Math 255: Real Analysis

The real numbers

Course information

- My name is Haim Grebnev
- I'm a postdoctoral scholar
- I research inverse problems, specialty in geometric analysis
- In my spare time I like to…

Undergraduate

Masters student

PhD student

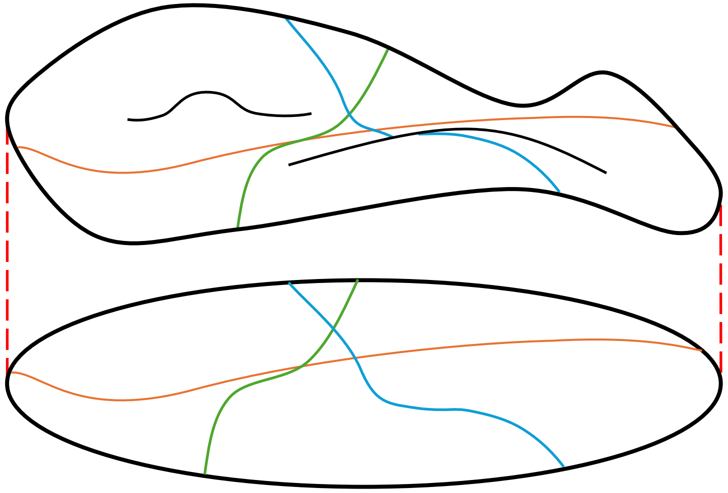
Postdoc

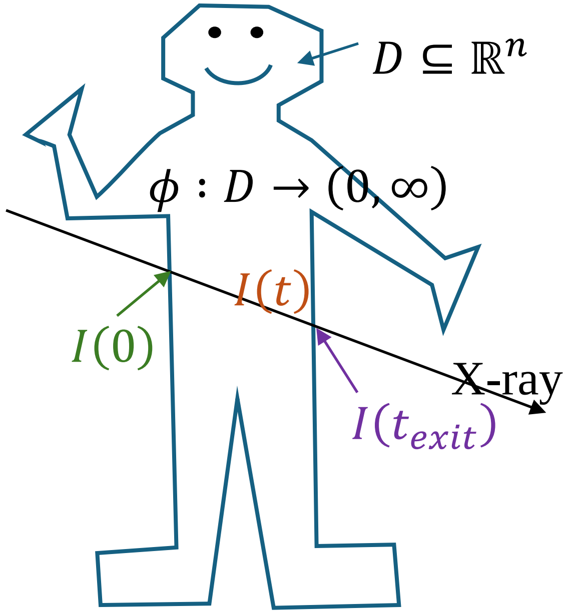
Assistant professor

Associate professor

Full professor

Full professor with tenure





Course information

- **Email:** haim.grebnev@yale.edu
- All materials on Canvas
- Homework assignments due every Friday at 11:59 p.m. (30% of grade)
- Two midterms (40% of grade) and final exam (30% of grade)

Classroom expectations

- Create a welcoming environment!
- Study together (discord channel?)
- Discuss homework together!
- But everyone must write their own solutions!
- No looking up solutions, no use of AI
- Use all the resources available to you
- Don’t be afraid to ask questions!

Homework

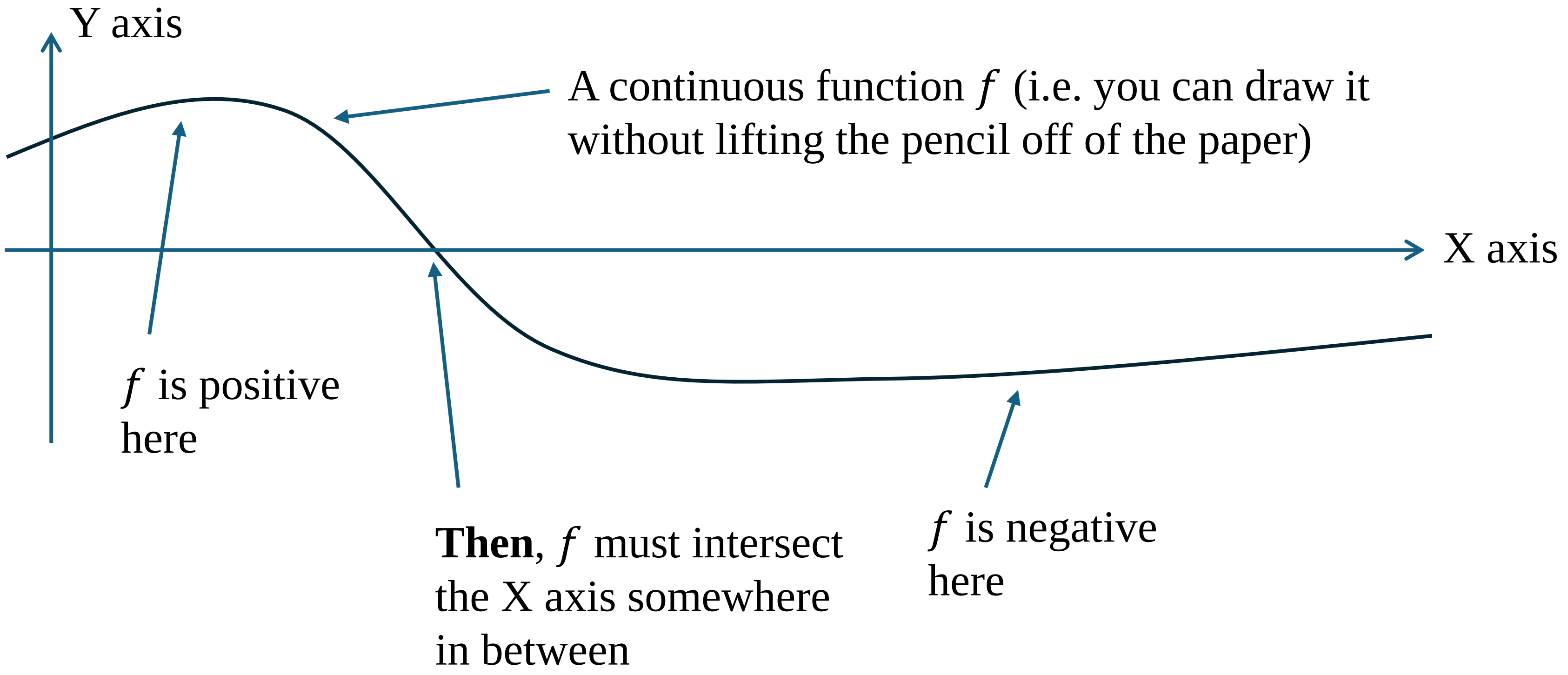
- Will be due on Fridays at 11:59 p.m.
- First homework due Friday Sept. 6th (plan to post it tomorrow)
- Unless stated otherwise, all solutions need to be written as proofs.
- You can only use results proved in class up to that point.
- If in doubt, ask!

What is Real Analysis?

- “Analysis” is an overarching field of math that stems from the idea of the “infinitesimal” (i.e. calculus)
- A few examples are:
  - “Real analysis”
  - “Complex analysis”
  - “Functional analysis”
- We will study “real analysis:” we study calculus while proving everything.
- We don’t emphasize computation

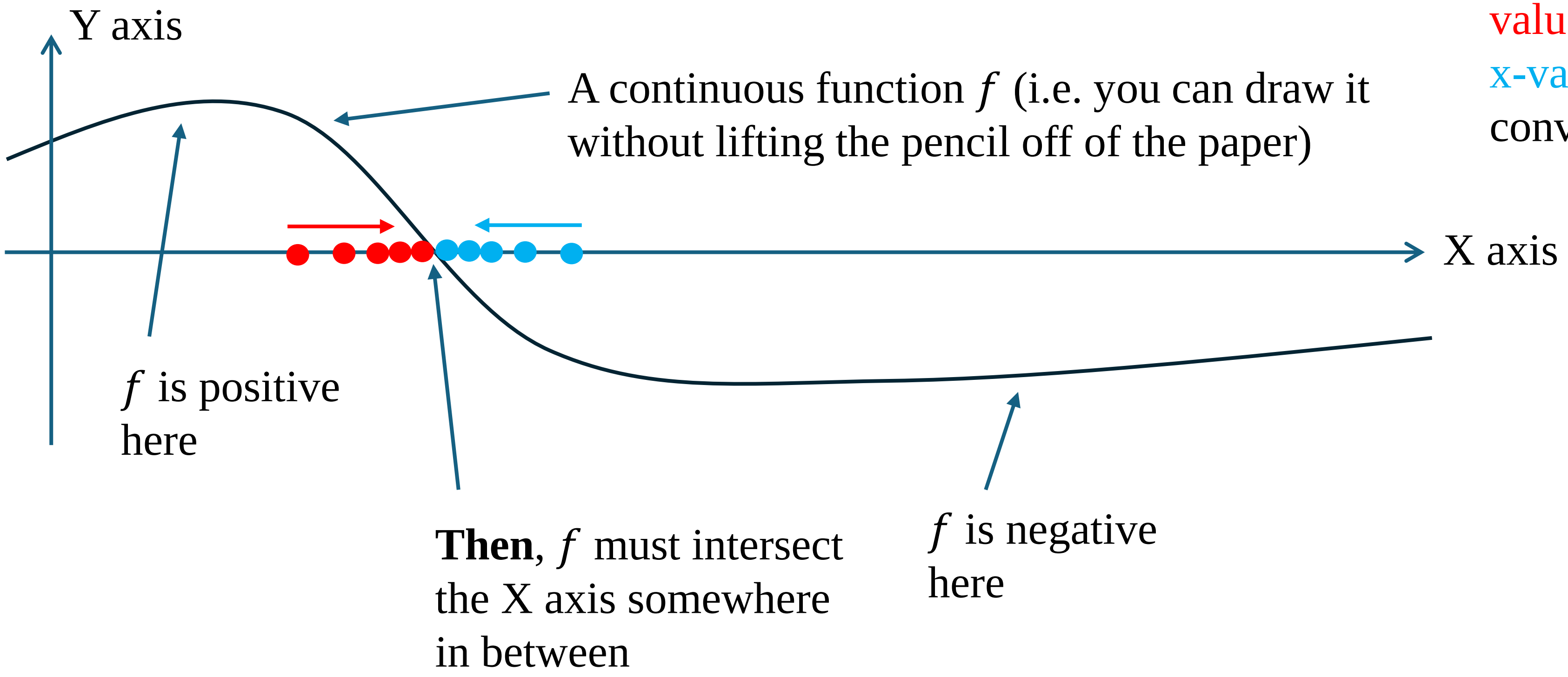
Real numbers

- Because we'll be dealing with the "infinitesimal" on the real number line, we need to study the nature of the real line on the tiniest scale
- A great starting point is to recall the intermediate value theorem:



The need for real numbers

- From a math point of view, this isn't an obvious theorem.
- It's not true if you think of the X-axis and Y-axis only consisting of rational numbers.



- To prove it, take a sequence of x values where f is positive, and x-values where f is negative that converge to where f **is zero**.
  - How do we know that the sequences converge to something?
  - We have to prove that there are no "holes" in the fabric of the "real line."

Defining the real numbers

- How does one define real numbers?
- A natural idea is to define them as infinite decimals:

$$\pi = 3.1415 \ldots, \qquad 5.7000\ldots, \qquad 2.000\ldots$$

- We can write these as sequences of **rational numbers** (i.e. fractions):

$$\frac{3}{1}, \frac{31}{10}, \frac{314}{100}, \frac{3141}{1000}, \frac{31415}{10000}, \ldots \to \pi, \qquad \frac{5}{1}, \frac{57}{100}, \frac{570}{1000}, \frac{5700}{10000}, \ldots \to 5.7000$$

- In other words, the idea is that we can approximate all real numbers by rational numbers better and better

Defining the real numbers

We assume we know this

- Numbers in math are constructed as follows:

Whole numbers
$\mathbb{Z} = \{\dots, -1, -2, -3, 0, 1, 2, 3, \dots\}$

Rational numbers
$\mathbb{Q}$ (e.g. $-\frac{1}{2}, \frac{4}{3}, 5 = \frac{5}{1}$)

Some numbers can't be written as fractions

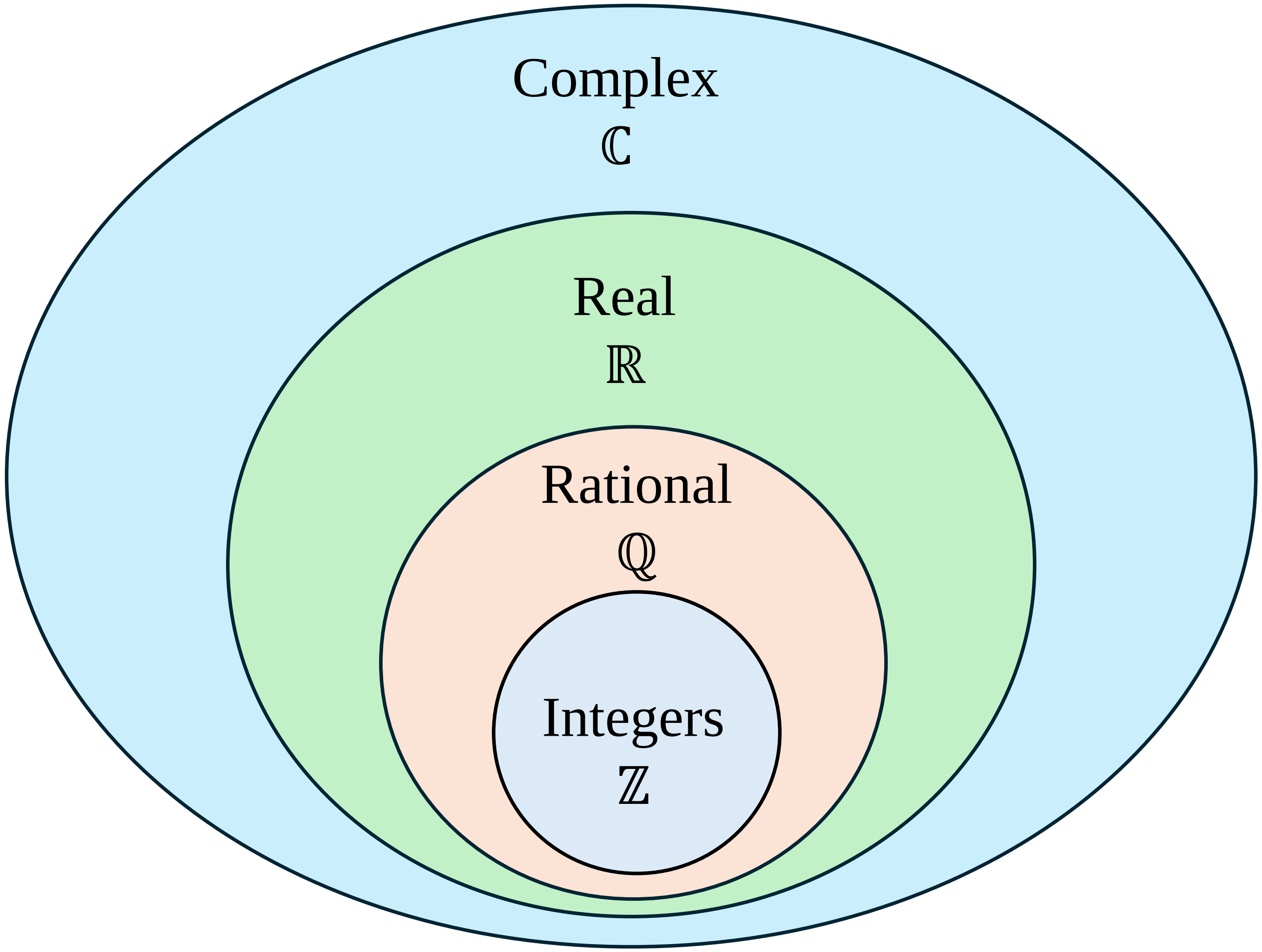
Real numbers
$\mathbb{R}$ (e.g. $3, \frac{3}{7}, \pi, \sqrt{2}$)

