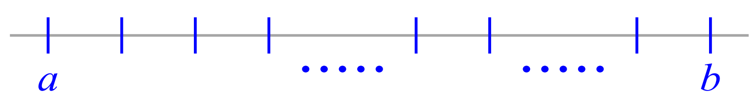
A set $\{c_1, c_2, ..., c_n\}$ of numbers is called a list of **test points** *for $\mathcal{P}$ if $c_k \in [x_{k-1}, x_k]$ for all $k \in \{1, 2, ..., n\}$.*

A partition, together with a list of test points for that partition, is called a **tagged partition** *of $[a, b]$. We denote tagged partitions by $\widehat{P} = \{x_0, ..., x_n\}; \{c_1, ..., c_n\}$.*

Given a partition $\mathcal{P} = \{x_0, ..., x_n\}$ of $[a, b]$, choosing test points $c_k \in [x_{k-1}, x_k]$ for each k is called **tagging** *$\mathcal{P}$ to create $\widehat{\mathcal{P}}$.*

EXAMPLE 2

Consider the tagged partition of $[a, b]$ into n equal-length subintervals, where the test points are the left endpoint of each subinterval.



We'll need this technical result later, which guarantees partitions of any interval of arbitrarily small norm:

Lemma 7.3 *Let $[a, b]$ be a closed, bounded interval with $a < b$. Given any $\epsilon > 0$, there is a partition $\mathcal{P}$ (and therefore also a tagged partition $\widehat{\mathcal{P}}$) with $\|\mathcal{P}\| < \epsilon$.*

PROOF Given $\epsilon > 0$, let $n > \dfrac{b-a}{\epsilon}$.

Consider the partition $\mathcal{P}$ of $[a, b]$ into n equal-length subintervals.

Each subinterval of $\mathcal{P}$ has width $\dfrac{b-a}{n} < \epsilon$, so $\|\mathcal{P}\| < \epsilon$.

If we need a tagged partition, choose test points for $\mathcal{P}$ arbitrarily. □

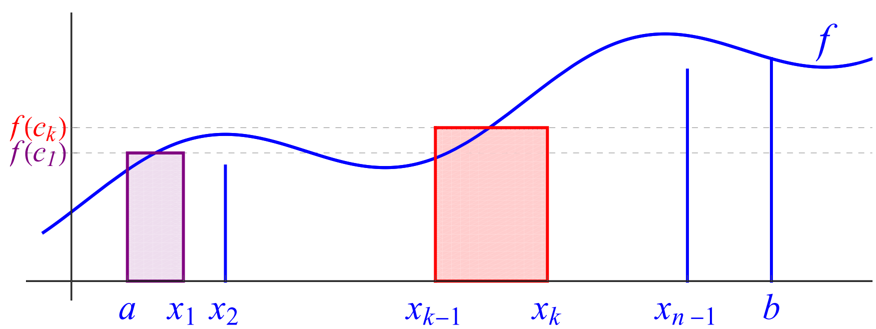
Riemann sums

Definition 7.4 *Let $a < b$ and suppose $f : [a, b] \to \mathbb{R}$.*

Given a tagged partition $\widehat{\mathcal{P}} = \{x_0, ..., x_n\}; \{c_1, ..., c_n\}$ of $[a, b]$, the **Riemann sum (for f) (associated to $\widehat{\mathcal{P}}$)** *is the number*

$$RS(f; \widehat{\mathcal{P}}) = \sum_{k=1}^{n} f(c_k)\Delta x_k.$$

Given an untagged partition $\mathcal{P} = \{x_0, ..., x_n\}$, **a Riemann sum (for f) associated to** *$\mathcal{P}$ is any Riemann sum associated to a tagged partition coming from $\mathcal{P}$, i.e. associated to some choice of test points for $\mathcal{P}$.*



Theorem 7.5 (Riemann sums are linear) *Let $a < b$ and suppose $f, g : [a, b] \to \mathbb{R}$. For any tagged partition $\widehat{\mathcal{P}} = \{x_0, ..., x_n\}; \{c_1, ..., c_n\}$ of $[a, b]$,*

1. *$RS(rf; \widehat{\mathcal{P}}) = r\, RS(f; \widehat{\mathcal{P}})$ for any constant $r \in \mathbb{R}$.*
2. *$RS(f + g; \widehat{\mathcal{P}}) = RS(f; \widehat{\mathcal{P}}) + RS(g; \widehat{\mathcal{P}})$.*

PROOF We prove statement (1) here.

$$RS(rf; \widehat{\mathcal{P}}) = \sum_{k=1}^{n} (rf)(c_k)\, \Delta x_k =$$

Statement 2 is a HW problem. □

Definition of the integral

Definition 7.6 *Let $a < b$ and suppose $f : [a, b] \to \mathbb{R}$.*
We say f is **(Riemann) integrable on** $[a, b]$ *if there is a real number, denoted*

$$\int_a^b f(x)\,dx \quad \text{or just} \quad \int_a^b f,$$

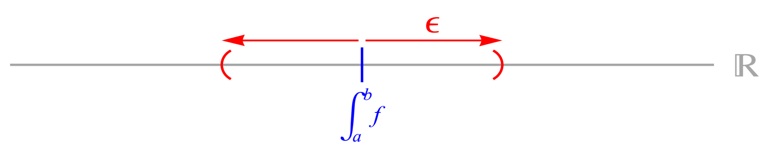
and called the **Riemann integral of f from a to b**, *such that $\forall\, \epsilon > 0\ \exists\, \delta > 0$ such that if $\widehat{\mathcal{P}}$ is any tagged partition of $[a, b]$ with $\|\widehat{\mathcal{P}}\| < \delta$, then*

$$\left| RS(f; \widehat{\mathcal{P}}) - \int_a^b f \right| < \epsilon.$$

Concept

This is another ϵ definition (like the definition of the limit of a sequence or the definition of open set or the definition of the limit of a function).

Here, the idea is that if $I = \displaystyle\int_a^b f$, then given any $\epsilon > 0$, the Riemann sums you get for f are always within ϵ of I, **no matter how you pick the test points**, and **no matter which partition you choose**, **if** the norm of the partition is small enough (where "small enough" means less than δ, which is allowed to depend on ϵ).

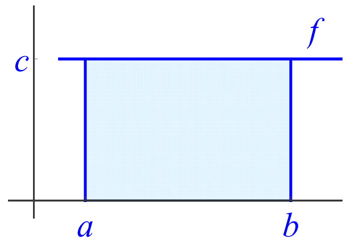


Drawbacks

- The definition of limit of a sequence gives you a decent way to check whether or not $a_n \to L$. But it doesn't tell you what L is (you have to guess L or figure L out some other way).
- Similarly, this definition of Riemann integral gives you a decent way to check whether or not some number $\displaystyle\int_a^b f$ is the integral of f from a to b. But it doesn't give you insight into how to find the number $\displaystyle\int_a^b f$.

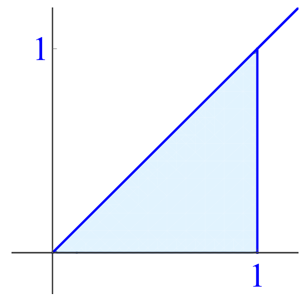
Example 3

Let $f : \mathbb{R} \to \mathbb{R}$ be a constant function $f(x) = c$. Prove f is integrable on $[a, b]$ for any $a < b$.



EXAMPLE 4

Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = x$. Prove f is integrable on $[0, 1]$.



How we will get around this dilemma

We will consider the biggest (more precisely, the $\sup$) and smallest (the $\inf$) values that Riemann sums associated to a partition $\mathcal{P}$ can take. This will take care of all of them.

Problem

The "biggest" Riemann sum associated to $\mathcal{P}$ may not actually be a Riemann sum.

Spoiler alert

We will eventually develop a much better way of studying this example.

Theorem 7.7 *If f is Riemann integrable on $[a, b]$, then $\int_a^b f$ is unique (there cannot be two different values for the integral).*

PROOF Suppose not, i.e. that $L = \int_a^b f$ and $M = \int_a^b f$, where $L \neq M$.

WLOG ______.

Now let $\epsilon =$

By def'n of Riemann integral, $\exists\, \delta_L > 0$ s.t. for any tagged partition $\widehat{\mathcal{P}}$ of $[a, b]$,

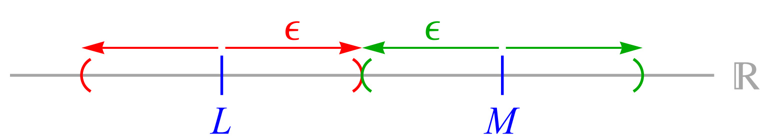
$$\|\widehat{\mathcal{P}}\| < \delta_L \text{ implies } \left|RS(f; \widehat{\mathcal{P}}) - L\right| < \epsilon.$$

Similarly, $\exists\, \delta_M > 0$ s.t. for any tagged partition $\widehat{\mathcal{P}}$ of $[a, b]$,

$$\|\widehat{\mathcal{P}}\| < \delta_M \text{ implies } \left|RS(f; \widehat{\mathcal{P}}) - M\right| < \epsilon.$$

Let $\delta =$ ______ . For any tagged partition $\widehat{\mathcal{P}}$ of $[a, b]$ with $\|\widehat{\mathcal{P}}\| < \delta$,

$$\left|RS(f; \widehat{\mathcal{P}}) - L\right| < \epsilon \quad \text{and} \quad \left|RS(f; \widehat{\mathcal{P}}) - M\right| < \epsilon.$$



So by the Triangle Inequality,

$$|L - M| \leq \left|L - RS(f; \widehat{\mathcal{P}})\right| + \left|RS(f; \widehat{\mathcal{P}}) - M\right| < \epsilon + \epsilon = 2\epsilon =$$

This is impossible. Therefore $L = M$, as wanted. □

Theorem 7.8 (Integrable functions are bounded) *Let $f : [a, b] \to \mathbb{R}$. If f is Riemann integrable on $[a, b]$, then f is bounded on $[a, b]$.*

PROOF We prove the contrapositive.

Toward that end, suppose that f is not bounded on $[a, b]$.

For now, let $I = \int_a^b f$ (we will show later that such an I cannot exist).

Using $\epsilon = 1$ in the def'n of the Riemann integral, $\exists\, \delta > 0$ so that

$$\|\widehat{\mathcal{P}}\| < \delta \text{ implies } \left|RS(f; \widehat{\mathcal{P}}) - I\right| < 1 \Rightarrow \qquad\qquad .$$

Now, let $\mathcal{Q}$ be any partition (not tagged yet) of $[a, b]$ with $\|\mathcal{Q}\| < \delta$.

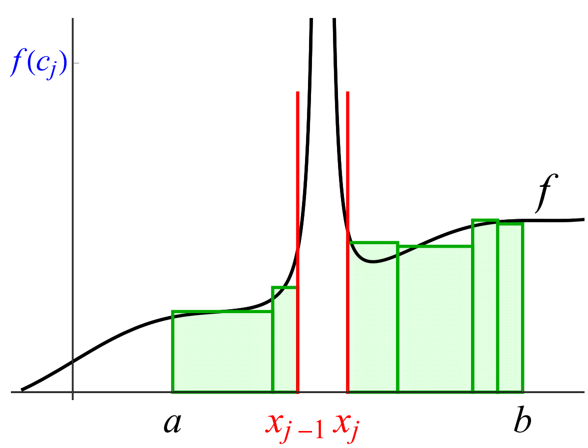
Since f is unbounded, so is $|f|$, $\exists\, j$ s.t. $|f|$ is unbounded on the j^{th} subinterval $[x_{j-1}, x_j]$.

To tag the partition $\mathcal{Q}$, first choose all the test points $\{c_1, c_2, \ldots, c_{j-1}, c_{j+1}, \ldots, c_n\}$ but the j^{th} one arbitrarily.

Then, choose the j^{th} test point c_j so that

$$|f(c_j)| > \frac{|I| + 1 + \left|\sum\limits_{k=1, k\neq j}^{n} f(c_k)\Delta x_k\right|}{\Delta x_j}. \tag{7.1}$$

(This can be done since $|f|$ is unbounded on the j^{th} subinterval.)



Now

$$\begin{aligned}|RS(f;\widehat{\mathcal{Q}})| &= \left|\sum_{k=1,k\neq j}^{n} f(c_k)\Delta x_k + f(c_j)\Delta x_j\right| \\ &\geq |f(c_j)\Delta x_j| - \left|\sum_{k=1,k\neq j}^{n} f(c_k)\Delta x_k\right|\end{aligned}$$

(by (7.1)

$$\begin{aligned}&> \left|\frac{|I|+1+\left|\sum_{k=1,k\neq j}^{n} f(c_k)\Delta x_k\right|}{\Delta x_j}\Delta x_j\right| - \left|\sum_{k=1,k\neq j}^{n} f(c_k)\Delta x_k\right| \\ &= |I|+1+\left|\sum_{k=1,k\neq j}^{n} f(c_k)\Delta x_k\right| - \left|\sum_{k=1,k\neq j}^{n} f(c_k)\Delta x_k\right| \\ &= |I|+1,\end{aligned}$$

so $\left|RS(f;\widehat{\mathcal{Q}}) - I\right| \geq 1 = \epsilon$.