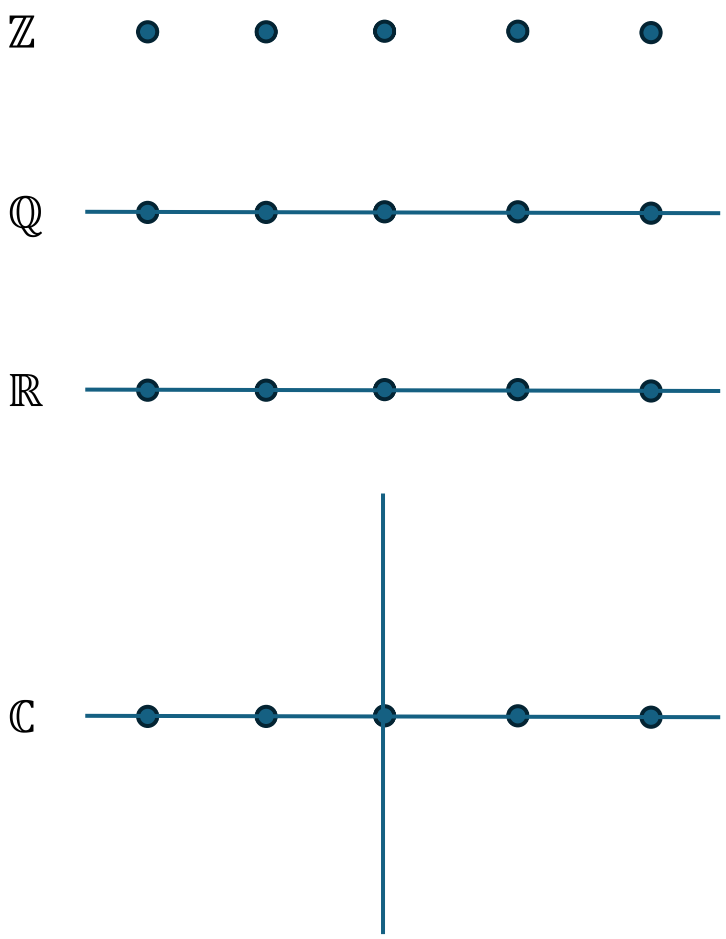
Complex numbers
$\mathbb{C}$ (e.g. $4, 2 + i, \pi - 7i$)

We will define this

We won't worry about this

Defining the real numbers

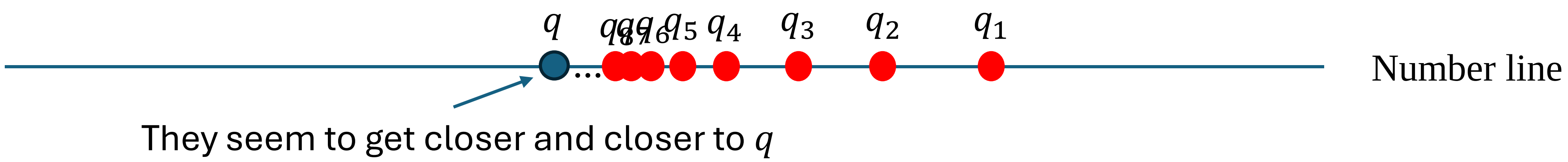




Limits of sequences

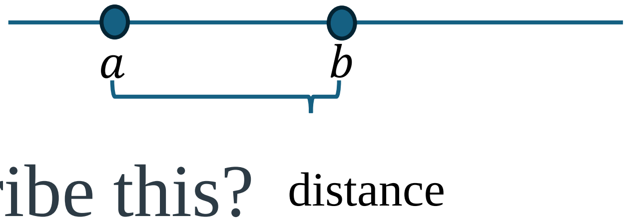
- To define real numbers as better and better approximations of rational numbers, we need notion of **limits**.
- Suppose you have an (infinite) **sequence** of rational numbers $\{q_1, q_2, q_3, \ldots\}$ (i.e. there's infinity many of them).
- For example, $\{q_1 = 1/1\,, q_2 = 1/2\,, q_3 = 1/3\,, q_4 = 1/4\,, q_5 = 1/5\,, \ldots\}$
- Suppose they get closer and closer to some rational number q.
- What does our example sequence above get closer and closer to to?

Answer: zero



Limits of sequences

Recall that on a number line $|a - b|$ is the "distance" between them.



Technically $\varepsilon > 0$ is rational (for now) because we don't yet "know" what real numbers are.

- How can we mathematically describe this?

Definition (Limit): Suppose that you have a sequence of rational numbers $\{q_1, q_2, q_3, \ldots\}$ (there's infinity many of them). We say that the limit of this sequence is q if

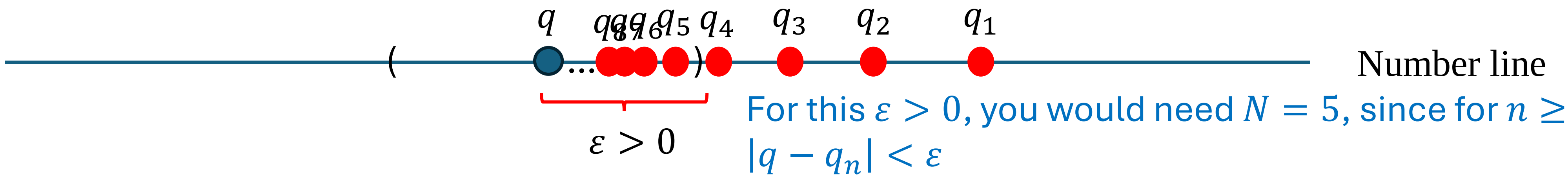
If you go far enough in the sequence

$$\underbrace{\boxed{\text{For any } \varepsilon > 0,}}_{\text{Choose an arbitrary distance}} \boxed{\text{there exists an } N > 0 \text{ such that for any } n > N,}$$

$$\boxed{|q - q_n| < \varepsilon}$$

Afterwards, you'll always be within that distance of the limit.

- We write $\lim_{n\to\infty} q_n = q$



For this $\varepsilon > 0$, you would need $N = 5$, since for $n \geq 5$, $|q - q_n| < \varepsilon$

Defining the real numbers

Definition (Limit): Suppose that you have a sequence of rational numbers $\{q_1, q_2, q_3, \dots\}$ (there's infinity many of them). We say that the limit of this sequence is q if

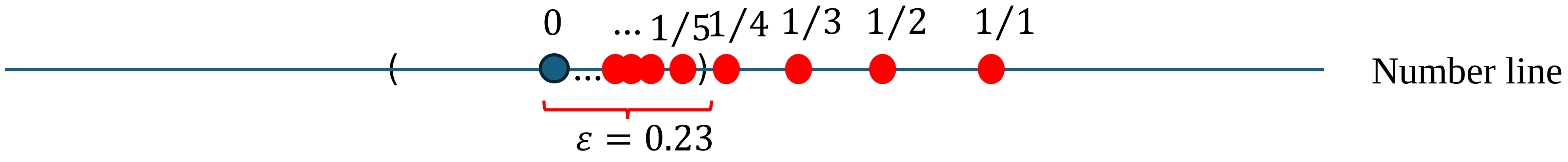
For any $\varepsilon > 0$, there exists an $N > 0$ such that for any $n > N$,
$$|q - q_n| < \varepsilon$$

- Let's try an example. Take $\{q_1 = 1/2\,, q_2 = 1/3\,, q_3 = 1/4\,, q_4 = 1/5\,, \dots\}$. In other words, $q_n = 1/n$.

- It looks like it goes to zero (i.e. it's limit is zero). In other it seems that $q = 0$.

- If $\varepsilon = 0.23$, what $N > 0$ do we need? Answer: $N = 5$, since if $n \geq 5$, then $|0 - q_n| = |0 - 1/n| = 1/n < 0.23$

- If $\varepsilon = 0.05$, what $N > 0$ do we need? Answer: $N = 21$, since if $n \geq 21$, then $|0 - q_n| = |0 - 1/n| = 1/n < 0.05$



Defining the real numbers

Definition (Limit): Suppose that you have a sequence of rational numbers $\{q_1, q_2, q_3, \ldots\}$ (there's infinity many of them). We say that the limit of this sequence is q if

For any $\varepsilon > 0$, there exists an $N > 0$ such that for any $n > N$,

$$|q - q_n| < \varepsilon$$

Play around:

$$|q - q_n| < \varepsilon$$
$$|0 - 1/n| < \varepsilon$$
$$\frac{1}{n} < \varepsilon$$
$$\frac{1}{\varepsilon} < n$$

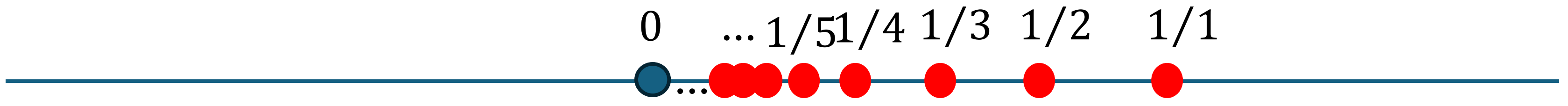
Aha! We need $\frac{1}{\varepsilon} < N$

- Let's try an example. Take $\{q_1 = 1/2\,, q_2 = 1/3\,, q_3 = 1/4\,, q_4 = 1/5\,, \ldots\}$. In other words, $q_n = 1/n$. We have that $q = 0$
- Let's **prove** that $\lim_{n\to\infty} 1/n = 0$ using the definition. Take any $\varepsilon > 0$.

Take any $N > 1/\varepsilon$. This works because for any $n > N > 1/\varepsilon$, by reversing the above work

$$|q - q_n| < \varepsilon$$

This proves that indeed $\lim_{n\to\infty} q_n = \lim_{n\to\infty} 1/n = 0!$ ■



Number line

Defining the real numbers

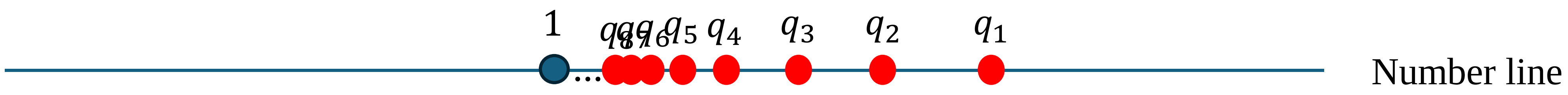
Definition (Limit): Suppose that you have a sequence of rational numbers $\{q_1, q_2, q_3, \ldots\}$ (there's infinity many of them). We say that the limit of this sequence is q if

$$\text{For any } \varepsilon > 0, \text{ there exists an } N > 0 \text{ such that for any } n > N,$$
$$|q - q_n| < \varepsilon$$

- Let's try another example: take the sequence $\left\{q_n = \frac{n}{n+1}\right\}$.
- Computing a few of the terms:

$$q_1 = \frac{1}{2} = 0.5, \qquad q_2 = \frac{2}{3} = 0.\bar{6}, \ldots \quad q_{20} = \frac{20}{21} = 0.9523\ldots, \ldots \quad q_{50} = \frac{50}{51} = 0.9803\ldots$$

- What does it look like it tends to: Answer: 1
- So we guess that the limit is 1, let's prove it!



Defining the real numbers

Definition (Limit): Suppose that you have a sequence of rational numbers $\{q_1, q_2, q_3, \ldots\}$ (there's infinity many of them). We say that the limit of this sequence is q if

For any $\varepsilon > 0$, there exists an $N > 0$ such that for any $n > N$,
$$|q - q_n| < \varepsilon$$

Play around:

$$|q - q_n| < \varepsilon$$
$$\left|1 - \frac{n}{n+1}\right| < \varepsilon$$
$$1 - \frac{n}{n+1} < \varepsilon$$
$$n + 1 - n < (n+1)\varepsilon$$
$$1 < n\varepsilon + \varepsilon$$
$$\frac{1-\varepsilon}{\varepsilon} < n$$

Aha! We need:
$$\frac{1-\varepsilon}{\varepsilon} < N$$

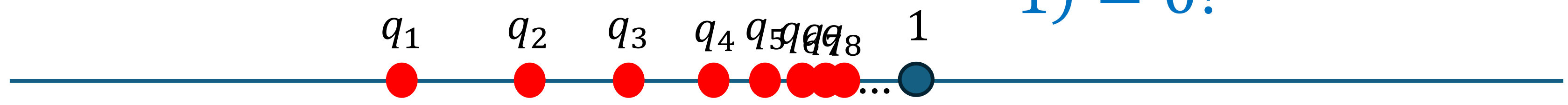
- Let's try another example: take the sequence $\left\{q_n = \frac{n}{n+1}\right\}$. Here $q = 0$.
- Let's **prove** that $\lim_{n\to\infty} \frac{n}{n+1} = 1$ using the definition. Take any $\varepsilon > 0$.