This is a contradiction to $I = \int_a^b f$, so f cannot be integrable on $[a,b]$. □

EXAMPLE 5

The function $f:\mathbb{R}\to\mathbb{R}$ defined by $f(x) = \begin{cases} \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$ is not integrable on $[0,1]$, since it is not bounded on $[0,1]$.

QUESTIONS WE WANT TO EVENTUALLY ADDRESS

1. Is there a bounded function that is not integrable?
2. Can you actually get an expression for $\int_a^b f$ symbolically, in terms of f?

7.2 Upper and lower Riemann sums

Definition 7.9 *Let* $a < b$ *and suppose* $f : [a, b] \to \mathbb{R}$ *is bounded. Let* $\mathcal{P} = \{x_0, \ldots, x_n\}$ *be a partition of* $[a, b]$.

- *The* **upper (Riemann) sum (of** f**) associated to** $\mathcal{P}$ *is*

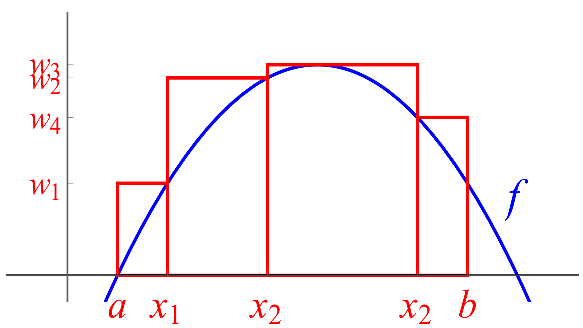
$$\mathcal{U}(f; \mathcal{P}) = \sum_{k=1}^{n} w_k \, \Delta x_k$$

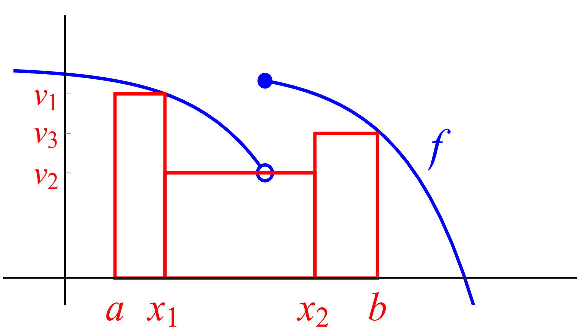
where $w_k = \sup\{f(x) : x \in [x_{k-1}, x_k]\}$.

- *The* **lower (Riemann) sum (of** f**) associated to** $\mathcal{P}$ *is*

$$\mathcal{L}(f; \mathcal{P}) = \sum_{k=1}^{n} v_j \, \Delta x_j$$

where $v_k = \inf\{f(x) : x \in [x_{k-1}, x_k]\}$.





Note: The upper and lower "Riemann" sums are not necessarily Riemann sums, because there may not be $c_j \in [x_{j-1}, x_j]$ such that $f(c_j) = v_j$ or w_j (see the lower sum picture on the previous page).

On the other hand, if f is continuous, then $\mathcal{U}(f;\mathcal{P})$ and $\mathcal{L}(f;\mathcal{P})$ are Riemann sums by the Max-Min Existence Theorem, because in this case f achieves its maximum and minimum on every compact interval like $[x_{j-1}, x_j]$.

One more note: There are no tags (test points) needed to define an upper or lower Riemann sum associated to a partition. In the long run, this will be an advantage of thinking about the integral in terms of upper and lower sums, as opposed to general Riemann sums.

A remark on the notation: Throughout this chapter, if $\mathcal{P}$ is a partition of $[a, b]$ and $f : [a, b] \to \mathbb{R}$ is bounded, then by "v_j" and "w_j" we mean

$$v_j = \inf\{f(x) : x \in [x_{j-1}, x_j]\} \text{ and } w_j = \sup\{f(x) : x \in [x_{j-1}, x_j]\}.$$

So we (and you) can use v_j and w_j in this chapter without defining them again.

If we need to specify the function or the partition from which the v_j and/or w_j come from, we'll use superscripts, like this:

different functions: v_j^f versus v_j^g

different partitions: $w_j^{\mathcal{P}}$ versus $w_j^{\mathcal{Q}}$

Lemma 7.10 *Let $f : [a, b] \to \mathbb{R}$ be bounded and let $\mathcal{P} = \{x_0, \ldots, x_n\}$ be a partition of $[a, b]$. Then, if we tag $\mathcal{P}$ with test points $\{c_1, \ldots, c_n\}$ to create $\widehat{\mathcal{P}}$, we have*

$$\mathcal{L}(f; \mathcal{P}) \le RS(f; \widehat{\mathcal{P}}) \le \mathcal{U}(f; \mathcal{P}).$$

PROOF By definition of infimum and supremum, we have $v_k \le f(c_k) \le w_k$ for all k.

Thus
$$L(f; \mathcal{P}) = \sum_{k=1}^{n} v_k \Delta x_k \le \sum_{k=1}^{n} f(c_k) \Delta x_k = RS(f; \widehat{\mathcal{P}})$$

and
$$RS(f; \widehat{\mathcal{P}}) = \sum_{k=1}^{n} f(c_k) \Delta x_k \le \sum_{k=1}^{n} w_k \Delta x_k \le U(f; \mathcal{P}). \square$$

The next result shows that you can get tag a partition to produce a Riemann sum that is arbitrarily close to its upper and lower sum:

Lemma 7.11 (Approximation of upper/lower sums by Riemann sums) *Let $a < b$ and suppose $f : [a, b] \to \mathbb{R}$ is bounded. Let $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ be any partition of $[a, b]$.*

1. *Given any $\epsilon > 0$, $\mathcal{P}$ can be tagged with test points to create $\widehat{\mathcal{P}}$ so that*
$$RS(f; \widehat{\mathcal{P}}) - \mathcal{L}(f; \mathcal{P}) < \epsilon.$$
2. *Given any $\epsilon > 0$, $\mathcal{P}$ can be tagged with test points to create $\widehat{\mathcal{P}}$ so that*
$$\mathcal{U}(f; \mathcal{P}) - RS(f; \widehat{\mathcal{P}}) < \epsilon.$$

PROOF We prove the first statement here.

Denote $\mathcal{P}$ as $\{x_0, \ldots, x_n\}$ and let $\epsilon > 0$.

Recall
$$v_j = \inf\{f(x) : x \in [x_{j-1}, x_j]\}.$$

By a characterization of infimum (Chapter 2), for each j there is c_j such that
$$f(c_j) < v_j + \frac{\epsilon}{b-a}$$

which implies
$$f(c_j) - v_j < \frac{\epsilon}{b-a}.$$

Tag $\mathcal{P}$ with the $\{c_j\}$ to make $\widehat{\mathcal{P}}$. Now

$$\begin{aligned} RS(f;\widehat{\mathcal{P}}) - \mathcal{L}(f;\mathcal{P}) &= \sum_{j=1}^{n} f(c_j)\,\Delta x_j - \sum_{j=1}^{n} v_j\,\Delta x_j \\ &= \sum_{j=1}^{n} \left[f(c_j) - v_j\right]\,\Delta x_j \\ &< \sum_{j=1}^{n} \frac{\epsilon}{b-a}\,\Delta x_j \\ &= \frac{\epsilon}{b-a}\sum_{j=1}^{n} \Delta x_j \\ &= \frac{\epsilon}{b-a}(b-a) \\ &= \epsilon. \end{aligned}$$

The second statement is similar and left as HW. □

MORE QUESTIONS WE WANT TO ADDRESS

Fix a bounded function $f : [a,b] \to \mathbb{R}$.

Lemma 7.10 shows that for any fixed partition $\mathcal{P}$, the lower sum associated to $\mathcal{P}$ is $\leq$ the upper sum associated to $\mathcal{P}$.

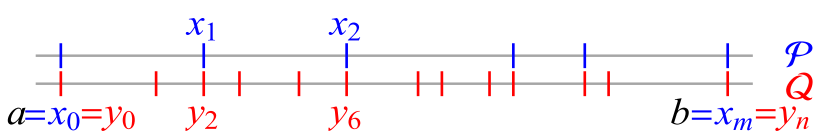
1. What if you took two different partitions $\mathcal{P}$ and $\mathcal{Q}$? Is the lower sum associated to $\mathcal{P}$ necessarily $\leq$ the upper sum associated to $\mathcal{Q}$?
2. More generally, how can we construct arguments that take into account multiple partitions at once?

7.3 Refinements and joins

Definition 7.12 *Let $\mathcal{P} = \{x_0, \ldots, x_m\}$ and $\mathcal{Q} = \{y_0, \ldots, y_n\}$ be two partitions of $[a, b]$. We say $\mathcal{Q}$ is a* **refinement** *of $\mathcal{P}$, and write $\mathcal{Q} \geq \mathcal{P}$, if either of the following equivalent conditions hold:*

- *$\mathcal{Q} \supseteq \mathcal{P}$ as sets, i.e. $\{x_0, \ldots, x_m\} \subseteq \{y_0, \ldots, y_n\}$;*
- *every subinterval of $\mathcal{Q}$ is contained in a single subinterval of $\mathcal{P}$.*

I think it is clear that these two conditions are equivalent, but if you don't believe me, here's a picture:



EXAMPLES AND NON-EXAMPLES

- $\mathcal{P} = \{0, 1, 5, 10\}$
  $\mathcal{Q} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$
- $\mathcal{P} = \{0, 2, 5\}$
  $\mathcal{Q} = \{0, 2, 5, 6\}$
- $\mathcal{P}$ = partition of $[a, b]$ into n equal-length subintervals;
  $\mathcal{Q}$ = partition of $[a, b]$ into m equal length subintervals.

Lemma 7.13 *If $\mathcal{Q} \geq \mathcal{P}$, then $\|\mathcal{Q}\| \leq \|\mathcal{P}\|$.*

PROOF Suppose $\mathcal{Q} \geq \mathcal{P}$.
By definition of norm, there is a subinterval I of $\mathcal{Q}$ with length $\|\mathcal{Q}\|$.
Since $\mathcal{Q} \geq \mathcal{P}$, I is contained in a single subinterval J of $\mathcal{P}$, so that subinterval J must have length at least $\|\mathcal{Q}\|$.
But the length of J is at most $\|\mathcal{P}\|$, so $\|\mathcal{Q}\| \leq \text{length}(J) \leq \|\mathcal{P}\|$ as wanted. □

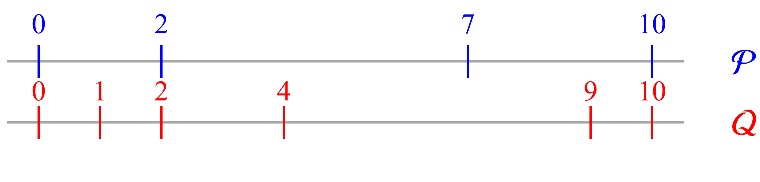
Definition 7.14 *Let $\mathcal{P} = \{x_0, \ldots, x_m\}$ and $\mathcal{Q} = \{y_0, \ldots, y_n\}$ be two partitions of $[a, b]$. Define the partition $\mathcal{P} \vee \mathcal{Q}$, called the* **least common refinement** *of $\mathcal{P}$ and $\mathcal{Q}$, also called the* **join** *of $\mathcal{P}$ and $\mathcal{Q}$, to be*

$$\mathcal{P} \vee \mathcal{Q} = \mathcal{P} \bigcup \mathcal{Q} = \{x_0, \ldots, x_m, y_0, \ldots, y_n\}$$

where duplicate numbers are removed and the remaining numbers are rewritten in increasing order as, say, $\{z_0, \ldots, z_p\}$.

EXAMPLE

Let $\mathcal{P} = \{0, 2, 7, 10\}$ and let $\mathcal{Q} = \{0, 1, 2, 4, 9, 10\}$.



Lemma 7.15 *Let $\mathcal{P}$ and $\mathcal{Q}$ be two partitions of $[a, b]$. Then $\mathcal{P} \vee \mathcal{Q} \geq \mathcal{P}$ and $\mathcal{P} \vee \mathcal{Q} \geq \mathcal{Q}$.*

PROOF This is immediate, since it is effectively a restatement of the two facts

$$\mathcal{P} \subseteq \mathcal{P} \bigcup \mathcal{Q} \text{ and } \mathcal{Q} \subseteq \mathcal{P} \bigcup \mathcal{Q}. \ \square$$

Using joins, we can show that any lower sum is $\leq$ any upper sum. This is shown in the next two results.