Take any $N > (1-\varepsilon)/\varepsilon$. This works because for any $n > N > (1-\varepsilon)\,/\,\varepsilon$, by reversing the above work

$$|q - q_n| < \varepsilon$$

This proves that indeed $\lim_{n\to\infty} q_n = \lim_{n\to\infty} n\,/\,(n+1) = 0$! ■

Number line



Math 255 Single Variable Analysis (All lectures starting with lecture 2)

Haim Grebnev

Limits continued

- **Theorem (Triangle Inequality for Rational Numbers):** Suppose that a and b are rational numbers. Then

$$|a+b| \le |a| + |b|$$

$$|a-b| \ge \big||a| - |b|\big|$$

Proof: $\sqrt{(a+b)^2} = \sqrt{a^2+2ab+b^2} \le \sqrt{|a|^2+2|a||b|+|b|^2} = \sqrt{(|a|+|b|)^2}$. So $|a+b| \le |a| + |b|$.

For the second one:

$$|a| \le |a-b| + |b| \quad \Rightarrow |a| - |b| \le |a-b|$$

$$|b| \le |b-a| + |a| \quad \Rightarrow |b| - |a| \le |b-a| = |a-b|$$

So $|a-b| \ge \big||a| - |b|\big|$.

■

- **Theorem (Uniqueness of limits):** Suppose $\{q_n\}$ is a sequence of rational numbers with limit rational number q. Then this limit is unique (i.e. you can't have a different rational $\tilde{q}$ that is also a limit of q_n).

Proof: Suppose $\tilde{q}$ is also a limit of $\{q_n\}$. We will show that $\tilde{q} = q$ by showing that for any $\varepsilon > 0$, $|\tilde{q} - q| < \varepsilon$. Take any $\varepsilon > 0$. Playing around:

$$|q - \tilde{q}| < |q - q_n| + |q_n - \tilde{q}|$$

There exists an $N_1 > 0$ such that for all $n > N_1$, $|q_n - q| < \varepsilon/2$ and for all $n > N_2$, $|q_n - \tilde{q}| < \varepsilon/2$. Take $N = \max\{N_1, N_2\}$. Then for all $n > N$,

$$|q - q_n| + |q_n - \tilde{q}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

and so $|q - \tilde{q}| < \varepsilon$.

■

- **Theorem (Operations with Limits):** Suppose that $\{q_n\}$ and $\{\tilde{q}_n\}$ are sequences of rational numbers with limits q and $\tilde{q}$ respectively (i.e. $\lim_{n\to\infty} q_n = q$ and $\lim_{n\to\infty} \tilde{q}_n = \tilde{q}$). Then
  - $\lim\limits_{n\to\infty} (q_n + \tilde{q}_n) = q + \tilde{q}$.
  - Same holds if you change "$+$" above to "$-$," "$\times$," and "$\div$" (last if $\tilde{q} \ne 0$).

Proof: Let's prove the "+" version. The rest are an exercise. Take any $\varepsilon > 0$. We need to show that there exists $N > 0$ such that for all $n > N$,

$$|(q_n + \tilde{q}_n) - (q + \tilde{q})| < \varepsilon,$$

Playing around:

$$|(q_n + \tilde{q}_n) - (q + \tilde{q})| = |(q_n - q) + (\tilde{q}_n - \tilde{q})| \leq |q_n - q| + |\tilde{q}_n - q|.$$

Because $\lim_{n\to\infty} q_n = q$ and $\lim_{n\to\infty} \tilde{q}_n = \tilde{q}$, there exist $N_1 > 0$ and $N_2 > 0$ such that for all $n > N_1$, $|q_n - q| < \varepsilon/2$ and for all $n > N_2$, $|\tilde{q}_n - \tilde{q}| < \varepsilon/2$. Take $N = \max\{N_1, N_2\}$. Then for all $n > N$,

$$|q_n - q| + |\tilde{q}_n - q| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

and so $|(q_n + \tilde{q}_n) - (q + \tilde{q})| < \varepsilon$.

■

Defining Real Numbers

- Recall the idea of defining real numbers as limits of rational numbers:

$$\frac{3}{10}, \frac{31}{100}, \frac{314}{1000}, \ldots \to \pi$$

- A few problems:
  - We only defined limits if the limit is rational, not a real number in general. So we need a notion of a sequence that "looks like" it has a limit.
  - Different sequences of rationales can converge to the same real number, for instance

    $\{2,2,2,2, \ldots\}$, $\{2.1, 2.01, 2.001, 2.0001, \ldots\}$, $\{1.9, 1.99, 1.999, 1.9999, \ldots\}$ all converge to 2.

- Cauchy came up with a notion of a sequence that "looks like" it has a limit:
- **Definition (Cauchy Sequence):** We say that a sequence of rational numbers $\{q_n\}$ is a **Cauchy sequence** if

  For any $\varepsilon > 0$, there exists an $N > 0$ such that for any $n > N$, $|q_N - q_n| < \varepsilon$

  (Alternative version): For any $\varepsilon > 0$, there exists an $N > 0$ such that for any $m, n > N$,

$$|q_m - q_n| < \varepsilon$$

- **Remark:** In the homework, you'll prove that these are equivalent definitions.
- **Definition (Set):** A **set** is a collection of elements (i.e. objects).
- **Definition (Real Numbers):** A real number "r" is a set of Cauchy sequences of rational numbers with the following properties

- If you take any two (Cauchy) sequences $\{q_n\}$ and $\{\tilde{q}_n\}$ in r, they must be equivalent in the sense

$$\lim_{n\to\infty}(q_n - \tilde{q}_n) = 0.$$

- The set r contains all Cauchy sequences of rational numbers equivalent to what it contains (i.e. if you think of another sequence $\{\tilde{q}_n\}$ that is equivalent to $\{q_n\}$ already in r, then $\{\tilde{q}_n\}$ is actually in r as well).

We denote the set of all real numbers by $\mathbb{R}$.

- **Notation:** If S_1 and S_2 are sets, then
  - $S_1 \cap S_2$ denotes their intersection.
  - $S_1 \cup S_2$ denotes their union.
  - $\emptyset$ denotes the empty set.
  - $S_1 \subseteq S_2$ means S_1 is contained in S_2.
  - If a is an element, then $a \in S_1$ means "a is in S_1." If a is <u>not</u> in S_1, we write $a \notin S_1$.
- **Proposition:** Different real numbers cannot have shared Cauchy sequences. In other words, if $r, \tilde{r} \in \mathbb{R}$ are not equal, then $r \cap \tilde{r} = \emptyset$.

Proof: Suppose that r are $\tilde{r}$ are real and have the same Cauchy sequence of rational numbers $\{q_n\}$. We will show that $r = \tilde{r}$. Take any Cauchy sequences of rational numbers $\{p_n\} \in r$ and $\{\tilde{p}_n\} \in \tilde{r}$. Then

$$\lim_{n\to\infty}(p_n - \tilde{p}_n) = \lim_{n\to\infty}(p_n - q_n) + \lim_{n\to\infty}(q_n - \tilde{p}_n) = 0 + 0 = 0.$$

So $\{p_n\}$ and $\{\tilde{p}_n\}$ are "equivalent." So $\{\tilde{p}_n\} \in r$ and $\{p_n\} \in \tilde{r}$. Since $\{p_n\}$ and $\{\tilde{p}_n\}$ were arbitrarily chosen, this shows that $r = \tilde{r}$.

■

- **Corollary:** For every Cauchy sequence of rational numbers $\{q_n\}$, there exists one and only one real number r containing it.
- We think of rational numbers sitting in real numbers as follows: if q is a rational number, we think of it as sitting in the real numbers as the set containing the constant Cauchy sequence $\{q, q, q, \dots\}$. For instance, we think of 0 as the set containing $\{0,0,0,\dots\}$.
- **Definition (Operations on Reals):** Suppose that r_1 and r_2 are real numbers. Take any Cauchy sequences of rational numbers $\{p_n\} \in r$ and $\{q_n\} \in r_2$ in them respectively.
  - We define $r_1 + r_2$ as the real number containing the Cauchy sequence $\{p_n + q_n\} = \{p_1 + q_1, p_2 + q_2, p_3 + q_3, \dots\}$.

- We define $r_1 - r_2$, $r_1 \cdot r_2$, $r \div r_2$ (if $r_2 \neq 0$) by replacing "$+$" with "$-$," "$\cdot$," and "$\div$" respectively in the above statement.
- We define $|r_1|$ as the real number containing the Cauchy sequence $\{|p_n|\}$
- We define $r_1 \leq r_2$ if $r_1 = r_2$ or eventually $p_n \leq q_n$ (i.e. there exists an $N > 0$ for all $n > N$, $p_n \leq q_n$).
- We define $r_1 < r_2$ if $r_1 \leq r_2$ and $r_1 \neq r_2$.

* **Proposition:** *The first three "∘" points above are well-defined (i.e. the stated sequences are indeed Cauchy and that the definitions don't depend on the $\{p_n\} \in r$ and $\{q_n\} \in r_2$ that we choose).*

  Proof: We'll prove that the third point is well-defined, the rest are an exercise. First we show that $\{|p_n|\}$ is Cauchy. Take any $\varepsilon > 0$. Playing around:

$$\big||p_N| - |p_n|\big| \leq |p_N - p_n|$$

  Since $\{p_n\}$ is Cauchy, there exists an $N > 0$ such that for $n > N$, $|p_N - p_n| < \varepsilon$. So by the above $\big||p_N| - |p_n|\big| < \varepsilon$. So indeed $\{|p_n|\}$ is Cauchy.

  Now suppose we choose another $\{\tilde{p}_n\} \in r_1$. We have to show that the real number containing $\{|\tilde{p}_n|\}$ is the same one that contains $\{|p_n|\}$ so that the definition of $|r_1|$ does not depend on the Cauchy sequence we choose. This will follow if we show that $\{|\tilde{p}_n|\}$ is equivalent to $\{|p_n|\}$, or in other words that $\lim_{n\to\infty}\big||p_n| - |\tilde{p}_n|\big| = 0$. Take any $\varepsilon > 0$. By the triangle inequality,

$$\big||p_n| - |\tilde{p}_n|\big| \leq |p_n - \tilde{p}_n|$$

  Since $\{p_n\}$ and $\{\tilde{p}_n\}$ are both in r_1, they are equivalent and so $\lim_{n\to\infty}|p_n - \tilde{p}_n| = 0$. So there exists an $N > 0$ such that for all $n > N$, $|p_n - \tilde{p}_n| < \varepsilon$. The above inequality then implies that $\big||p_n| - |\tilde{p}_n|\big| < \varepsilon$. So indeed $\lim_{n\to\infty}\big||p_n| - |\tilde{p}_n|\big| = 0$.

  ■

* **Theorem (Triangle Inequality for Real Numbers):** *Suppose that a and b are real numbers. Then*

$$|a + b| \leq |a| + |b|$$

$$|a - b| \geq \big||a| - |b|\big|$$

  Proof: Take any Cauchy sequences of rational numbers $\{p_n\} \in a$ and $\{q_n\} \in b$. We have by the triangle inequality for rational numbers that

$$|p_n + q_n| \leq |p_n| + |q_n|$$

  The left-hand side is the Cauchy sequence for $|a + b|$ and the right hand-side is the Cauchy sequence for $|a| + |b|$. So this not only "eventually holds," but always holds. So we conclude that $|a+b| \leq |a| + |b|$ by the above definition (see "Operations on Reals"). The second inequality follows similarly.

Completeness of real numbers

- Sequences of real numbers have the property that if they "look like" they have a limit, then they have a limit.
  - This is called **completeness**, in other words there are no "holes" in the "fabric" of the real numbers.
- **Remark:** We define limits of sequences of real numbers and Cauchy sequence of real numbers the same way as we did for rational numbers above. The theorem of "uniqueness of limits" and "operations on limits" also holds for real numbers with the same proofs.
- **Lemma:** *If you take any two distinct real numbers $r_1 \neq r_2$, then*
  - *either $r_1 < r_2$ or $r_1 > r_2$,*
  - *there always exists a rational number q between them (i.e. $r_1 < q < r_2$ or $r_1 > q > r_2$).*

  Proof: Homework or TA discussion section.
- Now we prove what intuition has been telling us all along: a Cauchy sequence that represents a real number actually converges to it. Note that it does not fall directly out of the definition since it relies on the nontrivial fact, provided by the above theorem, that for any real number $\varepsilon > 0$, there exists a rational number $\varepsilon' \in \mathbb{Q}$ satisfying $0 < \varepsilon' < \varepsilon$.
- **Lemma:** *If $\{p_n\} \in r$, then $r = lim_{n\to\infty} p_n$.*

  Proof: Take any $\varepsilon > 0$. By the previous lemma, we know that there exists a rational $\varepsilon' > 0$ so that $0 < \varepsilon' < \varepsilon$. We have to show that

  (1) $$\exists N > 0 \;\; \forall n > N, \;\; |p_n - r| \leq \varepsilon' < \varepsilon.$$

  By "Operations on Reals" part 4 this inequality will follow if we show that [1]

  (2) $$\exists K > 0 \;\; \forall k > K, \;\; |p_n - p_k| < \varepsilon'.$$

  Since $\{p_n\}$ is Cauchy, we know that

  $$\exists N' > 0 \;\; \forall n', k' > N', \;\; |p_{n'} - p_{k'}| < \varepsilon'.$$

  Notice then that setting $N = N'$ and $K = N'$ will make (1) and (2) hold true, proving the theorem.

  ■
- **Theorem (Completeness of Reals):** Suppose that $\{r_n\} = \{r_1, r_2, r_3, \dots\}$ is a Cauchy sequence of real numbers. Then this sequence has a limit r (i.e. there exists an r such that $\lim_{n\to\infty} r_n = r$).

[1] We're technically using here that ε' is rational here because then ε' contains the sequence $\{\varepsilon', \varepsilon', \varepsilon', \dots\}$.

Proof: Take a Cauchy sequence $\{r_n\}_{n=1}^{\infty}$ of real numbers. We will construct the limit $r \in \mathbb{R}$. There are many ways to do this proof: here is mine. This is called a "diagonal argument."

For any $n > 0$, take any $\{p_{n,k}\}_{k=1}^{\infty} \in r_n$. Fix any $n > 0$ and consider for the moment $\varepsilon_n = 1/n$. Since each $\{p_{n,k}\}_{k=1}^{\infty}$ is Cauchy there exists K_n such that for any $k > K_n$, $|p_{n,K_n} - p_{n,k}| < \varepsilon_n = 1/n$. Define the sequence of rational numbers $\{q_n\}_{n=1}^{\infty}$ to be $q_n = p_{n,K_n}$. We will show that $\{q_n\}_{n=1}^{\infty}$ is Cauchy and hence there exists an $r \in \mathbb{R}$ containing it (spoiler: this is the r we're looking for).