> **Theorem 7.16 (Refining makes upper sum smaller and lower sum bigger)** *Let $f:[a,b]\to\mathbb{R}$ be bounded and let $\mathcal{P}$ and $\mathcal{Q}$ be partitions of $[a,b]$. If $\mathcal{Q}\geq\mathcal{P}$, then*
>
> $$\mathcal{L}(f;\mathcal{P})\leq\mathcal{L}(f;\mathcal{Q}) \quad \text{and} \quad \mathcal{U}(f;\mathcal{P})\geq\mathcal{U}(f;\mathcal{Q}).$$

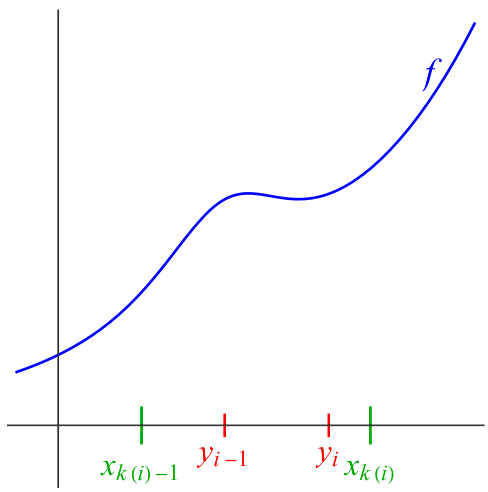
PROOF We prove the first inequality here. Let

$$\mathcal{P}=\{x_0,x_1,\ldots,x_{n_{\mathcal{P}}}\} \text{ and } \mathcal{Q}=\{y_0,y_1,\ldots,y_{n_{\mathcal{Q}}}\}.$$

Since $\mathcal{Q}\geq\mathcal{P}$, for every subinterval $[y_{i-1},y_i]$ of $\mathcal{Q}$, there is a subinterval $[x_{k(i)-1},x_{k(i)}]$ of $\mathcal{P}$ which contains it. Therefore, for each $i\in\{1,2,\ldots,n_{\mathcal{Q}}\}$,

$$v_i^{\mathcal{Q}}=\inf\{f(x):x\in[y_{i-1},y_i]\}\geq\inf\{f(x):x\in[x_{k(i)-1},x_{k(i)}]\}=v_{k(i)}^{\mathcal{P}}. \tag{7.2}$$

because there are more x's in $[x_{j(i)-1},x_{j(i)}]$ than in $[y_{i-1},y_i]$, hence more possible places for the function f to be small.



Now

$$\begin{aligned}
\mathcal{L}(f;\mathcal{Q}) &= \sum_{i=1}^{n_{\mathcal{Q}}} v_i^{\mathcal{Q}}\,\Delta y_i \\
&= \sum_{k=1}^{n_{\mathcal{P}}}\left(\sum_{\{i:k(i)=k\}} v_i^{\mathcal{Q}}\,\Delta y_i\right) \\
&\geq \sum_{k=1}^{n_{\mathcal{P}}}\left(\sum_{\{i:k(i)=k\}} v_k^{\mathcal{P}}\,\Delta y_i\right) \qquad \text{(from (7.2) above)} \\
&= \sum_{k=1}^{n_{\mathcal{P}}}\left(v_k^{\mathcal{P}}\sum_{\{i:k(i)=k\}} \Delta y_i\right) \\
&= \sum_{k=1}^{n_{\mathcal{P}}}\left(v_k^{\mathcal{P}}\Delta x_k\right)=\mathcal{L}(f;\mathcal{P}).
\end{aligned}$$

The inequality involving upper sums is similar and left as HW. □

Theorem 7.17 (Any lower sum is at most any upper sum) *Let $f : [a, b] \to \mathbb{R}$ be bounded and let $\mathcal{P} = \{x_0, \ldots, x_n\}$ and $\mathcal{Q} = \{y_0, \ldots, y_n\}$ be partitions of $[a, b]$. Then $\mathcal{L}(f; \mathcal{P}) \leq \mathcal{U}(f; \mathcal{Q})$.*

PROOF Let $\mathcal{R} = \mathcal{P} \vee \mathcal{Q}$. Then $\mathcal{L}(f; \mathcal{P}) \leq \mathcal{L}(f; \mathcal{R}) \leq \mathcal{U}(f; \mathcal{R}) \leq \mathcal{U}(f; \mathcal{Q})$. □

7.4 Integrability criteria

In this section, we take aim at one of our earlier questions: can you get an expression for $\int_a^b f$ symbolically, in terms of f?

Along the way, we will come up with a nice criterion that can be used to determine whether a function f is integrable on $[a, b]$.

Theorem 7.18 (Integrability criteria) *Let $a < b$ and suppose $f : [a, b] \to \mathbb{R}$ is bounded. Then, the following are equivalent:*

1. *f is Riemann integrable on $[a, b]$;*
2. *$\forall \epsilon > 0$, $\exists \delta > 0$ such that if $\mathcal{P}$ is any partition of $[a, b]$ with $\|\mathcal{P}\| < \delta$, then $\mathcal{U}(f; \mathcal{P}) - \mathcal{L}(f; \mathcal{P}) < \epsilon$;*
3. *$\forall \epsilon > 0$, $\exists$ partition $\mathcal{P}$ of $[a, b]$ with $\mathcal{U}(f; \mathcal{P}) - \mathcal{L}(f; \mathcal{P}) < \epsilon$;*
4. $\sup\{\mathcal{L}(f; \mathcal{P}) : \mathcal{P}$ *is a part. of* $[a, b]\} = \inf\{\mathcal{U}(f; \mathcal{P}) : \mathcal{P}$ *is a part. of* $[a, b]\}$

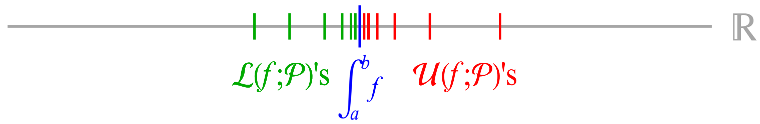
When statement (4) holds, both quantities in statement (4) are equal to $\int_a^b f$.

How to interpret these statements:

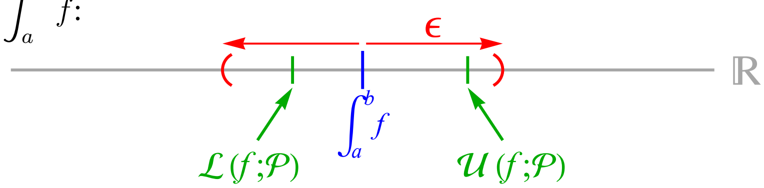
- Statement (3) gives a way (which is in most cases the best way) to check whether a function f is integrable on $[a, b]$, without having to actually compute a potential value of $\int_a^b f$:

- Statement (4) gives a way, once you know that f is integrable on $[a, b]$, to write a formula for the integral of f:

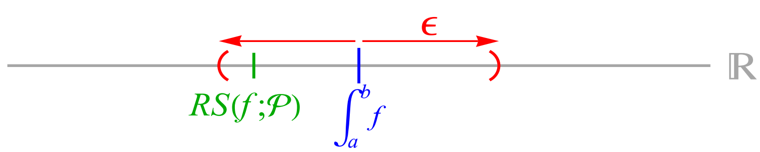
$$\int_a^b f = \sup_{\mathcal{P}}\{\mathcal{L}(f;\mathcal{P})\} = \inf_{\mathcal{P}}\{\mathcal{U}(f;\mathcal{P}\}.$$



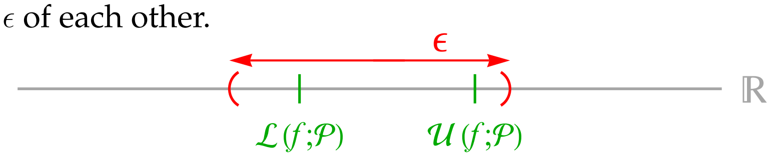
- Statement (3) says that once you know f is integrable, you can choose a partition $\mathcal{P}$ so that both the upper sum and the lower sum associated to that $\mathcal{P}$ is within ϵ of $\int_a^b f$:



- Statement (1) (the definition of the Riemann integral) says that once you know a function is integrable, then for <u>any</u> partition with suitably small norm, <u>any</u> Riemann sum associated to that partition is within ϵ of $\int_a^b f$.



- Statement (2) says that once you know a function is integrable, then for <u>all</u> partitions of suitably small norm, the upper and lower sums of that partition are within ϵ of each other.



PROOF First, we show $(1) \Rightarrow (2)$. Let $\epsilon > 0$.

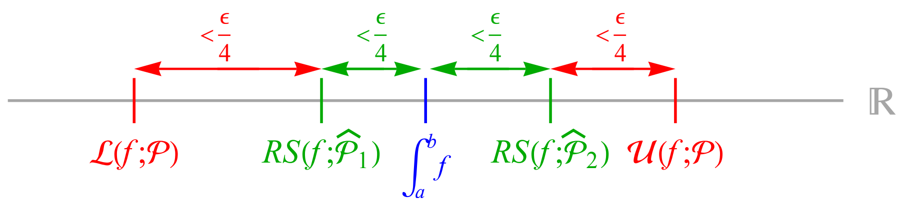
Since f is integrable, there is $\delta > 0$ such that

$$\|\widehat{\mathcal{P}}\| < \delta \text{ implies } \left| RS(f;\widehat{\mathcal{P}}) - \int_a^b f \right| < \frac{\epsilon}{4}.$$

Let $\mathcal{P}$ be a partition of $[a,b]$ with $\|\mathcal{P}\| < \delta$.

By Lemma 7.11,

- $\mathcal{P}$ can be tagged to make $\widehat{\mathcal{P}}_1$ so that $RS(f;\widehat{\mathcal{P}}_1) - \mathcal{L}(f;\mathcal{P}) < \frac{\epsilon}{4}$;
- $\mathcal{P}$ can also be tagged to make $\widehat{\mathcal{P}}_2$ so that $\mathcal{U}(f;\mathcal{P}) - RS(f;\widehat{\mathcal{P}}_2) < \frac{\epsilon}{4}$.



Now

$$\begin{aligned}
&\mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) \\
&= |\mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P})| \\
&\leq \mathcal{U}(f;\mathcal{P}) - RS(f;\widehat{\mathcal{P}}_2) + \left| RS(f;\widehat{\mathcal{P}}_2) - \int_a^b f \right| \\
&\qquad\qquad + \left| \int_a^b f - RS(f;\widehat{\mathcal{P}}_1) \right| + RS(f;\widehat{\mathcal{P}}_1) - \mathcal{L}(f;\mathcal{P}) \\
&< \frac{\epsilon}{4} + \frac{\epsilon}{4} + \frac{\epsilon}{4} + \frac{\epsilon}{4} = \epsilon.
\end{aligned}$$

This finishes the proof of $(1) \Rightarrow (2)$.

Next, we show $(2) \Rightarrow (3)$.

This is clear: just choose a partition $\mathcal{P}$ with $\|\mathcal{P}\| < \delta$ and apply (2).

Next, we prove $(3) \Rightarrow (4)$.

To do this, let

$$\begin{aligned}
A &= \sup_{\mathcal{P}}\{\mathcal{L}(f;\mathcal{P})\} = \sup\{\mathcal{L}(f;\mathcal{P}) : \mathcal{P} \text{ is a partition of } [a,b]\}; \\
B &= \inf_{\mathcal{P}}\{\mathcal{U}(f;\mathcal{P})\} = \inf\{\mathcal{U}(f;\mathcal{P}) : \mathcal{P} \text{ is a partition of } [a,b]\}.
\end{aligned}$$

Since every lower sum is less than or equal to every upper sum, $A \leq B$.

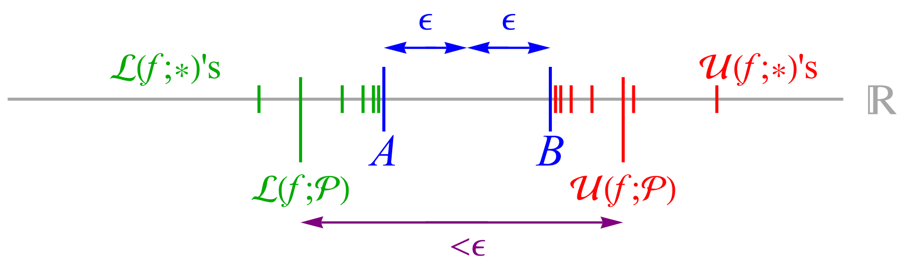
Suppose not, i.e. $A < B$. Then let $\epsilon = \dfrac{B-A}{2} > 0$.

Assuming (3), there is $\mathcal{P}$ such that

$$\mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) < \epsilon = \frac{B-A}{2}. \tag{7.3}$$

But $L(f;\mathcal{P}) \le A$ and $U(f;\mathcal{P}) \ge B$, so

$$\mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) \ge B - A. \tag{7.4}$$



(7.3) and (7.4) contradict one another, so A must equal B, proving (4).