Take any rational $\varepsilon > 0$. We're aiming for $|q_N - q_n| < \varepsilon$. Since $\{r_n\}_{n=1}^{\infty}$ is Cauchy, there exists an $N > 0$ such that for any $n > N$

$$|r_N - r_n| < \frac{\varepsilon}{3}$$

By "Operations on Reals" part 4, there exists a $K_{N,n}$ such that if $k > K_{N,n}$ [2]

$$|p_{N,k} - p_{n,k}| \leq \frac{\varepsilon}{3}$$

Now, we have that for all $k > 0$

$$|q_N - q_n| = |p_{N,K_N} - p_{n,K_n}| \leq |p_{N,K_N} - p_{N,k}| + |p_{N,k} - p_{n,k}| + |p_{n,k} - p_{n,K_n}|$$

Then, if we set $K = \max\{K_N, K_{N,n}, K_n\}$, then for $k > K$ the right-hand side is bounded from above by

$$\frac{1}{N} + \frac{\varepsilon}{3} + \frac{1}{n} < \frac{1}{N} + \frac{\varepsilon}{3} + \frac{1}{N}$$

Suppose we made $N > 0$ earlier big enough so that $1/N < \varepsilon'/3$. Then we get that

$$|q_N - q_n| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

And so indeed $\{q_n\}_{n=1}^{\infty}$ is Cauchy. Let $r \in \mathbb{R}$ be the real number that contains it.

We will show that $\lim_{n\to\infty}\{r_n\}_{n=1}^{\infty} = r$. Take any $\varepsilon > 0$. By a previous lemma, we know that there exists a rational $\varepsilon' > 0$ so that $0 < \varepsilon' < \varepsilon$. We need to find an $N > 0$ such that for all $n > N$

$$|r - r_n| \leq \varepsilon' < \varepsilon$$

By "Operations on Reals" part 4 this inequality holds for any n if there exists a K such that for $k > K$

$$|q_k - p_{n,k}| < \varepsilon'.$$

---

[2] We're technically using here that $\varepsilon/3$ is rational here because then $\varepsilon/3$ contains the sequence $\{\varepsilon/3\,, \varepsilon/3\,, \varepsilon/3\,, \dots\}$

We have that

$$\left|q_k - p_{n,k}\right| \le |q_k - q_N| + |q_N - q_n| + \left|q_n - p_{n,k}\right| = |q_k - q_N| + |q_N - q_n| + \left|p_{n,K_n} - p_{n,k}\right|$$

Because $\{q_n\}$ is Cauchy, there exists $N > 0$ be such that for any $n > N$, $|q_N - q_n| < \varepsilon'/3$. Increase N if necessary so that $1/N < \varepsilon'/3$. Take any $n > N$. Let $K = \max\{N, K_n\}$. Then for any $k > K$, the above is bounded above by

$$\frac{\varepsilon'}{3} + \frac{\varepsilon'}{3} + \frac{1}{n} < \frac{2\varepsilon'}{3} + \frac{1}{N} < \frac{2\varepsilon'}{3} + \frac{\varepsilon'}{3} = \varepsilon'.$$

So $\left|q_k - p_{n,k}\right| < \varepsilon'$. As discussed above, this shows that indeed $\lim_{n\to\infty}\{r_n\}_{n=1}^{\infty} = r$.

■

- We can go back to thinking about real numbers normally.
- **Notation:** The symbol "∃" means "there exists" and "∀" means "for all." The symbol ":" (or "|") means "such that."
- **Definition (Supremum/infimum):** Suppose that S is a nonempty set of real numbers. If S is bounded from above, then its **supremum** is the smallest real number r that is ≥ all numbers in S (also called **least upper bound (LUB)**) and is denoted by

  $$\sup S.$$

  If S is bounded from below, then its **infimum** is the biggest real number r that is ≤ than all numbers in S (also called **greatest lower bound (GLB)**) and is denoted by

  $$\inf S.$$

- **Lemma:** *Suppose that $\{x_n\}$ and $\{y_n\}$ are sequences of numbers with limits x and y respectively. If each $x_n \le y_n$, then $x \le y$.*

  Proof: Either homework problem or TA discussion.

- **Theorem:** *The above definition is well defined (i.e. the real numbers sup S and inf S actually exist in the situations described).*

  Proof: We'll do "sup," "inf" is left as an exercise. We construct the sequences of real numbers

  $$\{r_n \in S\} \quad \text{and} \quad \{u_n \text{ each is an upper bound for } S\}$$

  as follows. Take any $r_1 \in S$ and any upper bound u_1 for S. Let their distance be $L = |u_1 - r_1|$. Take their center $x = (r_1 + u_1)/2$.

  1. If x is an upper bound, set $u_2 = x$ and $r_2 = r_1$.
  2. If x is not an upper bound, take any $r_2 \in S$ such that $r_2 > x$ and set $u_2 = u_1$.

Do this again: take the center $x = (r_2 + u_2)/2$ and proceed as in 1) and 2) but replace "1" with "2" and "2" with "3" to get r_3 and u_3. Repeat iteratively to get the sequences $\{r_n\}$ and $\{u_n\}$. It's not hard to see by induction that:

- $|u_n - r_n| \leq \frac{L}{2^n}$.
- $\{r_n\}$ is nondecreasing (i.e. $r_1 \leq r_2 \leq r_3 \leq \cdots$) and $\{u_n\}$ is nonincreasing (i.e. $u_1 \geq u_2 \geq u_3 \geq \cdots$).
- Observe that for any $0 < N < n$

$$|r_n - r_N| = r_n - r_N \leq u_N - r_N = |u_N - r_N| < L/2^N$$

  which you can quickly use to prove that $\{r_n\}$ is Cauchy.
- Similarly use that for any $0 < N < n$

$$|u_N - u_n| = u_N - u_n \geq u_N - r_N = |u_N - r_N| < L/2^N$$

  to prove that $\{u_n\}$ is Cauchy.

So by completeness of real numbers, $\{r_n\}$ and $\{u_n\}$ have limits:

$$r = \lim_{n\to\infty} r_n \quad \text{and} \quad u = \lim_{n\to\infty} u_n.$$

Let's show that $r = u$. We will show this by showing that $\forall \varepsilon > 0$, $|r - u| < \varepsilon$. Playing around

$$|r - u| < |r - r_n| + |r_n - u_n| + |u_n - u|.$$

Now, $\exists N > 0 : \forall n > N$, all three terms on the right are $< \varepsilon/3$ and so $|r - u| < \varepsilon$. So indeed $r = u$.

Let's show that $r = u$ is $\sup S$. Let's show that u is bigger than everything in S. Take any $x \in S$. Since each $x \leq u_n$, by the previous lemma $x \leq u$ (you can use the constant sequence $\{x_n = x\}$ in that lemma if you wish). Next let's show that r is the smallest upper bound. Suppose y was another upper bound. Since each $r_n \leq y$, $r \leq y$. So indeed $r = u$ is $\sup S$.

■

Proposition: If S is nonempty, $\inf S \leq \sup S$.

Proof: For any $r \in S$, $\inf S \leq r \leq \sup S$.

■

- **Theorem (Monotone Convergence Theorem):** *Suppose that $\{x_n\}$ is a nondecreasing sequence of real numbers (i.e. $x_1 \leq x_2 \leq x_3 < \cdots$) and is bounded from above (i.e. there is a number u bigger than all the x_n's). Then this sequence has a limit, and in fact*

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

(technical point: on the left-hand side $\{x_n\}$ is thought of as a set, not a sequence). If $\{x_n\}$ is nonincreasing, then the same holds but change "sup" to "inf."

Proof: Homework or TA discussion session

Uncountability of the Reals

- Turns out that there are many types of "infinities:" some are bigger than others. In math we refer to this "size of the infinities" as the **cardinalities**.
- **Definition:** Suppose that A and B are sets. A map $f : A \to B$ is a rule that for every element $a \in A$ outputs an element in B which we denote by $f(a)$ (draw a picture!). We say that
  - f is **injective** (or "**into**") if $f(a) = f(b)$ implies that $a = b$ (i.e. f never hits anything twice)
  - f is **surjective** (or "**onto**") if you for every $b \in B$, there exists an $a \in A$ such that $f(a) = b$ (i.e. f hits everything).
  - f is **bijective** (or "**one-to-one**") if it is both injective and surjective.
- **Example:** Take the maps $f_1, f_2, f_3 : \mathbb{Z} \to \mathbb{Z}$ given by

$$f_1(n) = n + 1, \quad f_2(n) = 2n, \quad f_3(n) = \begin{cases} 0 & \text{if } n \text{ is odd} \\ n/2 & \text{if } n \text{ is even} \end{cases}$$

  Which of these are bijective, injective, and surjective (*answer*: they are bijective, injective, and surjective respectively).
- **Definition:** Suppose that A and B are sets.
  - If there exists a bijective map $f : A \to B$, then we say that A and B have the **same cardinality**.
  - If there exists an injective map $f : A \to B$, but there *does not* exist a surjective map $g : A \to B$, then we say that the cardinality of A is smaller than the cardinality of B.
- **Definition:** If a set has a bijection with $\mathbb{Z}_+ = \{1,2,3, \ldots\}$, it is called **countable**. If not, then it is **uncountable**.
- **Theorem:** The sets $\mathbb{Z}_+$ and $\mathbb{Q}$ have the same cardinality (i.e. $\mathbb{Q}$ is countable).

  Proof: Either TA discussion or homework.

  Theorem: The cardinality of $\mathbb{Z}_+$ is smaller than that of $\mathbb{R}$ (i.e. $\mathbb{R}$ is uncountable).

  Proof: The map $f : \mathbb{Z}_+ \to \mathbb{R}$ given by $f(n) = n$ is injective. So let's show that a surjective map $g : \mathbb{Z}_+ \to \mathbb{R}$ does not exist. Suppose that it does. Write it out as the following in decimal form (the right-hand sides below are just an example for demonstration) but don't write any number in the form where it ends in "9" repeated

$$g(0) = 0.163528464$$

$$g(2) = 34.847578403 83$$

$$g(3) = 7.758104731$$

$$g(4) = 4.348710743$$

$$g(5) = 0.0037417209$$

$$g(6) = 9.538753892$$

$$g(7) = 11.00000000000$$

$$\vdots$$

Align the right so that the decimal point "." are on top of each other so that everything is aligned as above. Go along the diagonal as above to construct a number, in the above example it is:

$$0.858750 \ldots$$

Now if the digit is not a 5, turn into a 5. If the digit is a 5, turn it into a 4.

$$5.545545 \ldots$$

But this won't be a number that g hits, which contradicts that it's a surjection. ■

Infinite Limits

- **Definition:** *Suppose that $\{x_n\}$ is a sequence of numbers. We say that x_n goes to $+\infty$*

$$\forall L \in \mathbb{R} \quad \exists N > 0 \quad \forall n > N, \quad x > L$$

In this case we can write $\infty = \lim_{n\to\infty} x_n$. We say that x_n goes to $-\infty$ if

$$\forall L \in \mathbb{R} \quad \exists N > 0 \quad \forall n > N, \quad x < L.$$

In this case we can write $-\infty = \lim_{n\to\infty} x_n$. In both cases we still say that $\{x_n\}$ do not converge.

- **Example:** $\{x_n = n^2\}$ goes to $+\infty$ and $\{y_n = -n^2\}$ goes to $-\infty$.

Let's prove that $\{x_n = n^2\}$ goes to $+\infty$, the other is left as an exercise. Take $\forall L \in \mathbb{R}$. Take any $N > L$ such that $N > 1$. Then $\forall n > N$,

$$x_n = n^2 > N^2 > N > L.$$

So indeed $\{x_n = n^2\}$ goes to $+\infty$. //

Euclidean space

- **Definition:** *The space $\mathbb{R}^n$ is the set of all n-tuples of real numbers:*

$$v = (v_1, v_2, \ldots, v_n)$$

The real numbers $v_1, v_2, \ldots, v_n$ are called the ***components*** *of v. We often call elements of $\mathbb{R}^n$ (such as v) "points" or "vectors" in $\mathbb{R}^n$.*

- For example, we live in $\mathbb{R}^3$ where a position in space could be $v = (3,4,6)$. //
- **Example:** The real line $\mathbb{R}$ is just $\mathbb{R}^1$.
- The reason we call elements of $\mathbb{R}^n$ "vectors" is because that's how we called them in physics. In math we call them vectors because $\mathbb{R}^n$ is easily shown to be a vector space.
- We're going to generalize the "analysis" that we've developed on $\mathbb{R}$ so far from the real line to $\mathbb{R}^n$. The basic notion that will get us started is defining distance:
- **Definition:** *For any vector $v \in \mathbb{R}^n$. We define its length to be*

$$\|v\| = \sqrt{(v_1)^2 + \cdots + (v_n)^2}.$$

We will often not write the double bar "$\|\cdot\|$" and just write "$|\cdot|$."

- **Theorem:** For any vectors $v, w \in \mathbb{R}^n$, the **triangle inequalities** holds:

$$|v + w| \leq |v| + |w|, \quad |v - w| \geq \big||v| - |w|\big|$$

And Cauchy's inequality holds:

$$|v \cdot w| \leq |v||w|$$

Proof: The triangle inequalities are proved exactly the same way as for numbers except interpret multiplication as a dot product. We won't prove Cauchy's inequality since we won't really use it, and you most likely proved it in a linear algebra course.

■

- **Definition:** *Suppose that $\{x_k\}$ is a sequence of points in $\mathbb{R}^n$ and that $x \in \mathbb{R}^n$. We say that x is the **limit** of $\{x_k\}$ if*

$$\forall \varepsilon > 0 \;\; \exists K > 0 \;\; \forall k > K, \;\; |x - x_k| < \varepsilon.$$

- **Definition:** *Suppose that $x \in \mathbb{R}^n$ and that $r > 0$ is a real number. The **(open) ball of radius r centered at x** is the set*

$$B_{\mathbb{R}^n}(x; r) = \{y \in \mathbb{R}^n : |x - y| < r\}$$

- In other words, a sequence $\{x_k\}$ in $\mathbb{R}^n$ has the limit x if for arbitrary small radius $\varepsilon > 0$, $\{x_k\}$ eventually enters the ball of radius ε centered at x.
- Although dealing with limits in multiple variables at the same time looks daunting, fortunately we can often deal with it component wise.
- **Proposition:** *Suppose that $\{x_k\}$ is a sequence in $\mathbb{R}^n$. Write each x_k as*

$$x_k = ((x_k)_1, \dots, (x_k)_n)$$

Suppose that each component sequence has a limit y_j separately;

$$y_1 = lim_{k\to\infty}((x_1)_k),$$

$$\vdots$$

$$y_n = lim_{k\to\infty}((x_n)_k).$$

Then if we take the point $y = (y_1, \dots, y_n)$, then point y is the limit of $\{x_k\}$ (i.e. $y = lim_{k\to\infty} x_k$).

Proof: Take any $\varepsilon > 0$. We have that

$$|x_k - y| = \sqrt{((x_1)_k - y_1)^2 + \cdots + ((x_n)_k - y_n)^2}.$$

Note that all of the "$(\,\cdot\,)^2$" above can be changed to "$|\cdot|$." Because each $y_j = \lim_{k\to\infty}((x_1)_k)$, for any $j = 1, \dots, n$ there exists an $K_j > 0$ such that for all $k > K_j$, $\left|(x_j)_k - y_j\right| < \varepsilon/\sqrt{n}$. Set $K = \max\{K_1, \dots, K_n\}$ and note that for all $k > K$ the right-hand side above is less than

$$< \sqrt{\left(\frac{\varepsilon}{\sqrt{n}}\right)^2 + \cdots + \left(\frac{\varepsilon}{\sqrt{n}}\right)^2} = \varepsilon.$$

So indeed $y = \lim_{k\to\infty} x_k$.

■

Continuity

- We will now discuss the subject of continuity. Intuitively, a function is called continuous if its graph can be drawn without lifting the pencil off the paper. In a calculus course, you may have defined mathematically as requiring that for all a

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

  To do this, we need to define what it means to write an expression of the form $\lim_{x\to a} f(x)$.

- Often functions are not defined on the whole possible set, but we still want to talk about the continuity. An example is $f(x) = 1/x$. That's why we'll often talk about functions of the form $f : U \subseteq \mathbb{R} \to \mathbb{R}$.

- **Definition:** *Suppose that $f : U \subseteq \mathbb{R} \to \mathbb{R}$ is a function. For any $a \in U$, we say that $y \in \mathbb{R}$ is the* ***limit*** *of f as x goes to a, which we write as write $y = lim_{x\to a} f(x)$, if*

$$\forall\varepsilon > 0 \;\; \exists\delta > 0 \;\; \forall x \in U : x \neq a \text{ and } |x - a| < \delta, \;\; |f(x) - y| < \varepsilon.$$

  We define left and right hand side limits respectively as

$$y = \lim_{x\to a^-} f(x) \quad \text{if} \;\; \forall\varepsilon > 0 \;\; \exists\delta > 0 \;\; \forall x \in U : x \neq a, x < a \text{ and } |x - a| < \delta, \;\; |f(x) - y| < \varepsilon.$$

$$y = \lim_{x\to a^+} f(x) \quad \text{if} \;\; \forall\varepsilon > 0 \;\; \exists\delta > 0 \;\; \forall x \in U : x \neq a, x > a \text{ and } |x - a| < \delta, \;\; |f(x) - y| < \varepsilon.$$

- **Proposition:** *Suppose that $f : U \subseteq \mathbb{R} \to \mathbb{R}$ is a function. If $a \in U$, the limit $lim_{x \to a} f(x)$ exists if and only if its left and right hand side limits exist and are equal: $lim_{x \to a^-} f(x) = lim_{x \to a^+} f(x)$.*

  Proof: TA discussion section, homework, or exercise.

- Most likely up to this point you've worked only with functions of the form $f : \mathbb{R} \to \mathbb{R}$. But it turns out that the theory of continuity is essentially the same for more general functions of the form $f : \mathbb{R}^m \to \mathbb{R}^n$, which in turn have much richer applications. So we'll study the latter. The explicit way to write $f : \mathbb{R}^m \to \mathbb{R}^n$ is as

$$f(x) = \big(f_1(x_1, \dots, x_m), \dots, f_n(x_1, \dots, x_m)\big).$$

  For example, $f : \mathbb{R}^2 \to \mathbb{R}^2$ takes points in $\mathbb{R}^2$ to points in $\mathbb{R}^2$.

- **Definition:** *Suppose that $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is a function. For any $a \in U$ we say that $y \in \mathbb{R}^n$ is the limit of f as x goes to a, which we write as write $y = lim_{x \to a} f(x)$, if*

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 \ \ \forall x \in U : x \neq a \text{ and } |x - a| < \delta, \ \ |f(x) - y| < \varepsilon$$

  (notice that it's exactly the same as before!).

- **Definition:** *Suppose that $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is a function. We say that f is **continuous at $\boldsymbol{a} \in \boldsymbol{U}$** if*

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

  *We say that f is **continuous** if it is continuous at every point where it is defined.*

- **Remark:** If we write out explicitly the condition of $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ being continuous at $a \in U$, one gets

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 \ \ \forall x \in U : |x - a| < \delta, \ \ |f(x) - f(a)| < \varepsilon.$$

  The reason we don't need to write the condition "$x \neq a$" present in the definition of limit is that if $x = a$, then $|f(x) - f(a)| < \varepsilon$ is automatically satisfied since the left-hand side is zero.

  //

- **Remark:** Observe that in the language of open balls, another equivalent way of writing down $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ being continuous at $a \in U$ is

$$\forall \varepsilon > 0 \ \ \exists \delta > 0 \ \ \forall x \in U : x \in B_{\mathbb{R}^n}(a; \delta), \ \ f(x) \in B_{\mathbb{R}^n}(f(a); \varepsilon).$$

  This sometimes proves to be a powerful reformulation, in particular when we study topology later.

  //

- **Proposition:** The following functions are continuous:
  - The function $f(x) = x$

- The function $g(x) = c$ where $c \in \mathbb{R}$ is some constant.

Proof: We'll prove the first one and leave the second to the reader. Take any $a \in \mathbb{R}$. We need to prove that $f(a) = \lim_{x\to a} f(x)$, or in other words $a = \lim_{x\to a} x$. Take any $\varepsilon > 0$. We consider

$$|f(x) - f(a)| = |x - a|$$

Letting $\delta = \varepsilon$, we have that for any $x \in \mathbb{R} : |x - a| < \delta = \varepsilon$, the above is trivially less than ε. So indeed $a = \lim_{x\to a} x$ and hence $f(x) = x$ is continuous.

■

- Now we study how continuity behaves under basic operations.
- **Notation:**
  - If f and g are functions, then $f \circ g$ denotes the composition (i.e. $f \circ g(x) = f\big(g(x)\big)$).
  - If $f : A \to B$ is a map, $f[A] = \{f(x) : x \in A\} \subseteq B$ is the range of f (i.e. everything f hits).
- **Theorem:** *Suppose that $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ and $g : V \subseteq \mathbb{R}^n \to \mathbb{R}^k$ are continuous functions such that $f[U] \subseteq V$. Then $g \circ f : U \subseteq \mathbb{R}^m \to \mathbb{R}^k$ is continuous.*

Proof: Take any $a \in U$ and take $f(a) \in V$. Now, take any $\varepsilon > 0$. We're aiming to show that

$$\exists \delta > 0 \;\; \forall x \in U : x \neq a \text{ and } |x - a| < \delta, \;\; \Big|g\big(f(x)\big) - g\big(f(a)\big)\Big| < \varepsilon$$

Since g is continuous,

$$\exists \hat{\delta} > 0 \;\; \forall y \in V : y \neq f(a) \text{ and } |y - f(a)| < \hat{\delta}, \;\; \Big|g(y) - g\big(f(a)\big)\Big| < \varepsilon. \tag{3}$$

Since f is continuous,

$$\exists \delta' > 0 \;\; \forall x \in U : x \neq a \text{ and } |x - a| < \delta', \qquad |f(x) - f(a)| < \hat{\delta}.$$

By setting $y = f(x)$ in (3), we see that setting $\delta = \delta'$ gives us what we want.

■

- **Proposition:** The functions $h_A, h_M : \mathbb{R}^2 \to \mathbb{R}$ and $h_D : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ given by

$$h_A(x, y) = x + y, \;\; h_M(x, y) = xy, \;\; h_D(x) = 1/x$$

are continuous.

Proof: The proofs are very similar to the proof of the "Operations with Limits" theorem. We'll do the third one, and leave the rest as an exercise. Because h_D is hard to write, let me instead write $g = h_D$. Take any $a \in \mathbb{R} \setminus \{0\}$ (i.e. $a \neq 0$). Take any $\varepsilon > 0$. We look at

$$|g(x) - g(a)| = \left|\frac{1}{x} - \frac{1}{a}\right| = \left|\frac{x - a}{xa}\right| = \frac{|x - a|}{|x||a|}.$$

Let $\delta = \min\{a^2\varepsilon/2\,, |a|/2\}$. Observe for any $x \neq 0 : |x - a| < \delta$,

$$|x| = |x - a - (-a)| \geq \big||x - a| - |-a|\big| = \big||a| - |x - a|\big| \geq |a| - |x - a| > \frac{|a|}{2}.$$

So for any $x \neq 0 : |x - a| < \delta$,

$$\frac{|x - a|}{|x||a|} < \frac{a^2\varepsilon/2}{\frac{|a|}{2}|a|} = \varepsilon.$$

So indeed g is continuous.

$\blacksquare$

- The following lemma and its variants are often useful when plugging multiple functions into one function.
- **Lemma:** *Suppose that $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ and $g : U \subseteq \mathbb{R}^m \to \mathbb{R}^k$, then $(f, g) : U \to \mathbb{R}^{n+k}$ is also continuous. Explicitly this is the function*

$$(f, g)(x) = \big(f_1(x_1, \dots, x_m), \dots, f_n(x_1, \dots, x_m), g_1(x_1, \dots, x_m), \dots, g_k(x_1, \dots, x_m)\big).$$

Proof: Take any $a \in U$. We need to show that $(f, g)(a) = \lim_{x\to a}(f, g)(x)$. Pick any $\varepsilon > 0$.

$$\exists\delta > 0 \;\; \forall x \in U : |x - a| < \delta, |f(x) - f(a)| < \frac{\varepsilon}{\sqrt{2}}$$

$$\exists\hat{\delta} > 0 \;\; \forall x \in U : |x - a| < \hat{\delta}, |g(x) - g(a)| < \frac{\varepsilon}{\sqrt{2}}$$

Set $\delta' = \min\{\delta, \hat{\delta}\} > 0$. Then for $x \in U : |x - a| < \delta'$,

$$|(f, g)(x) - (f, g)(a)|$$

$$= \sqrt{\big(f_1(x) - f_1(a)\big)^2 + \dots + \big(f_n(x) - f_n(a)\big)^2 + \big(g_1(x) - g_1(a)\big)^2 + \dots + \big(g_k(x) - g_k(a)\big)^2}$$

$$< \sqrt{\left(\frac{\varepsilon}{\sqrt{2}}\right)^2 + \left(\frac{\varepsilon}{\sqrt{2}}\right)^2} = \varepsilon.$$

So indeed $(f, g)(a) = \lim_{x\to a}(f, g)(x)$.

$\blacksquare$

Remark: The above theorem naturally generalizes when you need to group more functions together (e.g. $(f, g, \dots, h)$). //

- **Theorem:** *Suppose that $f : U \subseteq \mathbb{R}^m \to \mathbb{R}$ and $g : U \subseteq \mathbb{R}^m \to \mathbb{R}$ are continuous functions. Then $f + g$, $f - g$, and fg are all continuous on U. If S is the set where g is zero (i.e. $S = \{x \in dom\, g : g(x) = 0\}$), then f/g is continuous on $U \setminus S$.*

Proof: We'll prove that $f - g$ is continuous, the rest are an exercise. Take the function $h_A(x, y) = x + y$ and $h_M(x, y) = xy$ from the previous theorem. Consider also the continuous function $h_{-1}(x) = -1$. Then $f - g$ can be expressed as the composition

$$(f-g)(x) = h_A\Big(f, h_M\big(h_{-1}(x), g(x)\big)\Big).$$

By the previous theorem, this is indeed continuous.

■

- **Proposition:** *The **projection function** $P_k : \mathbb{R}^n \to \mathbb{R}$ given by*

  $$P_k(x_1, \dots, x_n) = x_k,$$

  where $1 \le k \le n$, is continuous.

  Proof: Take any $a \in \mathbb{R}^n$. We need to show that $P_k(a) = \lim_{x\to a} P_k(x)$. Take any $\varepsilon > 0$. We need to prove that

  $$\exists \delta > 0 \quad \forall x \in \mathbb{R}^n : x \neq a \text{ and } |x - a| < \delta, \;\; |P_k(x) - P_k(a)| = |x_k - a_k| < \varepsilon.$$

  We claim that $\delta = \varepsilon$ works. Then for any x satisfying the above,

  $$|x_k - a_k| = \sqrt{(x_k - a_k)^2} \le \sqrt{(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2} = |x - a| < \delta = \varepsilon.$$

  ■

- **Corollary:** *If $f, g : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ are continuous functions, then*

  $$f + g : U \subseteq \mathbb{R}^m \to \mathbb{R}^n \quad \textit{and} \quad f \cdot g : U \subseteq \mathbb{R}^m \to \mathbb{R}.$$

  are continuous functions (here "$\cdot$" is the dot product). If $n = 3$, then $f \times g : U \subseteq \mathbb{R}^m \to \mathbb{R}^3$ is continuous.

  Proof: They follow directly from our previous theorems. Let us demonstrate this for $f + g$. Letting h_A and P_k be as in the previous two propositions, using the previous lemma we see that $f + g$ is given by the following composition of continuous functions:

  $$f + g = (f_1 + g_1, \dots, f_n + g_n) = \Big(h_A\big(P_1(f), P_1(g)\big), \dots, h_A\big(P_n(f), P_n(g)\big)\Big).$$

  ■

- There is another very useful notion of continuity. Which we'll prove is equivalent to the notion of continuity that we proved above.

- **Definition:** A function $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is called **sequentially continuous at $\boldsymbol{a} \in \boldsymbol{U}$** if for any sequence $\{x_k \in U\}$ such that $a = \lim_{k\to\infty} x_k$,

  $$f(a) = \lim_{k\to\infty} f(x_k).$$

  A function is called **sequentially continuous** if it is sequentially continuous everywhere.