Next, we prove $(4) \Rightarrow (3)$. To do this, we let

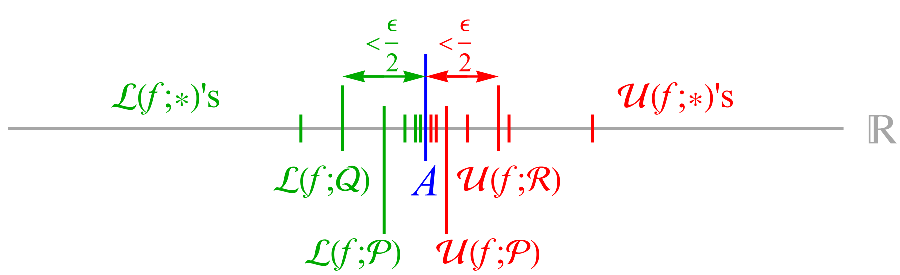
$$A = \sup_{\mathcal{P}}\{\mathcal{L}(f;\mathcal{P})\} = \inf_{\mathcal{P}}\{\mathcal{U}(f;\mathcal{P})\}.$$

By definition of sup (inf), $\exists$ partitions $\mathcal{Q}$, $\mathcal{R}$ such that

$$A - \mathcal{L}(f;\mathcal{Q}) < \frac{\epsilon}{2} \quad \text{and} \quad \mathcal{U}(f;\mathcal{R}) - A < \frac{\epsilon}{2}.$$

Let $\mathcal{P} = \mathcal{Q} \vee \mathcal{R}$; we have

$$A - \frac{\epsilon}{2} \;<\; \mathcal{L}(f;\mathcal{Q}) \;\le\; \mathcal{L}(f;\mathcal{P}) \;\le\; \mathcal{U}(f;\mathcal{P}) \;<\; \mathcal{U}(f;\mathcal{R}) \;<\; A + \frac{\epsilon}{2}$$



and this implies

$$\mathcal{U}(f;\mathcal{P}) - L(f;\mathcal{P}) < \left(A + \frac{\epsilon}{2}\right) - \left(A - \frac{\epsilon}{2}\right) = \epsilon,$$

showing (3) as wanted.

Next, let's prove $(3) \Rightarrow (2)$ (unfortunately, this is the hardest part).

To do this, let $\epsilon > 0$.

By (3), $\exists$ partition $\mathcal{Q} = \{y_0, \ldots, y_m\}$ of $[a,b]$ such that $U(f;\mathcal{Q}) - L(f;\mathcal{Q}) < \dfrac{\epsilon}{2}$.

Now let

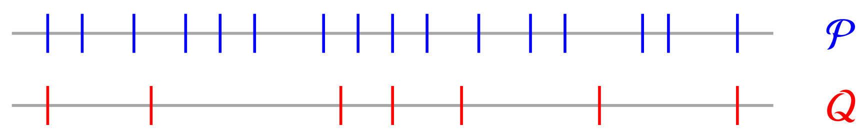
$$\delta = \frac{\epsilon}{4Mm}$$

where m is the size of $\mathcal{Q}$ and M is a bound for f (i.e. $|f(x)| \le M \; \forall \, x \in [a,b]$).

Now let $\mathcal{P} = \{x_0, \ldots, x_n\}$ be any partition of $[a,b]$ with $\|\mathcal{P}\| < \delta$.

We need to show $U(f;\mathcal{P}) - L(f;\mathcal{P}) < \epsilon$.

To do this, divide the indices of $\mathcal{P}$ (other than zero) into two types:



Note that $\#(I_2) \le (\text{size of } \mathcal{Q} - 2) < \text{size of } \mathcal{Q} = m$. Now,

$$\begin{aligned} U(f;\mathcal{P}) - L(f;\mathcal{P}) &= \sum_{k=1}^{n} w_k^{\mathcal{P}} \Delta x_k - \sum_{k=1}^{n} v_k^{\mathcal{P}} \Delta x_k \\ &= \sum_{k=1}^{n} \left[w_k^{\mathcal{P}} - v_k^{\mathcal{P}} \right] \Delta x_k \\ &= \sum_{k \in I_1} \left[w_k^{\mathcal{P}} - v_k^{\mathcal{P}} \right] \Delta x_k + \sum_{k \in I_2} \left[w_k^{\mathcal{P}} - v_k^{\mathcal{P}} \right] \Delta x_k \end{aligned}$$

So altogether, the expression on the previous page becomes

$$\begin{aligned}\mathcal{U}(f;\mathcal{P})-\mathcal{L}(f;\mathcal{P}) &< [\mathcal{U}(f;\mathcal{Q})-\mathcal{L}(f;\mathcal{Q})]+2M\delta m\\ &< \frac{\epsilon}{2}+2M\left(\frac{\epsilon}{4Mm}\right)m\\ &= \frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon.\end{aligned}$$

This proves $(3)\Rightarrow(2)$.

Finally, let's prove $(2)\Rightarrow(1)$. Again, we let

$$A=\sup_{\mathcal{P}}\{\mathcal{L}(f;\mathcal{P})\}=\inf_{\mathcal{P}}\{\mathcal{U}(f;\mathcal{P})\}.$$

To prove f is Riemann integrable, we need to show that

Toward that end, let $\epsilon>0$.
Assuming (2), we can choose $\delta>0$ so that

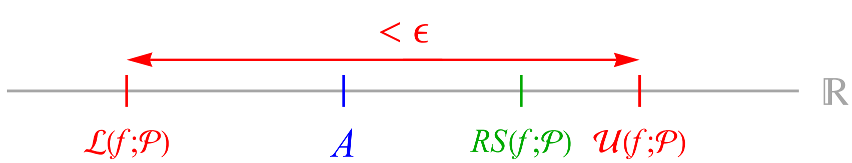
$$\|\mathcal{P}\|<\delta \text{ implies } \mathcal{U}(f;\mathcal{P})-\mathcal{L}(f;\mathcal{P})<\epsilon.$$

Now let $\mathcal{P}$ be any partition of $[a,b]$ with $\|\mathcal{P}\|<\delta$. We have

$$|\mathcal{U}(f;\mathcal{P})-A|=\mathcal{U}(f;\mathcal{P})-A\le\mathcal{U}(f;\mathcal{P})-\mathcal{L}(f;\mathcal{P})<\epsilon \text{ and}$$
$$|\mathcal{L}(f;\mathcal{P})-A|=A-\mathcal{L}(f;\mathcal{P})\le\mathcal{U}(f;\mathcal{P})-\mathcal{L}(f;\mathcal{P})<\epsilon.$$

Now, no matter how $\mathcal{P}$ is tagged to create $\widehat{\mathcal{P}}$,

$$RS(f;\widehat{\mathcal{P}})-A\le\mathcal{U}(f;\mathcal{P})-A<\epsilon \text{ and}$$
$$-(RS(f;\widehat{\mathcal{P}})-A)=A-RS(f;\widehat{\mathcal{P}})\le A-L(f;\mathcal{P})<\epsilon.$$

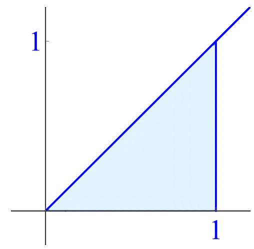


Therefore $|RS(f;\mathcal{P})-A|<\epsilon$, so $\int_a^b f=A$ by definition, proving (1). □

Examples

EXAMPLE 4, REVISITED

Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = x$. Prove f is integrable on $[0,1]$, and compute $\int_0^1 f$.

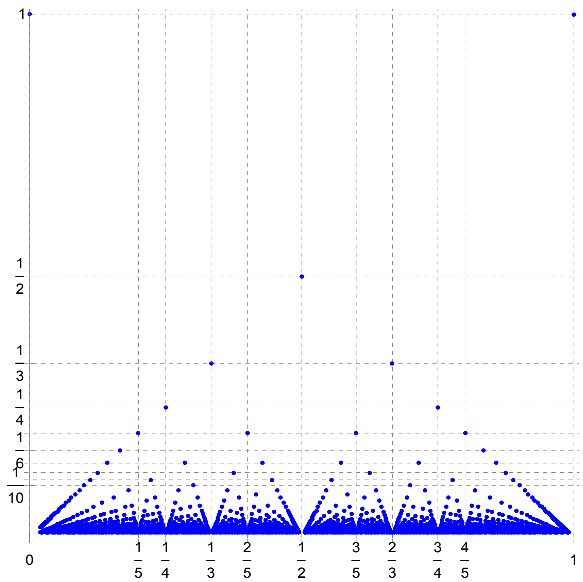


EXAMPLE 5

Determine if the Dirichlet function $\mathbb{1}_{\mathbb{Q}}$ is integrable on $[0,1]$. If so, compute $\int_0^1 \mathbb{1}_{\mathbb{Q}}$.

EXAMPLE 6

Determine if Thomae's function τ is integrable on $[0, 1]$. If so, compute $\int_0^1 \tau$.



Theorem 7.19 (Monotone functions are integrable) *Suppose $f : [a,b] \to \mathbb{R}$ is monotone. Then f is Riemann integrable on $[a,b]$.*

PROOF Suppose for now that f is increasing.

If $f(a) = f(b)$, then f is constant on $[a,b]$ so f is integrable on $[a,b]$ (This was Example 3, earlier in this chapter.)

So henceforth we'll assume $f(a) < f(b)$.

Now given $\epsilon > 0$, set $\delta = \dfrac{\epsilon}{f(b) - f(a)}$, so that $\delta[f(b) - f(a)] < \epsilon$.

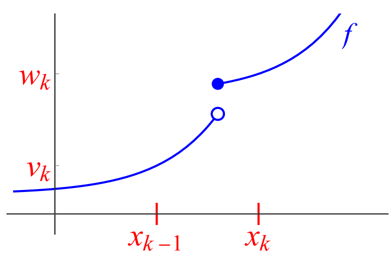
Let $\mathcal{P} = \{x_0, \dots, x_n\}$ be any partition of $[a,b]$ with $\|\mathcal{P}\| < \delta$.

We have

$$\mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) = \sum_{k=1}^{n} w_k \,\Delta x_k - \sum_{k=1}^{n} v_k \,\Delta x_k = \sum_{k=1}^{n} (w_k - v_k)\,\Delta x_k.$$

The key observation is that since f is increasing,

$$v_k =$$
$$w_k =$$



Therefore, from above we have

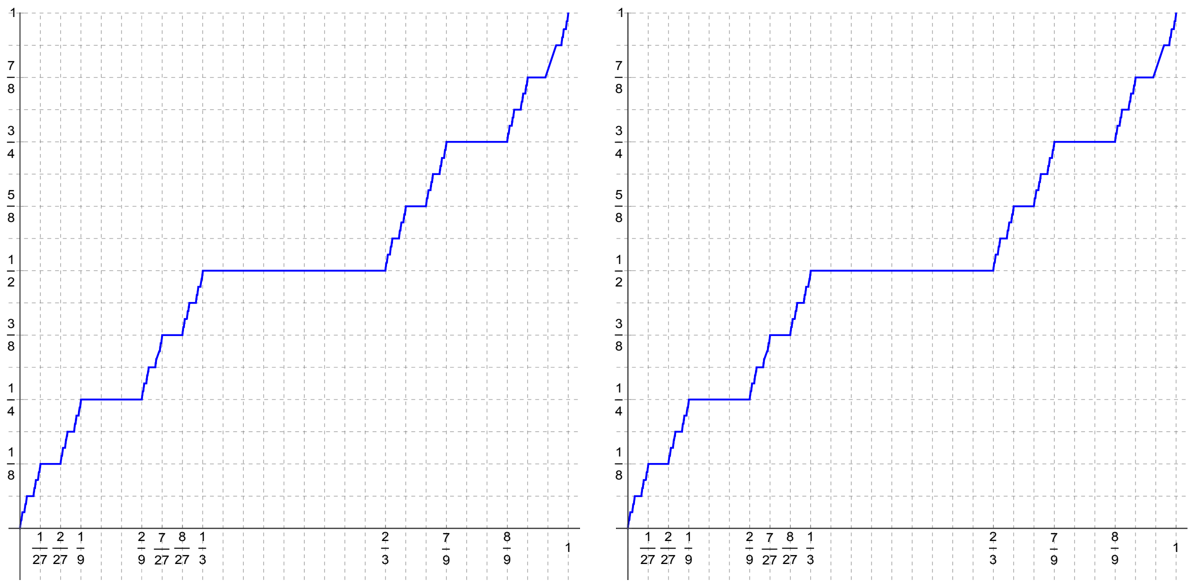
$$\begin{aligned}
&\mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) \\
&= \sum_{k=1}^{n} [f(x_k) - f(x_{k-1})]\,\Delta x_k \\
&< \sum_{k=1}^{n} [f(x_k) - f(x_{k-1})]\delta \\
&= \delta\left([f(x_1) - f(x_0)] + [f(x_2) - f(x_1)] + \dots + [f(x_n) - f(x_{n-1})]\right) \\
&= \delta[f(x_n) - f(x_0)] \\
&= \delta[f(b) - f(a)] \\
&< \epsilon.
\end{aligned}$$

By (3) of the integrabiilty criteria, f is integrable on $[a,b]$.

If f is decreasing, then $-f$ is increasing, so by the first part of this argument, $-f$ is integrable on $[a, b]$. By a theorem we haven't proven yet (but that you will do as HW), any multiple of an integrable function is integrable, so $-(-f) = f$ is integrable on $[a, b]$ as well. □

EXAMPLE 7

Let $c : [0, 1] \to \mathbb{R}$ be the Cantor function. Since c is monotone (earlier HW), c is Riemann integrable on $[0, 1]$. What is $\int_0^1 c$?



7.5 Properties of Riemann integrals

Theorem 7.20 (Integration is linear) *Let $a < b$ and suppose f and g are Riemann integrable on $[a, b]$. Then:*

1. *$f + g$ is Riemann integrable on $[a, b]$ and* $\int_a^b (f+g) = \int_a^b f + \int_a^b g$.
2. *For any $r \in \mathbb{R}$, rf is Riemann integrable on $[a, b]$ and* $\int_a^b (rf) = r\int_a^b f$.