- **Theorem:** *A function $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is continuous at $a \in U$ (in the usual sense we defined above) if and only if it is sequentially continuous at $a \in U$.*

  Proof: Suppose that f is continuous at $a \in U$. Take any any sequence $\{x_k \in U\}$ such that $a = \lim_{k\to\infty} x_k$. We will show that $f(a) = \lim_{k\to\infty} f(x_k)$. Take any $\varepsilon > 0$. Since f is continuous,

  $$\exists \delta > 0 \ \ \forall x \in U : x \neq a \text{ and } |x - a| < \delta, \ \ |f(x) - f(a)| < \varepsilon.$$

  Since $a = \lim_{k\to\infty} x_k$

  $$\exists K > 0 \ \ \forall k > K, \ \ |x_k - a| < \delta.$$

  Notice then that for $k > K$, $|f(x_k) - f(a)| < \varepsilon$. So indeed $f(a) = \lim_{k\to\infty} f(x_k)$.

  Now suppose that f is sequentially continuous at $a \in U$. We will show that f is continuous by contradiction. Suppose not! Then

  $$\exists \varepsilon > 0 \ \ \forall \delta > 0 \ \ \exists x \in U : x \neq a \text{ and } |x - a| < \delta, |f(x) - f(a)| \geq \varepsilon.$$

  Fix such an $\varepsilon > 0$. Construct a sequence $\{x_k \in U\}$ as follows. For every $k \geq 1$, consider $\delta = 1/k$ and so by the above statement there exists $x_k \in U : x_k \neq a$ and $|x_k - a| < 1/k$ such that $|f(x_k) - f(a)| \geq \varepsilon$. Consider the sequence $\{x_k\}$.

  First let's prove that $a = \lim_{k\to\infty} x_k$. Take any $\hat{\varepsilon} > 0$. Let $K > 0$ be so big so that $1/K < \hat{\varepsilon}$. Then for any $k > K$,

  $$|x_k - a| < \frac{1}{k} < \frac{1}{K} < \hat{\varepsilon}.$$

  So indeed $a = \lim_{k\to\infty} x_k$. Next let's prove that $f(a)$ is not the limit of $\{f(x_k)\}$, which will contradict that f is sequentially continuous. Notice that $|f(x_k) - f(a)| \geq \varepsilon$ always. So there cannot exist $K > 0$ so that for all $k > K$, $|f(x_k) - f(a)| < \varepsilon$. This contradiction proves the theorem.

  ■

- **Corollary:** *A function $f : U \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is continuous everywhere if and only if it is sequentially continuous everywhere.*

  Proof: Trivially follows from previous theorem. ■

Metric Spaces

- It turns out that generalizing the notion of continuity to more general spaces has proved to provide profound implications on other fields of mathematics, such as functional analysis and its applications to ODEs and PDEs. The appropriate generalization is the following:

- **Definition:** *Suppose that X is a set. Suppose we also have a **metric** (or **distance function**) $d_X : X \times X \to \mathbb{R}$, which means that d_X satisfies*

- $d_X(x,y) \geq 0$ *for all* $x, y \in X$, *and* $d_X(x,y) = 0$ *if and only if* $x = y$,
- $d_X(x,y) = d_X(y,x)$ *for all* $x, y \in X$,
- $d_X(x,z) \leq d_X(x,y) + d_X(y,z)$ *for all* $x, y, z \in X$ *(the triangle inequality!).*

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

- It turns out that most of the theory of sequences and continuity over $\mathbb{R}^n$ carry over to metric spaces verbatim upon changing $\|x - y\|_2$ to $d(x,y)$ because most proofs only use the above three properties. We note that the same does not hold for differential and integral calculus, since those typically require a bit more structure on X such as being a vector space or having what's called a measure.

- **Example:** $X \subseteq \mathbb{R}^n$ with $d(x,y) = \|x - y\|$ is a metric space.
  - From TA section: it turns out that setting $d(x,y) = \|x - y\|_1$ or $d(x,y) = \|x - y\|_\infty$ also gives a metric space. //

- **Example:** Suppose we set X to be a curved surface in $\mathbb{R}^3$ given as the graph of an infinitely differentiable function: $z = f(x,y)$. For any two points x, y on this surface, set $d(x,y)$ to be the length of the shortest curve connecting them. It turns out that this is a metric space.
  - You don't have the tools to rigorously prove that d is well-defined or satisfies the above properties. There is a way to generalize this to more general curved surfaces as well. //

- **Example:** Consider the space of bounded functions $X = \{f : \mathbb{R} \to \mathbb{R} \text{ bounded}\}$ and set $d(f,g) = \sup|f(x) - g(y)|$. This is a metric space. //

- **Note:** Until we specify otherwise, (X, d_X) and (Y, d_Y) represent metric space (we'll often need two). If in some setting we're only working with one metric space (X, d_X), we sometimes just write "d" instead of "d_X." //

- **Notation:** *For any point* $x \in X$ *and any radius* $r > 0$, *we denote the* ***(open) ball centered at x with radius r*** *as*

$$B_X(x;r) = \{x' \in X : d_X(x', x) < r\}.$$

- **Remark:** Observe that trivially $B_X(x;r_1) \subseteq B_X(x;r_2)$ if $r_1 \leq r_2$.

- The "X" in $B_X(x;r)$ is to remind us that this is a ball in the metric space X in case there are other metric spaces involved (e.g. Y). If in some setting we're only working with one metric space (X, d_X), we sometimes just write "B" instead of "B_X."

- **Definition:** *Suppose that* $\{x_k\}$ *is a sequence of points in* X. *We say that* $x \in X$ *is the* ***limit*** *of* $\{x_k\}$ *if*

$$\forall \varepsilon > 0 \ \ \forall K > 0 \ \ \forall k > K, \ \ d_X(x_k, x) < \varepsilon.$$

- **Definition:** *Suppose that* $f : X \to Y$ *is a function between metric spaces. We say that* $y = lim_{x\to a} f(x)$ *if*

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall x \in X : x \neq a \text{ and } d_X(x,a) < \delta, \;\; d_Y\big(f(x), f(a)\big) < \varepsilon$$

(note that $y \in Y$*).*

- **Definition:** *We say a function* $f : X \to Y$ *is* ***continuous at*** $\boldsymbol{a} \in \boldsymbol{X}$ *if*

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

We say that f *is* ***continuous*** *if it is continuous everywhere.*

- **Remark:** Notice that the condition of $f : X \to Y$ being continuous at $a \in X$ can be equivalently written in the following two ways:

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall x \in X : d_X(x,a) < \delta, \;\; d_Y\big(f(x), f(a)\big) < \varepsilon.$$

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall x \in X : x \in B(a;\delta), \;\; f(x) \in B_\varepsilon(f(a);\varepsilon).$$

The reason we don't need to write the condition "$x \neq a$" in the first line is that if $x = a$, then $d_Y\big(f(x), f(a)\big) < \varepsilon$ is automatically satisfied since $f(x) = f(a)$ and so the left-hand side is zero.

//

- **Definition:** *A function* $f : X \to Y$ *is called* ***sequentially continuous*** *at* $a \in X$ *if for any sequence* $\{x_k\}$ *in* X *such that* $a = lim_{k\to\infty} x_k$*,*

$$f(a) = \lim_{k\to\infty} f(x_k).$$

We say that f *is* ***sequentially continuous*** *if it is sequentially continuous everywhere.*

- **Theorem:**
  - *Suppose that* $f : X \to Y$ *and* $g : Y \to Z$ *are continuous. Then* $g \circ f : X \to Z$ *is continuous.*
  - $f : X \to Y$ *is continuous at* $a \in X$ *if and only if it is sequentially continuous at* a*.*
  - $f : X \to Y$ *is continuous if and only if it is sequentially continuous.*
  - *If* $f : X \to \mathbb{R}^n$ *and* $g : X \to \mathbb{R}^k$ *are continuous, then* $(f, g) : X \to \mathbb{R}^{n+k}$ *is continuous.*
  - *If* $f : X \to \mathbb{R}$ *and* $g : X \to \mathbb{R}$ *are continuous, then* $f + g$*,* $f - g$*,* fg*, are continuous, and* f/g *is continuous on* $X \setminus S$ *where* $S = \{x \in X : g(x) = 0\}$*.*
  - *If* $f : X \to \mathbb{R}^n$ *and* $g : X \to \mathbb{R}^n$ *are continuous, then* $f + g$*,* $f \cdot g$ *and* $f \times g$ *are continuous (latter only makes sense in case of* $n = 3$*).*

Proof: Proofs are the same as the previous analogous theorems, except replace "$\|\cdot\|$" with "d_X" or "d_Y."

∎

- **Metric Topology**

- Now we begin studying a type of geometry one could call "topological geometry" (not an actual name). The difference with other types of geometry, such as high school geometry, is that this type of geometry cares only about properties of shapes that are unchanged by continuous deformations. This is a beautiful field of its own and surprisingly has essential applications to analysis.

- **Definition:** For any set S, the **complement of S** is everything *not* contained in S (as a technical point: we think of S as sitting in another bigger set). We denote the complement of S by S^c (the little "c" stands for "complement").

- **Definition:** *Consider a subset $U \subseteq X$ of a metric space.*
  - *A point $x \in X$ is called an **interior point** of U if*
  $$\exists r > 0, \;\; B_X(x; r) \subseteq U.$$
  Note that this automatically implies that $x \in U$.
  - *A point $x \in X$ is called a **boundary point** of U if*
  $$\forall r > 0, \;\; B_X(x; r) \cap U \neq \emptyset \;\; \textit{and} \;\; B_X(x; r) \cap U^c \neq \emptyset.$$
  - *A point $x \in X$ is called an **exterior point** of U if*
  $$\exists r > 0, \;\; B_X(x; r) \subseteq U^c.$$

- **Remark:** Simply using logic, one can reason that the above definition classifies all points of $x \in X$ (i.e. all points $x \in X$ are either interior, boundary, or exterior points of U - they cannot be two or three of the types, just one).

  Furthermore, it immediately follows form the definition that an interior point of U is an exterior point of U^c and vice versa: an interior point of U^c is an exterior point of U.

- **Definition:** *For any subset $U \subseteq X$, the set of boundary points is called the **boundary of U** and is denote by ∂U.*

- **Remark:** Since the definition of boundary point is symmetric with respect to U and U^c, a point $x \in X$ is a boundary point of U if and only if it is a boundary point of U^c. So we get the equation:
$$\partial U = \partial(U^c).$$
//

- **Example:** In $\mathbb{R}^n$ (i.e. the metric space $X = \mathbb{R}^n$, $d_X(x, y) = \|x - y\|$), consider the (open) ball

$$B_{\mathbb{R}^n}(x;r) = \{y \in \mathbb{R}^n : \|x - y\| < r\}$$

- Every point $y \in B_{\mathbb{R}^n}(x;r)$ is an interior point of $B_{\mathbb{R}^n}(x;r)$.
- Every point $y \in \mathbb{R}^n$ such that $\|x - y\| = r$ is a boundary point of $B_{\mathbb{R}^n}(x;r)$. More precisely

$$\partial B_{\mathbb{R}^n}(x;r) = \{y \in \mathbb{R}^n : \|x - y\| = r\}.$$

- Every point $y \in \mathbb{R}^n$ such that $\|x - y\| > r$ is an exterior point of $B_{\mathbb{R}^n}(x;r)$.

Let's prove the first one, the rest are exercises. Take any $x' \in B_{\mathbb{R}^n}(x;r)$. We need to show that exists $r' > 0$ such that $B_{\mathbb{R}^n}(x';r') \subseteq B_{\mathbb{R}^n}(x;r)$. Take any $r' > 0 : r' < r - \|x - x'\|$. Then for any point $y' \in B_{\mathbb{R}^n}(x';r')$

$$\|x - y'\|_2 \leq \|x - x'\| + \|x' - y'\| < \|x - x'\| + r' < \|x - x'\| + r - \|x - x'\| = r.$$

So indeed $B_{\mathbb{R}^n}(x';r') \subseteq B_{\mathbb{R}^n}(x;r)$. //

- Repeating the reasoning verbatim involved in the above example but replacing "$\| \cdot \|_2$" with "d_X" gives:
- **Lemma:** *Consider an (open) ball $B_X(x;r)$ in a metric space X.*
  - *Every point $x' \in B_X(x;r)$ is an interior point of $B_X(x;r)$.*
  - *Every point $x' \in X : d_X(x',x) > r$ is an exterior point of $B_X(x;r)$.*
- The reason points $x' \in X : d_X(x',x) = r$ are not necessarily boundary points can be seen by the example of the subset $\mathbb{Z} \subseteq \mathbb{R}^1$.
- **Definition:** *A sequence $\{x_n\}$ in X is called **Cauchy** if for any (the following are two equivalent statements)*

$$\forall \varepsilon > 0 \;\; \exists N > 0 \;\; \forall n > N, \;\; d_X(x_n, x_N) < \varepsilon.$$

$$\forall \varepsilon > 0 \;\; \exists N > 0 \;\; \forall m,n > N, \;\; d_X(x_m, x_n) < \varepsilon.$$

- **Definition:** *A metric space (X,d) is called **complete** if every Cauchy sequence $\{x_k\}$ in X has a limit: $x = lim_{k\to\infty} x_k$ where $x \in X$.*
- **Theorem:** $\mathbb{R}^n$ *with the usual metric is complete.*
- **Proof:** Homework assignment. ■
- **Definition:**
  - *A subset $U \subseteq X$ is called **open** if all of its points are interior points.*
  - *A subset $A \subseteq X$ is called **closed** if it contains all of its boundary points (note that a set automatically contains its interior points and never contains its exterior points).*

- **Lemma:** *A (open) ball $B_X(x, r)$ in a metric space X is open (that's why we call it an "open" ball).*

  Proof: By the previous lemma, every point in $B_X(x; r)$ is an interior point. ■

- It's sometime awkward to rigorously prove that sets other than the balls are open, such as the set of all points strictly above the parabola $y = x^2$. We'll develop tools later that make this easier.

- Riddle: What's the difference between a set and a door. *Answer:* A door is either open or closed, while a set can be either or both.