PROOF First, let's prove (1). Let $\epsilon > 0$.

Since f and g are integrable, there is $\delta_f > 0$ and $\delta_g > 0$ such that

$$||\mathcal{P}|| < \delta_f \text{ implies } \left| RS(f;\widehat{\mathcal{P}}) - \int_a^b f \right| <$$
$$||\mathcal{P}|| < \delta_g \text{ implies } \left| RS(g;\widehat{\mathcal{P}}) - \int_a^b g \right| <$$

Now let $\delta =$

If $||\mathcal{P}|| < \delta$, then

$$\left| RS(f+g;\widehat{\mathcal{P}}) - \left(\int_a^b f + \int_a^b g \right) \right| =$$

The proof of statement (2) is left as a HW problem.
As a hint, start by letting $\epsilon > 0$.
You have to figure out what $\delta > 0$ has to be (in terms of ϵ) so that

$$||\mathcal{P}|| < \delta \text{ implies } \left| RS(rf;\widehat{\mathcal{P}}) - r\int_a^b f \right| < \epsilon.$$

Theorem 7.21 (Additivity of integrals) *Let $a < c$ and suppose $f : [a,c] \to \mathbb{R}$. Let $b \in (a,c)$. Then, the following two statements are equivalent:*

1. *f is Riemann integrable on $[a,c]$.*
2. *f is Riemann integrable on both $[a,b]$ and $[b,c]$.*

Furthermore, when these statements are true, we have

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

PROOF (1) ⇒ (2): Suppose f is Riemann integrable on $[a,c]$.

Let $\epsilon > 0$.

By the integrability criterion, $\exists\,\delta > 0$ so that for any partition $\mathcal{P}$ of $[a,b]$,

$$\|\mathcal{P}\| < \delta \text{ implies } \mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) < \epsilon.$$

Let $\mathcal{P}$ be any partition of $[a,c]$ with $\|\mathcal{P}\| < \delta$ that contains b. Write

$$\mathcal{P} = \{a = x_0, x_1, x_2, \ldots, x_n = b, x_{n+1}, \ldots, x_m\}$$

and then set

$$\mathcal{P}_1 = \{a = x_0, x_1, x_2, \ldots, x_n = b\} \text{ and } \mathcal{P}_2 = \{b = x_n, x_{n+1}, \ldots, x_m = c\}.$$

These are partitions of $[a,b]$ and $[b,c]$, respectively. Now

$$\begin{aligned}\mathcal{U}(f;\mathcal{P}_1) - \mathcal{L}(f;\mathcal{P}_1) &= \sum_{k=1}^{n} \left(w_k^{\mathcal{P}_1} - v_k^{\mathcal{P}_1}\right) \Delta x_k \\ &\le \sum_{k=1}^{m} \left(w_k^{\mathcal{P}} - v_k^{\mathcal{P}}\right) \Delta x_k \\ &= \mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) < \epsilon,\end{aligned}$$

so by the integrability criterion, f is Riemann integrable on $[a,b]$.

Similarly,

$$\begin{aligned}\mathcal{U}(f;\mathcal{P}_2) - \mathcal{L}(f;\mathcal{P}_2) &= \sum_{k=n+1}^{m} \left(w_k^{\mathcal{P}_2} - v_k^{\mathcal{P}_2}\right) \Delta x_k \\ &\le \sum_{k=1}^{m} \left(w_k^{\mathcal{P}} - v_k^{\mathcal{P}}\right) \Delta x_k \\ &= \mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) < \epsilon,\end{aligned}$$

so by the integrability criterion, f is Riemann integrable on $[b,c]$.

(2) ⇒ (1): Suppose f is Riemann integrable on both $[a,b]$ and $[b,c]$. Let

$$I_1 = \int_a^b f \text{ and } I_2 = \int_b^c f.$$

To show $\int_a^c f = I_1 + I_2$, let $\epsilon > 0$.

By hypothesis, there is a partition $\mathcal{P}_1 = \{a = x_0, x_1, \ldots, x_{n_1} = b\}$ of $[a,b]$ s.t.

$$\mathcal{U}(f;\mathcal{P}_1) - I_1 < \frac{\epsilon}{4} \text{ and } I_1 - \mathcal{L}(f;\mathcal{P}_1) < \frac{\epsilon}{4},$$

and a partition $\mathcal{P}_2 = \{b = x_{n_1}, x_{n_1+1}, \ldots, x_m = c\}$ of $[b,c]$ s.t.

$$\mathcal{U}(f;\mathcal{P}_2) - I_2 < \frac{\epsilon}{4} \text{ and } I_2 - \mathcal{L}(f;\mathcal{P}_2) < \frac{\epsilon}{4}.$$

Now let $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2 = \{a = x_0, x_1, x_2, \ldots, x_m = c\}$.

For this partition of $[a,c]$,

$$\begin{aligned}
\mathcal{U}(f;\mathcal{P}) &= \sum_{k=1}^{n} w_k^{\mathcal{P}} \, \Delta x_j \\
&= \sum_{k=1}^{n} w_k^{\mathcal{P}} \, \Delta x_k + \sum_{k=n+1}^{m} w_k^{\mathcal{P}} \, \Delta x_j \\
&= \sum_{k=1}^{n} w_k^{\mathcal{P}_1} \, \Delta x_k + \sum_{k=n+1}^{m} w_k^{\mathcal{P}_2} \, \Delta x_j \\
&= \mathcal{U}(f;\mathcal{P}_1) + \mathcal{U}(f;\mathcal{P}_2)
\end{aligned}$$

Similarly, $\mathcal{L}(f;\mathcal{P}) = \mathcal{L}(f;\mathcal{P}_1) + \mathcal{L}(f;\mathcal{P}_2)$ (same proof with v_k's instead of w_k's).

Therefore

$$\begin{aligned}
\mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) &= [\mathcal{U}(f;\mathcal{P}_1) + \mathcal{U}(f;\mathcal{P}_2)] - [\mathcal{L}(f;\mathcal{P}_1) + \mathcal{L}(f;\mathcal{P}_2)] \\
&= [\mathcal{U}(f;\mathcal{P}_1) - \mathcal{L}(f;\mathcal{P}_1)] + [\mathcal{U}(f;\mathcal{P}_2) - \mathcal{L}(f;\mathcal{P}_2)] \\
&= [\mathcal{U}(f;\mathcal{P}_1) - I_1] + [I_1 - \mathcal{L}(f;\mathcal{P}_1)] + [\mathcal{U}(f;\mathcal{P}_2) - I_2] + [I_2 - \mathcal{L}(f;\mathcal{P}_2)] \\
&< \frac{\epsilon}{4} + \frac{\epsilon}{4} + \frac{\epsilon}{4} + \frac{\epsilon}{4} \\
&= \epsilon.
\end{aligned}$$

By the integrability criterion, f is Riemann integrable on $[a,c]$.

Furthermore,

$$\mathcal{U}(f;\mathcal{P}) - (I_1 + I_2) = [\mathcal{U}(f;\mathcal{P}_1) - I_1] + [\mathcal{U}(f;\mathcal{P}_2) - I_2] < \frac{\epsilon}{4} + \frac{\epsilon}{4} = \frac{\epsilon}{2},$$

and since $\epsilon > 0$ is arbitrary, this implies

$$I_1 + I_2 \geq \inf_{\mathcal{P}} \mathcal{U}(f;\mathcal{P}) = \int_a^c f. \tag{7.5}$$

Similarly

$$(I_1 + I_2) - \mathcal{L}(f;\mathcal{P}) = [I_1 - \mathcal{L}(f;\mathcal{P}_1)] + [I_2 - \mathcal{L}(f;\mathcal{P}_2)] < \frac{\epsilon}{4} + \frac{\epsilon}{4} = \frac{\epsilon}{2},$$

and since $\epsilon > 0$ is arbitrary, we have

$$I_1 + I_2 \leq \sup_{\mathcal{P}} \mathcal{L}(f;\mathcal{P}) = \int_a^c f. \tag{7.6}$$

Inequalities (7.5) and (7.6) together imply

$$I_1 + I_2 = \int_a^b f + \int_b^c f = \int_a^c f. \square$$

Definition 7.22 *If $a < b$ and f is Riemann integrable on $[a, b]$, then we define*

$$\int_b^a f = -\int_a^b f.$$

Definition 7.23 *If f is Riemann integrable on any interval I of positive length containing a, then we define*

$$\int_a^a f = 0.$$

Corollary 7.24 (Additivity of integrals (general situation)) *For any $a, b, c \in \mathbb{R}$, so long as all these integrals exist, we have*

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

PROOF This is just a bunch of cases depending on which of a, b, c is the least and which is the greatest. For instance, if $c \leq a \leq b$, then

$$\int_c^b f = \int_c^a f + \int_a^b f$$

by Theorem 7.21. Restated, this is

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

which rearranges into

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

as wanted. The other cases are similar and omitted. □

7.6 Uniform continuity and the integrability of continuous functions

Uniform continuity

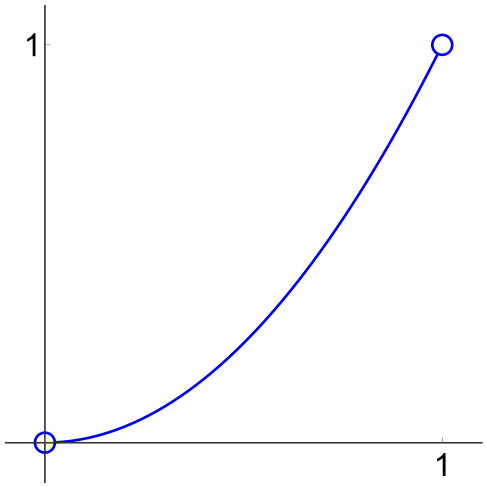
Recall what it means for $f : \mathbb{R} \to \mathbb{R}$ to be continuous on a set $E \subseteq \mathbb{R}$:

Sometimes, we need to be able to choose one δ that works for all $x \in E$, rather than choosing a different δ for each x. (We'll see one reason why when we prove that any continuous function on a compact interval is Riemann integrable on that interval.)

Unfortunately, this is not assured just because a function is continuous on E. Let's consider two examples to learn more:

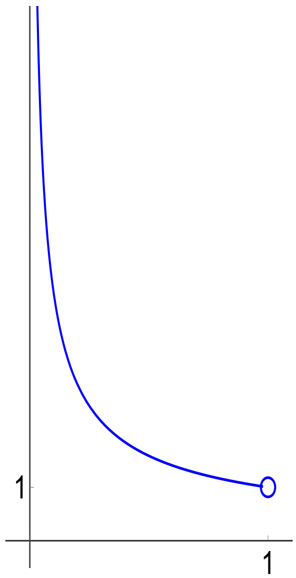
EXAMPLE A

Let $f(x) = x^2$ and suppose $E = (0,1)$. Prove f is continuous on E.



EXAMPLE B

Let $f(x) = \dfrac{1}{x}$ and suppose $E = (0,1)$. Prove f is continuous on E.



Despite the similar looking arguments, there is a big difference between Examples A and B above.

Definition 7.25 *Let $f : \mathbb{R} \to \mathbb{R}$ and suppose $E \subseteq \mathbb{R}$. We say f is* **uniformly continuous (unif. cts.)** *on E if for every $\epsilon > 0$, there is $\delta > 0$ such that for all $a, x \in E$,*

$$|x - a| < \delta \textit{ implies } |f(x) - f(a)| < \epsilon.$$

On the previous pages, we showed:

- $f(x) = x^2$ is unif. cts. on $(0, 1)$, but
- $f(x) = \dfrac{1}{x}$, despite being cts on $(0, 1)$, is not unif. cts. on $(0, 1)$.

Theorem 7.26 *Let $f : \mathbb{R} \to \mathbb{R}$. If $E \subseteq \mathbb{R}$ is compact and f is continuous on E, then f is uniformly continuous on E.*

PROOF Let $\epsilon > 0$.

Since f is continuous at each $a \in E$, there is $\delta(a) > 0$ such that

$$|x - a| < \delta(a) \text{ implies } |f(x) - f(a)| < \frac{\epsilon}{2}. \tag{7.7}$$

Consider the open cover $\{B_{\frac{1}{2}\delta(a)}(a) : a \in E\}$ of E.

Since E is compact, there is a finite subcover:

$$\{B_{\frac{1}{2}\delta(a_1)}(a_1), B_{\frac{1}{2}\delta(a_2)}(a_2), ..., B_{\frac{1}{2}\delta(a_n)}(a_n)\}.$$

Let $\delta = \min\left\{\frac{1}{2}\delta(a_1), .., \frac{1}{2}\delta(a_n)\right\}$.

Now, let $x, y \in E$ be such that $|x - y| < \delta$.
Since the sets $B_{\frac{1}{2}\delta(a_j)}(a_j)$ cover E, $x \in B_{\frac{1}{2}\delta(a_j)}(a_j)$ for some j.

Therefore $|x - a_j| < \frac{1}{2}\delta(a_j)$, meaning $|f(x) - f(a_j)| < \frac{\epsilon}{2}$ by (7.7).

Furthermore, $|y - a_j| \le |y - x| + |x - a_j| < \delta + \frac{1}{2}\delta(a_j) < \frac{1}{2}\delta(a_j) + \frac{1}{2}\delta(a_j) = \delta(a_j)$.

So by (7.7) again, $|f(y) - f(a_j)| < \frac{\epsilon}{2}$.

Putting this together with the triangle inequality,

$$|f(x) - f(y)| \le |f(x) - f(a_j)| + |f(a_j) - f(y)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

making f unif. cts on E by definition. □

APPLICATION

$f(x) = x \sin 3x$ is unif. cts. on $[-3, 7]$.

The important consequence of Theorem 7.26 is that we can use uniform continuity to show that continuous functions must be integrable, for if $f : [a, b] \to \mathbb{R}$ is continuous, it must be uniformly continuous since $[a, b]$ is compact.

Theorem 7.27 *Suppose f and g are Riemann integrable on $[a, b]$.*

1. *For any $\varphi : \mathbb{R} \to \mathbb{R}$ which is continuous, $\varphi \circ f$ is integrable on $[a, b]$.*
2. *f^2 is integrable on $[a, b]$.*
3. *fg is integrable on $[a, b]$.*

PROOF Let's start with the first statement. Let $\epsilon > 0$.

The eventual goal is to make a partition $\mathcal{P}$ such that $\mathcal{U}(\varphi \circ f; \mathcal{P}) - \mathcal{L}(\varphi \circ f; \mathcal{P}) < \epsilon$.

Part 1: Our goal in this part is to find a bound on φ.

f is integrable on $[a, b]$, so f is bounded on $[a, b]$, meaning

$$f([a, b]) \subseteq [c, d]$$

for suitable $c, d \in \mathbb{R}$.

Since φ is continuous on $[c, d]$, it is unif. cts. on $[c, d]$.

Thus $\exists\, \delta > 0$ such that for all $y, y_0 \in [c, d]$,

$$|y - y_0| < \delta \text{ implies } |\varphi(y) - \varphi(y_0)| < \frac{\epsilon}{2(b-a)}. \tag{7.8}$$

Also, since φ is cts on $[c, d]$, the image $\varphi([c, d])$ is compact, hence bounded.

So there is $M \geq 0$ such that $|\varphi(y)| \leq M$ for all $y \in [c, d]$.

Part 2: Now we define our partition $\mathcal{P}$.

Since f is integrable on $[a, b]$, $\exists$ partition $\mathcal{P} = \{x_0, \ldots, x_n\}$ of $[a, b]$ with

$$\mathcal{U}(f; \mathcal{P}) - \mathcal{L}(f; \mathcal{P}) < \frac{\delta\epsilon}{4M}. \tag{7.9}$$

Part 3: We split the subintervals of $\mathcal{P}$ into two types.

- let $I_1 = \{k \in \{1, ..., n\} : w_k^f - v_k^f < \delta\}$;
- let $I_2 = \{k \in \{1, ..., n\} : w_k^f - v_k^f \geq \delta\}$.

Part 4: We show that for the first type of subinterval, $w_k^{\varphi \circ f} - v_k^{\varphi \circ f}$ is small.

For every $k \in I_1$, and every $x, y \in [x_{k-1}, x_k]$, $|f(x) - f(y)| < \delta$ so by (7.8),

$$|(\varphi \circ f)(x) - (\varphi \circ f)(y)| < \frac{\epsilon}{2(b-a)}$$

Therefore for $k \in I_1$, $w_k^{\varphi \circ f} - v_k^{\varphi \circ f} \leq \frac{\epsilon}{2(b-a)}$.

Part 5: We show that the second type of subintervals are very skinny, collectively.

Claim: The total length of the subintervals corresponding to indices in I_2 must be at most $\frac{\epsilon}{4M}$.

Proof of claim: Suppose not. Then we would have

$$\begin{aligned}
\mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) &= \sum_{k=1}^{n} \left(w_k^f - v_k^f\right) \Delta x_k \\
&\geq \sum_{k \in I_2} \left(w_k^f - v_k^f\right) \Delta x_k \\
&\geq \sum_{k \in I_2} \delta \, \Delta x_k \\
&= \delta \sum_{k \in I_2} \Delta x_k \geq \delta \frac{\epsilon}{4M}, \quad \text{contradicting (7.9).}
\end{aligned}$$

Part 6: Put everything together and show $\mathcal{U}(\varphi \circ f;\mathcal{P}) - \mathcal{L}(\varphi \circ f;\mathcal{P}) < \epsilon$.

$$\begin{aligned}
\mathcal{U}(\varphi \circ f;\mathcal{P}) - \mathcal{L}(\varphi \circ f;\mathcal{P}) &= \sum_{k=1}^{n} \left(w_k^{\varphi\circ f} - v_k^{\varphi\circ f}\right) \Delta x_k \\
&= \sum_{k \in I_1} \left(w_k^{\varphi\circ f} - v_k^{\varphi\circ f}\right) \Delta x_k + \sum_{k \in I_2} \left(w_k^{\varphi\circ f} - v_k^{\varphi\circ f}\right) \Delta x_k \\
&\leq \sum_{k \in I_1} \left(w_k^{\varphi\circ f} - v_k^{\varphi\circ f}\right) \Delta x_k + \sum_{k \in I_2} (M - (-M)) \Delta x_k \quad \text{(from part 1)} \\
&< \sum_{k \in I_1} \frac{\epsilon}{2(b-a)} \Delta x_k + \sum_{k \in I_2} (M - (-M)) \Delta x_k \quad \text{(from part 4)} \\
&= \frac{\epsilon}{2(b-a)} \sum_{k \in I_1} \Delta x_k + 2M \sum_{k \in I_2} \Delta x_k \\
&\leq \frac{\epsilon}{2(b-a)} (b-a) + 2M \sum_{k \in I_2} \Delta x_k \\
&\leq \frac{\epsilon}{2} + 2M \left(\frac{\epsilon}{4M}\right) \quad \text{(from the claim in part 5)} \\
&= \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

By (3) of the integrability criteria, $\varphi \circ f$ is integrable on $[a, b]$.

For (2), note that $\varphi(x) = x^2$ is cts. Therefore $\varphi \circ f = f^2$ is integrable by (1).
For (3), observe

$$fg = \frac{1}{2}\left[(f+g)^2 - f^2 - g^2\right],$$

which is integrable by statement (2), together with linearity. □

Corollary 7.28 (Continuous functions are integrable) *Suppose $\varphi : [a, b] \to \mathbb{R}$ is continuous. Then φ is Riemann integrable on $[a, b]$.*

PROOF The function $f(x) = x$ is integrable on $[a, b]$ (earlier example). Apply statement (1) of the previous theorem.

7.7 Fundamental Theorem of Calculus

Order properties of the Riemann integral

Theorem 7.29 (Order properties) *Let $a < b$ and suppose f and g are Riemann integrable on $[a, b]$. Then:*

1. *if $f \geq 0$ on $[a, b]$, then $\int_a^b f \geq 0$.*
2. *if $f \leq g$ on $[a, b]$, then $\int_a^b f \leq \int_a^b g$.*
3. **Max-Min Inequality for Integrals:** *if $m, M \in \mathbb{R}$ are such that $m \leq f(x) \leq M$ on $[a, b]$, then*
$$m(b-a) \leq \int_a^b f \leq M(b-a).$$
4. **Triangle Inequality for Integrals:** *$|f|$ is Riemann integrable on $[a, b]$, and*
$$\left|\int_a^b f\right| \leq \int_a^b |f|.$$

PROOF For statement (1), first notice that $f \geq 0$ implies that all the $v_j \geq 0 \; \forall \, j$, so any lower sum for f is nonnegative, because

$$\mathcal{L}(f;\mathcal{P}) = \sum_{j=1}^{n} v_j \, \Delta x_j \geq 0.$$

Thus $\int_a^b f = \sup_{\mathcal{P}} \mathcal{L}(f;\mathcal{P}) \geq 0$.

To prove (2), suppose $f \leq g$. Let $h = g - f$; then $h \geq 0$ on $[a,b]$, so by (1),

$$0 \leq \int_a^b h = \int_a^b (f - g) = \int_a^b f - \int_a^b g.$$

Rearrange this inequality to get (2).

For statement (3), apply (2): consider the constant functions m and M; by (2) and our previous computation of the integral of a constant function, we have

$$m(b-a) = \int_a^b m \leq \int_a^b f \leq \int_a^b M = M(b-a)$$

as wanted.

The first part of statement (4) follows from Theorem 7.27, since $\varphi(x) = |x|$ is cts.

Finally, since $-|f| \leq f \leq |f|$, we have (by (2) of this theorem)

$$-\int_a^b |f| \leq \int_a^b f \leq \int_a^b |f|,$$

so

$$\left| \int_a^b f \right| \leq \int_a^b |f|. \; \square$$

The Fundamental Theorem of Calculus

QUESTION

How do you actually compute integrals?

Theorem 7.30 (Fundamental Theorem of Calculus (Part 1)) *Let $E \subseteq \mathbb{R}$ be open, and suppose $f : E \to \mathbb{R}$ is cts. Let $a \in E$ and define $F : E \to \mathbb{R}$ by*

$$F(x) = \int_a^x f(t)\,dt.$$

Then F is differentiable on E and $F' = f$.

PROOF Let $x_0 \in E$. We need to show the following limit statement:

$$\lim_{h\to 0} \frac{F(x_0+h) - F(x_0)}{x_0} = f(x_0).$$

To do this, let $\epsilon > 0$.

We start by using the continuity of f to get a bound on the integral of f on a small interval near x_0. Since f is cts at x_0, there is $\delta > 0$ s.t.

$$|t - x_0| < \delta \Rightarrow |f(t) - f(x_0)| < \epsilon.$$

Thinking of t as $x_0 + h$, this means that if $|h| < \delta$, then for all $t \in [x_0, x_0 + h]$,

$$|t - x_0| \le |h| < \delta \text{ so } |f(t) - f(x_0)| < \epsilon.$$

Now, we can verify the limit statement from earlier. For h such that $0 < |h - 0| < \delta$,

$$\begin{aligned}
\left|\frac{F(x_0+h) - F(x_0)}{h} - f(x_0)\right| &= \left|\frac{\int_a^{x_0+h} f(t)\,dt - \int_a^{x_0} f(t)\,dt}{h} - f(x_0)\right| \\
&= \left|\frac{\int_{x_0}^{x_0+h} f(t)\,dt}{h} - f(x_0)\right| \\
&= \left|\frac{\int_{x_0}^{x_0+h} f(t)\,dt - hf(x_0)}{h}\right| \\
&= \left|\frac{\int_{x_0}^{x_0+h}[f(t) - f(x_0)]\,dt}{h}\right| \\
&= \frac{1}{|h|}\left|\int_{x_0}^{x_0+h} [f(t) - f(x_0)]\,dt\right| \\
&\le \frac{1}{|h|}\int_{x_0}^{x_0+h} |f(t) - f(x_0)|\,dt \\
&< \frac{1}{|h|}\int_{x_0}^{x_0+h} \epsilon\,dt \\
&\le \frac{1}{|h|}\epsilon\,|(x_0+h) - x_0| = \frac{1}{|h|}\epsilon\,|h| = \epsilon.
\end{aligned}$$

Therefore
$$\lim_{h\to 0} \frac{F(x_0+h)-F(x)}{h} = f(x_0),$$
i.e. $F'(x_0) = f(x_0)$ as wanted. Since $x_0 \in U$ was arbitrarily chosen, $F' = f$ on E, proving the FTC. □

Corollary 7.31 (Cts functions have antiderivatives) *If $E \subseteq \mathbb{R}$ is open and $f : E \to \mathbb{R}$ is cts, then there is a differentiable function $'f : E \to \mathbb{R}$ such that $('f)' = f$ on E.*

PROOF Choose $a \in E$ and let $'f(x) = \int_a^x f(x)\,dx$. Apply the FTC. □

Corollary 7.32 (Fundamental Theorem of Calculus (Part 2)) *Let $E \subseteq \mathbb{R}$ be open and $f : E \to \mathbb{R}$ be continuous. If $'f : E \to \mathbb{R}$ is any differentiable function such that $('f)' = f$ on E, then for any $a, b \in E$, we have*
$$\int_a^b f = [\,'f]_a^b = \,'f(b) - \,'f(a).$$

Significance: To compute an integral, it is sufficient to find any antiderivative of f.

PROOF Let $'f$ be any antiderivative of f. as in the theorem.

Define $F(x) = \int_a^x f(t)\,dt - \,'f(x)$. Note
$$F'(x) = f(x) - f(x) = 0$$
so by the ____________, F is constant.

Since $F(a) = \int_a^a f(t)\,dt - \,'f(a)$, we have $F(a) = -\,'f(a)$.

As F is constant, we have, for any $b \in U$,
$$\begin{aligned} 0 &= F(b) - F(a) \\ &= \left[\int_a^b f(t)\,dt - \,'f(b)\right] - \left[\int_a^a f(t)\,dt - \,'f(a)\right] \\ &= \int_a^b f(t)\,dt - \,'f(b) - 0 + \,'f(a) \end{aligned}$$

Rearrange this to get
$$\int_a^b f = \,'f(b) - \,'f(a). \;\square$$

APPLICATION

$$\int_1^2 x^2\,dx =$$

Remarks on the inverse relationship between differentiation and integration

In Calculus 1, you are told that differentiation and integration are inverse operations:

This story you are told is a lie (or at the very least, it's a gross oversimplification).