- **Theorem:**
  - *A subset $U \subseteq X$ is open if and only if U^c is closed.*
  - *A subset $A \subseteq X$ is closed if and only if A^c is open.*

  Proof: To prove the first point, take any $x \in \partial(U^c)$ (i.e. a boundary point of U^c). We have to show that $x \in U^c$. Earlier we observed that $\partial U = \partial(U^c)$. So $x \in \partial U$. Since U contains only interior points, $x \notin U$ or equivalently $x \in U^c$. So indeed U^c is closed.

  Now suppose that U^c is closed. Take any $x \in U$. Since U^c contains all of its boundary points and interior points, we have that x is an exterior point of U^c, which means that it is an interior point of U. Hence indeed U is open.

  The second point of the theorem is logically equivalent to the first point since $A = (A^c)^c$.

  ■

- If you really want to make a set closed, fret not! There is an operation to help you:

- **Definition:** *For any set $A \subseteq X$ (not necessarily open/closed). We define its* ***closure*** *by*

$$\bar{A} = A \cup \{\textit{all boundary points of } A\}.$$

- **Lemma:** *For any set $A \subseteq X$, it's closure $\bar{A}$ is closed (hence "closure" was a good name)*

  Proof: We will prove that $\bar{A}$ is closed by showing that $(\bar{A})^c$ is open. Take any point in $x \in (\bar{A})^c$ (i.e. $x \notin \bar{A}$). We need to show that there exists $r > 0$ such that $B_X(x; r) \subseteq (\bar{A})^c$. We know that x is not an interior point of A since $x \notin A$. Similarly, x is not a boundary point of A since $x \notin \bar{A}$. So x is an exterior point of A. So there exists $r > 0$ such that $B_X(x; r) \subseteq A^c$. Every point $y \in B_X(x, r)$ is also an exterior point of A because the fact that $B_X(x; r)$ is open implies that there exists $r' > 0$ such that

$$B_X(y, r') \subseteq B_X(x, r) \subseteq A^c.$$

  So every point $y \in B_X(x, r)$ is not in $\bar{A}$. Hence $B_X(x, r) \subseteq (\bar{A})^c$.

  ■

- **Example:** By a previous example we get that

$$\overline{B_{\mathbb{R}^n}(x;r)} = \{x' \in \mathbb{R}^n : \|x' - x\| \leq r\}.$$

Note, this is not true in a general metric space!

//

- **Theorem:** *Suppose that $\{U_1, \dots, U_m\}$ is a finite set of open sets in X. Then*

$$\bigcap_{j\in\{1,\dots,m\}} U_j = U_1 \cap \dots \cap U_m \quad \textit{is also open.}$$

Suppose that $\{U_\gamma\}_{\gamma\in\Gamma}$ is a (possibly infinite) set of open sets in X. Then

$$\bigcup_{\gamma\in\Gamma} U_\gamma \quad \textit{is also open.}$$

Proof: Let's begin with the first part. Take any point $x \in U_1 \cap \dots \cap U_m$. Because it's an interior point of every U_j, for any $j = 1, \dots, m$ there exists $r_j > 0$ such that $B_X(x, r_j) \subseteq U$. Let $r = \min\{r_j : j = 1, \dots, m\}$. Then observe that for each $j = 1, \dots, n$,

$$B_X(x,r) \subseteq B_X(x, r_j) \subseteq U_j$$

And so

$$B_X(x,r) \subseteq U_1 \cap \dots \cap U_m.$$

Hence x is an interior point of $U_1 \cap \dots \cap U_m$. So $U_1 \cap \dots \cap U_m$ is open.

Let's prove the second part. Take any $x \in \bigcup_{\gamma\in\Gamma} U_\gamma$. Then there exists $\gamma_0 \in \Gamma$ such that $x \in U_{\gamma_0}$. Since U_{γ_0} is open, there exists $r > 0$ such that

$$B_X(x,r) \subseteq U_{\gamma_0} \subseteq \bigcup_{\gamma\in\Gamma} U_\gamma.$$

So x is an interior point of $\bigcup_{\gamma\in\Gamma} U_\gamma$, and hence $\bigcup_{\gamma\in\Gamma} U_\gamma$ is open.

■

Example and nonexample: An example of the above theorem is:

$$\{U_j\}_{j\in\{1,\dots,10\}} = \left\{B_{\mathbb{R}^n}\left(x, \frac{1}{j}\right)\right\}_{j\in\{1,\dots,10\}} \quad \text{and} \quad \{U_r\}_{r\in\{b\geq 0: b<1\}} = \{B_{\mathbb{R}^n}(x,r)\}_{r\in\{b\geq 0: b<1\}}.$$

Then

$$U_1 \cap \dots \cap U_{10} = B_{\mathbb{R}^n}\left(x, \frac{1}{10}\right), \quad \text{which is open!}$$

$$\bigcup_{r\in\{b\geq 0:b<1\}} U_r = B_{\mathbb{R}^n}\left(x,\frac{1}{10}\right), \quad \text{which is open!}$$

Example of when the theorem doesn't apply is

$$\bigcap_{r\in\{b\geq 0:b<1\}} U_r = \{0\}, \quad \text{which is not open!}$$

//

- In point set topology - a more advanced course - the above theorem is in fact taken to be the definition of open sets since in which case you don't need the presence of a metric. Closed sets are defined to be the compliments of open sets. A rich theory can be constructed out of this, despite not having a metric!

- **Theorem:** *Suppose that $\{A_1, \dots, A_m\}$ is a finite set of closed sets in X. Then*

$$\bigcup_{j\in\{1,\dots,n\}} A_j = A_1 \cup \dots \cup A_j \quad \textit{is also closed.}$$

Suppose that $\{A_\gamma\}_{\gamma\in\Gamma}$ is a (possibly infinite) set of closed sets in X. Then

$$\bigcap_{\gamma\in\Gamma} A_\gamma \quad \textit{is also closed.}$$

Proof: You can do a similar argument as for the previous theorem. Alternatively, observe that

$$\left(A_1 \cup \dots \cup A_j\right)^c = A_1^c \cap \dots \cap A_j^c$$

which is open since each A_j^c open and by the previous theorem. So $A_1 \cup \dots \cup A_j$ is closed. The second part is proven similarly using

$$\left(\bigcap_{\gamma\in\Gamma} A_\gamma\right)^c = \bigcup_{\gamma\in\Gamma} A_\gamma^c.$$

■

- **Example and nonexample:** Examples of the above theorem are:

$$\overline{B_{\mathbb{R}^n}(x,1)} \cup \dots \overline{B_{\mathbb{R}^n}(x,10)} = \overline{B_{\mathbb{R}^n}(x,10)} \quad \text{which is closed!}$$

$$\bigcap_{r\in\{b\geq 0: b<1\}} \overline{B_{\mathbb{R}^n}(x,r)} = \{0\} \quad \text{which is closed!}$$

An example of where the theorem does not apply:

$$\bigcup_{r\in\{b\geq 0: b<1\}} \overline{B_{\mathbb{R}^n}(x,r)} = B_1(x,1), \quad \text{which is not closed!}$$

//

Theorem: *A function $f : X \to Y$ between metric spaces is continuous if and only if for any open set $U \subseteq Y$,*

$$f^{-1}[U] = \{x \in X : f(x) \in U\}$$

is open.

Proof: Suppose that f is continuous. Take any open set $U \subseteq Y$. Take any point $a \in f^{-1}[U]$. We have to show that there exists an $r > 0$ such that $B_X(a;r) \subseteq f^{-1}[U]$. Take $f(a) \in U$. Since U is open, there exists $\varepsilon > 0$ such that $B_Y(f(a);\varepsilon) \subseteq U$. Since f is continuous,

$$\exists \delta > 0 \ \ \forall x \in X : d_X(x,a) < \delta, \ \ d_Y\big(f(x), f(a)\big) < \varepsilon.$$

Equivalently:

$$\exists \delta > 0 \ \ \forall x \in B_X(a;\delta), \ \ f(x) \in B_Y(f(a);\varepsilon) \subseteq U.$$

Fix such a $\delta > 0$. Observe that the above statement says that any point in $B_X(a;\delta)$ get mapped into U by f. Hence $B_X(a;\delta) \subseteq f^{-1}[U]$. Setting $r = \delta$, we see that a is an interior point of $f^{-1}[U]$.

Now suppose that for any open set $U \subseteq Y$, $f^{-1}[U]$ is open. Take any $a \in X$. We will show that f is continuous at a. Take any $\varepsilon > 0$. Consider $f(a) \in U$. Since U is open, there exists an $\varepsilon' > 0 :$ $\varepsilon' < \varepsilon$ such that $B_Y(f(a), \varepsilon') \subseteq U$. Since $B_Y(f(a), \varepsilon')$ is open, $f^{-1}[B_Y(f(a), \varepsilon')]$ is open by assumption. Since trivially $a \in f^{-1}[B_Y(f(a), \varepsilon')]$, we have that (here the "⇒" applies to the part indicated by "$\underbrace{\cdot}$").

$$\exists \delta > 0 : \underbrace{B_X(a,\delta) \subseteq f^{-1}[B_Y(f(a), \varepsilon')]}$$

$$\Longrightarrow \forall x \in X : d_X(x,a) < \delta, \underbrace{f(x) \in B_Y(f(a), \varepsilon')}$$

$$\Longrightarrow d_Y\big(f(x), f(a)\big) < \varepsilon' < \varepsilon.$$

So

$$\forall x \in X : d_X(x,a) < \delta, \ \ d_Y\big(f(x), f(a)\big) < \varepsilon$$

and hence f is indeed continuous.

■

- In point set topology the above theorem is in fact taken to be the definition of continuous functions.
- **Definition:** *A continuous function $f : X \to Y$ that has a continuous inverse $f^{-1} : Y \to X$ is called a* ***homeomorphism*** *(in particular, it is a bijection because the inverse exists).*

- Intuitively speaking, if there is a homeomorphism between two metric spaces X and Y, then their topologies (i.e. phenomenon surrounding open sets, boundary points, etc.) are equivalent. This is analogous to the fact that if there is an isomorphism between vector spaces V and W, then they're linear structures are equivalent (for example: their dimension). The following theorem describes what we said rigorously:

- **Theorem:** *Suppose that $f : X \to Y$ is a homeomorphism. Take a subset $U \subseteq X$ and consider its **image** under f:*

$$V = f[U] = \{f(x) : x \in U\} \subseteq Y.$$

Then

1. *A point $x \in U$ is an interior point of U if and only if $f(x)$ is an interior point of V.*
2. *A point $x \in U$ is a boundary point of U if and only if $f(x)$ is a boundary point of V.*
3. *A point $x \in U$ is an exterior point of U if and only if $f(x)$ is an exterior point of V.*

Proof: Quiz section or homework assignment. ■

- If you're in $\mathbb{R}^n$, intuitively speaking two shapes are homeomorphic if one of the shapes can be bent and stretched continuously (i.e. without creating rips) to transform into the other. You can't use this for rigorous arguments, but it is the idea behind why we chose homeomorphisms to describe such phenomenon.

Connected Sets

- **Definition:** Suppose that $E \subseteq X$ is a subset of a metric space X. We say that E is **disconnected** if there exist nonempty subsets $A, B \subseteq E$ such that $E = A \cup B$ and

$$\bar{A} \cap B = \emptyset \quad \text{and} \quad A \cap \bar{B} = \emptyset.$$

In this case we call A and B a **disconnection** of S. If no such disconnection exists, then we say that E is **connected**.

- **Examples:** Examples of disconnected sets include (draw these out!)

$$E = B_{\mathbb{R}^2}\left(0; \frac{1}{2}\right) \cup B_{\mathbb{R}^2}\left(2; \frac{1}{2}\right)$$

and less trivially:

$$E = B_{\mathbb{R}^2}(-1; 1) \cup B_{\mathbb{R}^2}(1; 1) \quad \text{or} \quad E = \text{Graph of } 1/x^2 \text{ (set value at } x = 0 \text{ to be anything)}$$

An interval is an example of a connected set, as we will prove further below.

- **Notation:** *If $a < b$ are real numbers, we let*
  - *(a, b) denote the interval from a to b not including the endpoints a and b,*
  - *$[a, b)$ denote the interval from a to b including a but not b,*

- $(a, b]$ *denote the interval from* a *to* b *not including* a *but including* b,
- $[a, b]$ *denote the interval from* a *to* b *including both* a *and* b.

Similarly:

- $(-\infty, b)$ *denote the interval of all numbers* $< b$
- (a, ∞) *denote the interval of all numbers* $> a$
- $(-\infty, b]$ *denote the interval of all numbers* $\leq b$
- $[a, \infty)$ *denote the interval of all numbers* $\geq a$

- In other words, we write "(" or ")" if do <u>not</u> want to include the endpoint, and "[" or "]" if we <u>do</u> want to include the endpoint.
- **Note:** Recall that in the homework you will show that a point $x \in X$ is in the closure of a set A if you can find a sequence $\{x_k\}$ in X with each $x_k \in A$ such that $x = \lim_{k\to\infty} x_k$. This gives a good geometric way of thinking about closure.
- **Theorem:** The only connected subsets of $\mathbb{R}$ are the intervals.

  Proof: Take a set $S \subseteq \mathbb{R}$. We show that it is connected if and only if it is an interval. First, suppose that S is not an interval. We will show that S is disconnected. Then there exist points $a, b \in S$ and $c \notin S$ such that $a < c < b$. Define

$$A = S \cap (-\infty, c) \quad \text{and} \quad B = S \cap (c, \infty).$$

  Notice that $S = A \cup B$ and that $A \subseteq (-\infty, c)$ and $B \subseteq (c, \infty)$. Next, the closure of A is contained in $(-\infty, c]$ because for any point $x > c$ you can't find a sequence $\{x_k\}$ in A such that $x = \lim_{k\to\infty} x_k$ since each $x_k \leq c$, and so any point $x > c$ is not in the closure of A. For similar reasons, the closure of B is contained in $[c, \infty)$. Hence both

$$\bar{A} \cap B \subseteq (-\infty, c] \cap (c, \infty) = \emptyset \quad \text{and} \quad A \cap \bar{B} \subseteq (-\infty, c) \cap [c, \infty) = \emptyset$$

  Thus A and B are a disconnection of S, and so S is indeed disconnected.