EXAMPLE 8

Let $\tau : [0,1] \to \mathbb{R}$ be Thomae's function.

1. Let $F : [0,1] \to \mathbb{R}$ be $F(x) = \int_0^x \tau(t)\,dt$. What is F?
2. For the function F in the previous question, what is F'?

EXAMPLE 9

Let $f : \mathbb{R} \to \mathbb{R}$ be $f(x) = \begin{cases} x^2 \sin \dfrac{1}{x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$.

1. Prove f is differentiable at all x, and compute $f'(x)$.
2. Is $\int_0^1 f'(x)\,dx = f(1) - f(0)$?

Integration techniques

We end this section by verifying two Calc 2 techniques for computing integrals:

u-substitutions

Theorem 7.33 *Let $E_1, E_2 \subseteq \mathbb{R}$ be open sets, and suppose*

1. $g : E_1 \to E_2$ *is differentiable on E (hence continuous on E);*
2. $g' : E_1 \to \mathbb{R}$ *is continuous; and*
3. $f : E_2 \to \mathbb{R}$ *is continuous.*

Then, for all $a < b$ in E_1, $\displaystyle\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du.$

PROOF Let $F : E_2 \to \mathbb{R}$ be defined by $F(y) = \displaystyle\int_{g(a)}^{y} f(u)\,du$.

By the FTC, F is [ ] and $F' =$ [ ].

Now let $G = F \circ g$. This makes $G(x) \boxed{=} \displaystyle\int_{g(a)}^{g(x)} f(u)\,du$.

Notice $G(a) = \displaystyle\int_{g(a)}^{g(a)} f(u)\,du = 0$.

By the Chain Rule, $G : E_1 \to \mathbb{R}$ is differentiable and

$$G'(x) = F'(g(x))g'(x) = f(g(x))g'(x).$$

On the one hand, this makes G an antiderivative of $f(g(x))g'(x)$.

But also, by the FTC $\displaystyle\int_a^x f(g(t))g'(t)$ is an antiderivative of $f(g(x))g'(x)$.

So by the [ ] Theorem, $\exists C$ so that

$$G(x) = \int_a^x f(g(t))g'(t)\,dt + C.$$

Plug in $x = a$ to both sides of this to get $G(a) = C$, i.e. $0 = C$. That means

$$G(x) \boxed{=} \int_a^x f(g(t))g'(t)\,dt.$$

Plugging in $x = b$, we get

$$\int_a^b f(g(t))g'(t)\,dt \boxed{=} G(b) \boxed{=} \int_{g(a)}^{g(b)} f(u)\,du$$

as wanted. □

Parts

Theorem 7.34 *Let $E \subseteq \mathbb{R}$ be open and suppose $f, g : E \to \mathbb{R}$ are continuous. Suppose $F, G : E \to \mathbb{R}$ are differentiable with $F' = f$ and $G' = g$ on E. Then, $\forall\, a < b \in E$, $\int_a^b fG = (FG)(b) - (FG)(a) - \int_a^b Fg$.*

PROOF By the Product Rule,

$$(FG)' = F'G + FG' = fG + Fg.$$

Since F and G are continuous on E, fG and Fg are products of continuous functions, hence continuous on E. By the FTC,

$$(FG)(b) - (FG)(a) = \int_a^b (FG)' = \int_a^b (fG + Fg) = \int_a^b fG + \int_a^b Fg.$$

This rearranges into the statement of the theorem. □

7.8 Interchange of limit and integral

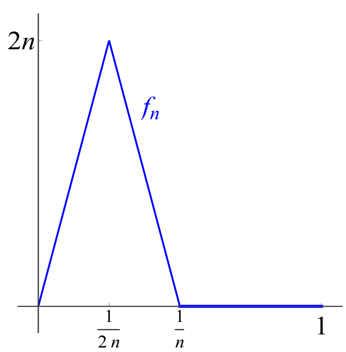
QUESTION 1

Let $E = [a, b] \subseteq \mathbb{R}$, and $\{f_n\}$ a sequence of integrable functions $E \to \mathbb{R}$.

If $f_n \to f$ on E, does $\int_a^b f_n \to \int_a^b f$?

(In other words, is $\lim \int_a^b f_n = \int_a^b (\lim f_n)$?)

EXAMPLE $E = [0, 1]$; $f_n(x) = \begin{cases} 2n - 2n\left|x - \dfrac{1}{2n}\right| & 0 \le x \le \dfrac{1}{n} \\ 0 & x > \dfrac{1}{n} \end{cases}$



QUESTION 2

Let $E = [a,b] \subseteq \mathbb{R}$, and $\{f_n\}$ a sequence of integrable functions $E \to \mathbb{R}$.

If $f_n \rightrightarrows f$ on E, does $\int_a^b f_n \to \int_a^b f$?

Theorem 7.35 (Interchange of limit and integral) *Suppose $\{f_n\}$ is a sequence of integrable functions on $E = [a,b]$, and suppose $f_n \rightrightarrows f$ on E.*

Then f is integrable on E and $\lim_{n\to\infty} \int_a^b f_n = \int_a^b f$.

PROOF First, we claim $\left\{\int_a^b f_n\right\}$ is a Cauchy sequence *of numbers*.

To show this, let $\epsilon > 0$.

Since $f_n \rightrightarrows f$, $\{f_n\}$ is uniformly Cauchy, so $\exists\, M$ s.t.

$$m, n \geq N \Rightarrow |f_m(x) - f_n(x)| < \frac{\epsilon}{b-a} \quad \forall\, x \in E.$$

Restated, we have $\forall\, x \in E$,

$$\begin{aligned} \frac{-\epsilon}{b-a} < f_m(x) - f_n(x) < \frac{\epsilon}{b-a} \\ \Rightarrow \int_a^b \frac{-\epsilon}{b-a} < \int_a^b [f_m(x) - f_n(x)] < \int_a^b \frac{\epsilon}{b-a} \\ \Rightarrow -\epsilon < \int_a^b f_m - \int_a^b f_n < \epsilon \end{aligned}$$

which means $\left|\int_a^b f_m - \int_a^b f_n\right| < \epsilon$.

This proves the claim, so by completeness, $\exists L \in \mathbb{R}$ so that $\int_a^b f_n \to L$.

Now, we will show $\int_a^b f = L$ using the definition of integral.

To do this, let $\epsilon > 0$.

Since $f_n \rightrightarrows f$, $\exists N_1$ so that

$$n \geq N_1 \Rightarrow |f_n(x) - f(x)| < \frac{\epsilon}{3(b-a)} \quad \forall\, x \in [a,b].$$

Since $\int_a^b f_n \to L$, $\exists N_2$ so that

$$n \geq N_2 \Rightarrow \left|\int_a^b f_n - L\right| < \frac{\epsilon}{3}.$$

Let $N = \max\{N_1, N_2\}$. Since f_N is integrable on $[a,b]$, $\exists\,\delta > 0$ so that

$$\|\widehat{\mathcal{P}}\| < \delta \Rightarrow \left| RS(f_N;\widehat{\mathcal{P}}) - \int_a^b f_N \right| < \frac{\epsilon}{3}.$$

Now, let $\widehat{\mathcal{P}}$ be a tagged partition of $[a,b]$ with $\|\widehat{\mathcal{P}}\| < \delta$.

$$\begin{aligned}
|RS(f;\widehat{\mathcal{P}}) - L| &\le |RS(f;\widehat{\mathcal{P}}) - RS(f_N;\widehat{\mathcal{P}})| + \left| RS(f_N;\widehat{\mathcal{P}}) - \int_a^b f_N \right| + \left| \int_a^b f_N - L \right| \\
&< \left| \sum_{k=1}^{n} [f(c_k) - f_N(c_k)] \right| \Delta x_j + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&\le \sum_{k=1}^{n} |f(c_k) - f_N(c_k)|\, \Delta x_j + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&< \sum_{k=1}^{n} \frac{\epsilon}{3(b-a)} \Delta x_j + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&= \frac{\epsilon}{3(b-a)} \sum_{k=1}^{n} \Delta x_j + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&= \frac{\epsilon}{3(b-a)} (b-a) + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&= \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\
&= \epsilon.
\end{aligned}$$

By definition, $\displaystyle\int_a^b f = L = \lim_{n\to\infty} \int_a^b f_n$. □

7.9 Chapter 7 Summary

DEFINITIONS TO KNOW

Nouns

- A **partition** $\mathcal{P}$ of $[a,b]$ is a finite list of numbers $\{x_0, x_1, ..., x_n\}$ with $a = x_0 < x_1 < x_2 < \cdots < x_n = b$.
  n is the **size** of $\mathcal{P}$; the k^{th} **subinterval** of $\mathcal{P}$ is $[x_{k-1}, x_k]$; the **width** of the k^{th} subinterval is $\Delta x_k = x_k - x_{k-1}$; the **norm** of $\mathcal{P}$ is $\|\mathcal{P}\| = \max\{\Delta x_k : 1 \le k \le n\}$.
- A **tagged partition** $\widehat{\mathcal{P}}$ is a partition together with a list of **test points** $\{c_1, ..., c_n\}$, i.e. points where $c_k \in [x_{k-1}, x_k]$ for all k.
- A **refinement** of partition $\mathcal{P}$ is another partition $\mathcal{Q}$ s.t. $\mathcal{Q} \supseteq \mathcal{P}$ as sets.
  The **join** $\mathcal{P} \vee \mathcal{Q}$ is the least common refinement of $\mathcal{P}$ and $\mathcal{Q}$, i.e. $\mathcal{P} \vee \mathcal{Q} = \mathcal{P} \bigcup \mathcal{Q}$ as sets.
- The **Riemann sum** for $f : [a,b] \to \mathbb{R}$ associated to tagged partition $\widehat{\mathcal{P}}$ is the number $RS(f; \widehat{\mathcal{P}}) = \sum_{k=1}^{n} f(c_k)\Delta x_k$.
- The **upper Riemann sum** for for bounded $f : [a,b] \to \mathbb{R}$ associated to (untagged) partition $\mathcal{P}$ is the number $\mathcal{U}(f; \mathcal{P}) = \sum_{k=1}^{n} w_k \Delta x_k$, where $w_k = \sup\{f(x) : x \in [x_{k-1}, x_k]\}$.
  The **lower Riemann sum** for for bounded $f : [a,b] \to \mathbb{R}$ associated to (untagged) partition $\mathcal{P}$ is the number $\mathcal{U}(f; \mathcal{P}) = \sum_{k=1}^{n} v_k \Delta x_k$, where $v_k = \inf\{f(x) : x \in [x_{k-1}, x_k]\}$.
- If $a < b$, the **Riemann integral of** f **from** a **to** b is a number $\int_a^b f$ so that $\forall\, \epsilon > 0\, \exists\, \delta > 0$ s.t. if $\|\widehat{\mathcal{P}}\| < \delta$, then $\left| RS(f; \widehat{\mathcal{P}}) - \int_a^b f \right| < \epsilon$.
  If $a = b$, then $\int_a^a f = 0$.
  If $a > b$, then $\int_a^b f = -\int_b^a f$.

Adjectives that describe functions $f : \mathbb{R} \to \mathbb{R}$

- f is called **integrable on** $[a,b]$ if $\int_a^b f$ exists (see above).
- f is called **uniformly continuous on** $E \subseteq \mathbb{R}$ if $\forall\, \epsilon > 0\, \exists\, \delta > 0$ s.t. for all $x, a \in E$, $|x - a| < \delta$ implies $|f(x) - f(a)| < \epsilon$.

THEOREMS WITH NAMES

Integrability criteria Suppose $f : [a,b] \to \mathbb{R}$ is bounded. TFAE:

1. f is integrable on $[a,b]$.
2. $\forall \epsilon > 0$, $\exists \delta > 0$ s.t. if $\mathcal{P}$ is a partition of $[a,b]$ with $\|\mathcal{P}\| < \delta$, then $\mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) < \epsilon$;
3. $\forall \epsilon > 0$, $\exists$ partition $\mathcal{P}$ of $[a,b]$ with $\mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) < \epsilon$;
4. $\sup\{\mathcal{L}(f;\mathcal{P}) : \mathcal{P}$ is a part. of $[a,b]\} = \inf\{\mathcal{U}(f;\mathcal{P}) : \mathcal{P}$ is a part. of $[a,b]\}$

When statement (4) holds, both quantities in statement (4) equal $\int_a^b f$.

Max-Min Inequality for Integrals If f is integrable on $[a,b]$ and $m \le f(x) \le M$, then $m(b-a) \le \int_a^b f \le M(b-a)$.

Triangle Inequality for Integrals If f is integrable on $[a,b]$, then so is $|f|$ and

$$\left| \int_a^b f \right| \le \int_a^b |f|.$$

Fundamental Theorem of Calculus (FTC) Part 1 If $f : E \to \mathbb{R}$ is continuous where E is open, if $a \in E$ and $F(x) = \int_a^x f$, then F is differentiable on E and $F' = f$.

Fundamental Theorem of Calculus (FTC) Part 2 If $f : E \to \mathbb{R}$ is continuous where E is open, and if $'f$ is any antiderivative of f on E, then for any $a < b$ in E, $\int_a^b f = 'f(b) - 'f(a)$.

Continuous functions have antiderivatives If $f : E \to \mathbb{R}$ is continuous where E is open, then $\exists$ differentiable function $'f : E \to \mathbb{R}$ which is an antiderivative of f.

(★) Interchange of limit and integral If $\{f_n\}$ is a sequence of integrable functions on $[a,b]$ with $f_n \rightrightarrows f$ on $[a,b]$, then f is integrable on E and $\int_a^b f_n \to \int_a^b f$.

OTHER THEOREMS TO REMEMBER

- Partitions of arbitrarily small norm exist.
- Integrable functions must be bounded.
- Any lower sum of f is less than or equal to any upper sum of f on the same interval.

- Refining a partition decreases the upper sum and increases the lower sum.
- Upper (lower) sums can be approximated arbitrarily well by Riemann sums: $\forall\, \mathcal{P}$ and $\forall\, \epsilon > 0$, $\exists$ tagging $\widehat{\mathcal{P}}$ of $\mathcal{P}$ so that $RS(f;\widehat{\mathcal{P}}) - \mathcal{L}(f;\mathcal{P}) < \epsilon$ and $\exists$ tagging $\widehat{\mathcal{P}}$ of $\mathcal{P}$ so that $\mathcal{U}(f;\mathcal{P}) - RS(f;\widehat{\mathcal{P}}) < \epsilon$.
- Riemann sums and integrals are linear (they preserve $+$, $-$ and constant multiples).
- Integrals are additive: $\displaystyle\int_a^c f = \int_a^b f + \int_b^c f$.
- Monotone functions are integrable.
- Continuous functions are integrable.

  More generally, if f is integrable and ϕ is continuous, then $\phi \circ f$ is integrable.
- Products of integrable functions are integrable.
- A continuous function on a compact domain is automatically uniformly continuous.
- Integrals preserve soft inequalities.

Facts about specific functions

- Dirichlet's function $\mathbb{1}_{\mathbb{Q}}$ is not integrable on any interval of positive length.
- Thomae's function f is integrable on $[0,1]$ and $\displaystyle\int_0^1 f = 0$.

  CAUTION: If f is Thomae's function, then $\left[\displaystyle\int_a^x f\right]' \neq f$.
- The Cantor function c is integrable on $[0,1]$ and $\displaystyle\int_0^1 c = \frac{1}{2}$.
- If $f(x) = \begin{cases} x^2 \sin\frac{1}{x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$, then f is differentiable at 0, but f' is not integrable on any interval containing 0.

  CAUTION: For this function, if $a < 0 < b$ then $\displaystyle\int_a^b f' \neq f(b) - f(a)$.

PROOF TECHNIQUES

To prove that f is integrable on $[a, b]$, do <u>one</u> of these things:

1. Show that f is a sum/difference/product of functions already known to be integrable.

2. Show f is a continuous composition of a function already know to be integrable.

3. Show f is monotone.

4. Show f is continuous.

5. (★) Show f is the <u>uniform</u> limit of integrable functions.

6. Use an integrability criterion: show f is bounded and then let $\forall\, \epsilon > 0$. Come up with a partition $\mathcal{P}$ so that $\mathcal{U}(f;\mathcal{P}) - \mathcal{L}(f;\mathcal{P}) < \epsilon$.

7. Show f is bounded and then take a sequence of partitions $\mathcal{P}_n$ with $\|\mathcal{P}_n\| \to 0$ and show $\lim_{n\to\infty} \mathcal{U}(f;\mathcal{P}) = \lim_{n\to\infty} \mathcal{L}(f;\mathcal{P})$; then apply the result of a HW problem.

8. Use the definition (requires guessing what $\int_a^b f$ is): let $\epsilon > 0$ and come up with $\delta > 0$ so that if $\|\widehat{\mathcal{P}}\| < \delta$, then $\left| RS(f;\widehat{\mathcal{P}}) - \int_a^b f \right| < \epsilon$.

To prove that f is <u>not</u> integrable on $[a, b]$, do <u>one</u> of these things:

1. Show f is unbounded on $[a, b]$

2. Show $\sup_{\mathcal{P}} \mathcal{L}(f;\mathcal{P}) \neq \inf_{\mathcal{P}} \mathcal{U}(f;\mathcal{P})$.

3. Take a sequence of partitions $\mathcal{P}_n$ with $\|\mathcal{P}_n\| \to 0$ and show $\lim_{n\to\infty} \mathcal{U}(f;\mathcal{P}) > \lim_{n\to\infty} \mathcal{L}(f;\mathcal{P})$; then apply the result of a HW problem.

To prove that f is uniformly continuous on E, do <u>one</u> of these things:

1. Show f is continuous and E is compact.

2. Use the definition: let $\epsilon > 0$ and come up with $\delta > 0$ so that $|x - a| < \delta$ implies $|f(x) - f(a)| < \epsilon$.

7.10 Chapter 7 Homework

Exercises from Section 7.1

1. Prove Statement 2 of Theorem 7.5, which says that if $f, g : [a,b] \to \mathbb{R}$ and $\widehat{\mathcal{P}} = \{x_0, ..., x_n\}; \{c_1, ..., c_n\}$ is a tagged partition of $[a,b]$, then $RS(f+g;\widehat{\mathcal{P}}) = RS(f;\widehat{\mathcal{P}}) + RS(g;\widehat{\mathcal{P}})$.

Exercises from Section 7.2

2. Prove the second statement of Lemma 7.11, which says that if $a < b$ and $f : [a,b] \to \mathbb{R}$ is bounded, and if $\mathcal{P} = \{x_0, x_1, ..., x_n\}$ is any partition of $[a,b]$, then for any $\epsilon > 0$, $\mathcal{P}$ can be tagged with test points to create $\widehat{\mathcal{P}}$ so that $\mathcal{U}(f;\mathcal{P}) - RS(f;\widehat{\mathcal{P}}) < \epsilon$.

Exercises from Section 7.3

3. Prove the second inequality in Theorem 7.16, which says that if $f : [a,b] \to \mathbb{R}$ is bounded and $\mathcal{P}$ and $\mathcal{Q}$ are partitions of $[a,b]$ such that $\mathcal{Q} \geq \mathcal{P}$, then $\mathcal{U}(f;\mathcal{P}) \geq \mathcal{U}(f;\mathcal{Q})$.

Exercises from Section 7.4

4. Let $f : [a,b] \to \mathbb{R}$ be bounded. Prove that the following are equivalent:

   a) f is integrable on $[a,b]$.

   b) For any sequence $\{\mathcal{P}_n\}$ of partitions of $[a,b]$ with $\|\mathcal{P}_n\| \to 0$, $\lim\limits_{n\to\infty} \mathcal{U}(f;\mathcal{P}_n) = \lim\limits_{n\to\infty} \mathcal{L}(f;\mathcal{P}_n)$.

   Also, prove that if statement (2) holds, then the common values of these limits is $\displaystyle\int_a^b f$.

5. Let $a < b$. Without using the Fundamental Theorem of Calculus (or any other results after the integrability criteria), prove that $f(x) = x$ is integrable on $[a,b]$ and determine $\displaystyle\int_a^b x$.

   Hint: You may use the result of Exercise 4.

6. Let $a < b$. Without using the Fundamental Theorem of Calculus (or any other results after the integrability criteria), prove that $f(x) = x^2$ is integrable on $[0,b]$ and determine $\displaystyle\int_0^b x^2$.

Hints: You may use the result of Exercise 4, and you may use without proof the summation formula $\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

(P.S. if you are curious how to prove this summation formula, use induction.)

7. Let $f, g : \mathbb{R} \to \mathbb{R}$ be so that f is integrable on $[a, b]$, and let $z \in [a, b]$. Suppose $g(x) = f(x)$ for all $x \in [a, b] - \{z\}$. Prove g is integrable on $[a, b]$ and $\int_a^b g = \int_a^b f$.

Exercises from Section 7.5

8. Prove the second statement of Theorem 7.20, which says that if f is integrable on $[a, b]$ (with $a < b$), then for any $r \in \mathbb{R}$, rf is integrable on $[a, b]$ and

$$\int_a^b (rf) = r \int_a^b f.$$

Exercises from Section 7.6

9. Let $a > 0$ be a constant. Prove that the function $f(x) = \frac{1}{x^2}$ is uniformly continuous on $[a, \infty)$.

10. Prove that $f(x) = \frac{1}{x^2}$ is not uniformly continuous on $(0, \infty)$.

11. Let $E \subseteq \mathbb{R}$. A function $f : E \to \mathbb{R}$ is called **Lipschitz** if there is a constant $K > 0$ so that

$$|f(x) - f(y)| \leq K|x - y|$$

for all $x, y \in E$.

 a) Prove that every Lipschitz function is uniformly continuous on E.
 b) Prove that $f(x) = \sqrt{x}$ is uniformly continuous on $[0, 1]$, but not Lipschitz on $[0, 1]$.

12. Suppose f is continuous and $f(x) \geq 0$ for all $x \in [a, b]$. Prove that if $\int_a^b f = 0$, then $f(x) = 0$ for all $x \in [a, b]$.

Hints: Suppose not, i.e. that there is a $z \in [a, b]$ such that $f(z) > 0$. Explain why this implies that there is a $c \in (a, b)$ (not $[a, b]$) such that $f(c) > 0$. Let $\epsilon = \frac{f(c)}{2}$ and use the uniform continuity of f to find a $\delta > 0$ such that $|x - c| < \delta$ implies $|f(x) - f(c)| < \epsilon$. Then use additivity to split the integral into pieces; show one piece must be strictly positive and use this to derive a contradiction.

13. Show, by providing a specific counterexample with proof, that if $f(x) \geq 0$ for all $x \in [a,b]$ (but f is not assumed continuous), then $\int_a^b f = 0$ does not necessarily imply $f(x) = 0$ for all $x \in [a,b]$.

14. Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable and suppose f' is bounded. Prove that f is uniformly continuous.

15. Give an example of a function $g : [0,1] \to \mathbb{R}$ which is uniformly continuous and differentiable on $(0,1)$, but for which g' is not bounded on $(0,1)$.

Exercises from Section 7.7

16. Prove that if f and g are both continuous and $\int_a^b f = \int_a^b g$, then there is a $c \in [a,b]$ where $f(c) = g(c)$.

17. Prove that if $f : [0,1] \to \mathbb{R}$ is continuous and $\int_0^x f(t)\,dt = \int_x^1 f(t)\,dt$ for all $x \in [0,1]$, then $f(x) = 0$ for all $x \in [0,1]$.

18. Prove the **Mean Value Theorem for Integrals** (not to be confused with the regular Mean Value Theorem), which says that if $a < b$ and if $f : [a,b] \to \mathbb{R}$ is continuous, then there is a $c \in [a,b]$ such that

$$\frac{1}{b-a}\int_a^b f(x)\,dx = f(c).$$

Hint: Apply the Mean Value Theorem to an appropriately defined function g. If you want $g'(c)$ to be $f(c)$, how should g be defined?

19. Prove the **Weighted Law of the Mean**, which says that if $f, g : [a,b] \to \mathbb{R}$ are such that g and fg are integrable on $[a,b]$ and $g(x) \geq 0$ for all $x \in [a,b]$, then $\exists\, c \in \mathbb{R}$ such that

$$\int_a^b fg = c\int_a^b g.$$

20. Suppose $f : [a,b] \to \mathbb{R}$ is continuous and $g : [a,b] \to [0,\infty)$ is integrable on $[a,b]$. Prove that $\exists\, t \in [a,b]$ such that

$$\int_a^b fg = f(t)\int_a^b g.$$

21. In this problem we prove the **Schwarz Inequality** for integrable functions, which says that if $f, g : [a,b] \to \mathbb{R}$ are integrable, then

$$\left(\int_a^b fg\right)^2 \leq \left(\int_a^b f^2\right)\left(\int_a^b g^2\right).$$

a) Prove that for all $t > 0$,

$$2\left|\int_a^b fg\right| \leq t\int_a^b f^2 + \frac{1}{t}\int_a^b g^2.$$

Hint: Consider $\int_a^b (tf+g)^2$ and $\int_a^b (tf-g)^2$. What is true about these two integrals (think in terms of inequalities)?

b) Prove that if $\int_a^b f^2 = 0$, then $\int_a^b fg = 0$.

c) Prove the Schwarz Inequality.

Hints: Most of the time, you can choose a particular value of t in the inequality you proved in part (a) and the Schwarz Inequality will follow from algebra. Sometimes, however, your formula for t won't work–part (b) of this question helps you handle that situation.

22. Prove **Jensen's Inequality** (this name is pronounced "yen-sen"), which says that if $\phi : \mathbb{R} \to \mathbb{R}$ is a differentiable and convex function (see the Chapter 6 homework for a definition of *convex*), then for any $f \in R([0,1])$,

$$\phi\left(\int_0^1 f\right) \leq \int_0^1 \phi \circ f.$$