  Now suppose that S is an interval. We will show that it is connected by contradiction. Suppose that it isn't connected. Then there exists a disconnection A and B of S. Let $a \in A$ and $b \in B$, and relabeling A and B if necessary, assume $a < b$. Observe that the whole interval $[a, b]$ must be contained in S because both $a, b \in S$ and S is an interval and hence contains everything in between a and b. Consider the number

$$c = \sup\{x \in A : a \leq x \leq b\}.$$

  Since $a \leq c \leq b$, we have that $c \in S$ and hence c either belong to A or B. Now we ask: does $c \in A$ or $c \in B$? We try both cases:

  1. Suppose c belongs to A. We cannot have $c = b$ or else $A \cap B \neq \emptyset$ and thus $\bar{A} \cap B \neq \emptyset$ (and $A \cap \bar{B} \neq \emptyset$). So we must have that $c < b$. So we can choose a sequence $\{x_k\}$ such

that each $x_k \in (c, b)$ (and hence $x_k \in S$) where x_k decreases to c from above with $c = \lim_{k\to\infty} x_k$. Each $x_k \in B$ because otherwise $x_k \in A$ along with $x_k > c$ would contradict the above supremum property of c. Thus c is in the closure of B. But then we get that $A \cap \bar{B} \neq \emptyset$, contradiction!

2. Suppose c is in B. Because of c's supremum property, we can choose a sequence $x_k \in A$ increasing from below such that $c = \lim_{k\to\infty} x_k$. Thus c is in the closure of A. But then we get that $\bar{A} \cap B \neq \emptyset$, contradiction!

So c can't belong to A nor B, while we observed that it has to belong to at least one of them. So S must indeed be connected.

∎

- The next theorem says that continuous functions take connected sets to connected sets.
- **Theorem:** Suppose that $f : X \to Y$ is a continuous function and that $U \subseteq X$ is a connected set. Then the image of U under f:

$$f[U] = \{f(x) : x \in U\}$$

is also connected.

Proof: We prove this by contradiction: suppose $f[U]$ is not connected. Then there exists a disconnection A and B of $f[U]$:

$$f[U] = A \cup B, \quad \bar{A} \cap B = \emptyset, \quad A \cap \bar{B} = \emptyset.$$

Consider the preimages

$$f^{-1}[A] = \{x \in X : f(x) \in A\} \quad \text{and} \quad f^{-1}[B] = \{x \in X : f(x) \in B\}.$$

Notice that $U \subseteq f^{-1}[A] \cup f^{-1}[B]$ because everything in U maps into $f[U]$ and hence must map into either A or B since $f[U] = A \cup B$. We can then write

$$U = (U \cap f^{-1}[A]) \cup (U \cap f^{-1}[B])$$

Since U is connected, $U \cap f^{-1}[A]$ and $U \cap f^{-1}[B]$ cannot be a disconnection of it. So either

$$\overline{(U \cap f^{-1}[A])} \cap (U \cap f^{-1}[B]) \neq \emptyset \quad \text{or} \quad (U \cap f^{-1}[A]) \cap \overline{(U \cap f^{-1}[B])} \neq \emptyset.$$

By relabeling A and B if necessary, suppose that the first one holds. Then there exists

$$\text{point } x \in U \cap f^{-1}[B] \text{ and a}$$

$$\text{sequence } \{x_k\}, \text{ each } x_k \in U \cap f^{-1}[A], \text{ and } x = \lim_{k\to\infty} x_k.$$

Then $f(x) \in B$ and $\{f(x_k)\}$ is a sequence of points in A. Since f is continuous, $f(x) = \lim_{k\to\infty} f(x_k)$. Hence $f(x) \in \bar{A}$. But then $\bar{A} \cap B \neq \emptyset$, contradiction!

∎

- **Theorem (Intermediate Value Theorem):** Suppose that $f : X \to \mathbb{R}$ is a continuous function where $E \subseteq X$ is a connected set. Suppose that $a, b \in E$ are such that $f(a) < f(b)$. Then for every value y satisfying $f(a) < y < f(b)$, there exists an $x \in E$ such that $f(x) = y$.

  Proof: Take any y satisfying $f(a) < y < f(b)$. Since E is connected, by the previous theorem the image $f[E]$ is also connected. Since $f[E]$ is a subset of $\mathbb{R}$, it must be an interval. Since $f(a), f(b) \in f[E]$, $f[E]$ must contain every value in between $f(a)$ and $f(b)$, in particular it contains y. Hence there exists an $x \in E$ such that $f(x) = y$.

  ∎

- Next we discuss another notion of connectedness.

- **Definition:** A **continuous curve** (or **path**) is a continuous function of the form $\gamma : [a, b] \to X$. Such a curve is also sometimes, but not very often, referred to as an **arc**.

- **Definition:** A set $E \subseteq X$ is called **arcwise connected** (or **path connected**) if for any points $x, y \in E$, there exists a continuous curve $\gamma : [a, b] \to X$ such that $\text{ran}(\gamma) \subseteq E$ and

$$\gamma(a) = x \quad \text{and} \quad \gamma(b) = y.$$

- Intuitively a set being "arcwise connected" means that every two points can be connected by a continuous curve that stays inside that set. As we'll prove below, all arcwise connected sets are connected. However the other direction is not true and hence arcwise connected is a slightly weaker notion of connectedness (look up the "topological sine curve"). Nevertheless it does provide a powerful way to prove that something is connected because in many cases it's easy to demonstrate that there is a continuous curve connecting any two points inside the set.

- **Theorem:** If a set $E \subseteq X$ is arcwise connected then it is connected.

  Proof: Suppose E is arcwise connected. We prove that it is connected by contradiction. Suppose not! Then there exists a disconnection A and B of E:

$$E = A \cup B, \quad \bar{A} \cap B = \emptyset, \quad A \cap \bar{B} = \emptyset.$$

  Take $\alpha \in A$ and $\beta \in B$. Because E is arcwise connected, there exists a continuous curve $\gamma : [a, b] \to M$ such that $\gamma(a) = \alpha$ and $\gamma(b) = \beta$. Consider the sets $\gamma^{-1}[A]$ and $\gamma^{-1}[B]$. Because every time $t \in [a, b]$ gets mapped into $E = A \cup B$ and hence into either A or B, we have that $[a, b] = \gamma^{-1}[A] \cup \gamma^{-1}[B]$. Since $[a, b]$ is connected, $\gamma^{-1}[A]$ and $\gamma^{-1}[B]$ cannot be a disconnection of $[a, b]$ and so either

$$\overline{\gamma^{-1}[A]} \cap \gamma^{-1}[B] \neq \emptyset \quad \text{or} \quad \gamma^{-1}[A] \cap \overline{\gamma^{-1}[B]} \neq \emptyset.$$

  By relabeling A and B if necessary, we can assume that the first one holds. Then there exists a $x \in \gamma^{-1}[B]$ such that $x \in \overline{\gamma^{-1}[A]}$ as well. The latter implies that there exists a sequence $\{x_k \in \gamma^{-1}[A]\}$ so that $x = \lim_{k\to\infty} x_k$. Observe then that $\gamma(x) \in B$ and each $\gamma(x_k) \in A$, and that γ's continuity implies that

$$\gamma(x) = \lim_{k\to\infty} \gamma(x_k).$$

So $\gamma(x) \in \bar{A}$ as well. So $\bar{A} \cap B \neq \emptyset$. But that contradicts that A and B are a disconnection. So E is indeed connected.

■

- Although connectedness does not imply path connectedness, there is a theorem in $\mathbb{R}^n$ that sort of gives something along these lines. We won't make use of the following theorem, so we won't cover the proof. We include a reference for the interested reader.

- **Theorem:** If a subset $U \subseteq \mathbb{R}^n$ is both open and connected, then it is path connected.

  Proof: Theorem 1.30 in *Advanced Calculus 2nd Ed* by Gerald Folland:

  https://sites.math.washington.edu/~folland/Homepage/AdvCalc24.pdf

Compactness

- Compactness, intuitively speaking, is an attribute of sets that plays at least two important roles in analysis: controlling the behavior of continuous maps and giving the ability to identify a point around which a subset seems to accumulate/cluster at. We start with the latter idea to define a type of compactness:

- **Definition:** A subset $K \subseteq X$ is called **sequentially compact** if for any sequence $\{x_k \in K\}$ in K, there exists a subsequence $\{\tilde{x}_n\}$ of $\{x_k\}$ such that has $\{\tilde{x}_n\}$ has a limit $x \in K$: $x = \lim_{n\to\infty} \tilde{x}_n$.

  Remark: $\{\tilde{x}_n\}$ being a subsequence of $\{x_k\}$ means that $\{x_n\}$ is an infinite subset of $\{x_k\}$ (i.e. $\{\tilde{x}_n\} \subseteq \{x_k\}$) that preserves the order of the elements.

- The main definition of compactness is the following:

- **Definition:** A subset $K \subseteq X$ is called **compact** if the following holds. Suppose that $\{U_\gamma\}_{\gamma\in\Gamma}$ is an open cover of K, which means that each U_γ is an open set and they cover K:

$$K \subseteq \bigcup_{\gamma\in\Gamma} U_\gamma.$$

  Then there is a finite subcover $\{U_k\}_{k=1}^m \subseteq \{U_\gamma\}_{\gamma\in\Gamma}$ of K:

$$K \subseteq \bigcup_{k\in\{1,\dots,m\}} U_k = U_1 \cup \dots \cup U_m$$

- In general topologies (which we don't study), compactness and sequential compactness are not necessarily the same thing. However, metric spaces are special in that the two notions are in fact equivalent. As we will prove, in $\mathbb{R}^n$ it turns out that being compact is equivalent to being both closed and bounded. To prove the first theorem, we need two important lemmas

- **Definition:** Suppose that $\{x_k\}$ is a sequence in X. We say that $x \in X$ is an **accumulation point** (or **cluster point**) if for any $r > 0$ there exists infinitely many x_k's in $B_X(x, r)$.

- **Lemma:** Suppose that $\{x_k \in K\}$ is a sequence in $K \subseteq X$. Then it has a subsequence with a limit $x \in K$ if and only if it has an accumulation point $\tilde{x} \in K$.

  Proof: Homework.

- **Lemma:** Suppose that $K \subseteq X$ is compact and that $A \subseteq K$ is closed. Then A is also compact.

  Proof: Suppose that $\{U_\gamma\}_{\gamma\in\Gamma}$ is an open cover of A. We have to show that there exists a finite subcover $\{U_k\}_{k=1}^m \subseteq \{U_\gamma\}_{\gamma\in\Gamma}$ of A. Add on the open set A^c to this collection of open sets to get $\{U_\gamma\}_{\gamma\in\Gamma} \cup \{A^c\}$ and notice that this is now an open cover of K since

$$K \subseteq A \cup A^c \subseteq \left(\bigcup_{\gamma\in\Gamma} U_\gamma\right) \cup A^c.$$

  Since K is compact, we can choose a finite subcover $\{U_k\}_{k=1}^m \cup \{A^c\} \subseteq \{U_\gamma\}_{\gamma\in\Gamma} \cup \{A^c\}$ of K where we "threw in" $\{A^c\}$ into this finite subcover just in case. Since $A \subseteq K$, this finite subcover of K also covers A:

$$A \subseteq K \subseteq \left(\bigcup_{k\in\{1,\ldots,m\}} U_k\right) \cup A^c = U_1 \cup \ldots \cup U_m \cup A^c.$$

  Since A does not intersect A^c, you can throw $\{A^c\}$ out of this finite subcover of A to get that simply $\{U_k\}_{k=1}^m$ is a finite subcover of A. Hence A is indeed compact.

  ∎

- **Theorem:** A subset $K \subseteq X$ is compact if and only if it is sequentially compact (this is only true because X is a metric space).

  Proof: First suppose that K is compact. Take any sequence $\{x_k \in K\}$ in K. We will show that it has a subsequence $\{\tilde{x}_n\}$ that has a limit $x \in K$. Suppose not! Then by the second to last lemma, $\{x_k\}$ has no accumulation point $x \in K$, which means that

$$(\exists x \in K \quad \forall r > 0 \quad B_X(x, r_x) \text{ only has finitely many } x_k\text{'s}) \quad \text{is not true}$$
$$\Leftrightarrow \quad \forall x \in K \quad \exists r_x > 0 \quad B_X(x, r_x) \text{ only has finitely many } x_k\text{'s}$$

  Consider the open cover

$$\{B_X(x, r_x)\}_{x\in K} \quad \text{of} \quad K.$$

(i.e. $K \subseteq \bigcup_{x \in K} B_X(x, r_x)$, which is true because every $x \in K$ is contained in $B_X(x, r_x)$). Because K is compact, we can choose a finite subcover $\left\{B\left(x_j, r_{x_j}\right)\right\}_{j \in \{1,\dots,m\}} \subseteq \{B(x, r_x)\}_{x \in K}$ of K. But then since each $B_X\left(x_j, r_{x_j}\right)$ has finitely many x_k's and the $B_X\left(x_j, r_{x_j}\right)$'s cover K, this implies that there are a finite number of x_k's: contradiction!

Now suppose that K is sequentially compact. Take any open cover $\{U_\gamma\}_{\gamma \in \Gamma}$ of K. We prove the following claim which is called the *Lebesgue number lemma* in the theory of compact subsets of metric spaces.

Claim:

$$\exists \delta > 0 \quad \forall x \in K \quad \exists \gamma \in \Gamma, \;\; B_X(x; \delta) \subseteq U_\gamma$$

(the $\delta > 0$ is called a **Lebesgue number** of the open cover $\{U_\gamma\}_{\gamma \in \Gamma}$).

Proof of claim: Suppose not! Then

$$\forall \delta > 0 \quad \exists x \in K \quad \forall \gamma \in \Gamma, \;\; B_X(x; \delta\,) \nsubseteq U_\gamma$$

Thus for any integer $k > 0$ there exists an x_k such that $B_X(x_k, 1/k)$ is not contain in any U_γ. Since K is sequentially compact, there is a subsequence $\{\tilde{x}_n\}$ of $\{x_k\}$ with limit $x \in K$: $x = \lim_{n \to \infty} \tilde{x}_n$. Let $1/k_n$ denote the radius of the ball $B_X(\tilde{x}_n, 1/k_n)$ associated to $\tilde{x}_n$. Since $\{U_\gamma\}_{\gamma \in \Gamma}$ covers K, x is contained in some U_{γ_0}. Since U_{γ_0} is open, there exists an $r > 0$ such that $B_X(x, r) \subseteq U_{\gamma_0}$. Now choose N big enough so that

$$\forall n > N \quad \tilde{x}_n \text{ is within } r/2 \text{ of } x$$

$$\forall n > N \quad 1/k_n < r/2.$$

By an argument we've done many times (which we leave to the reader), it follows from the triangle inequality that for any $n > N$,

$$B_X(\tilde{x}_n, 1/k_n) \subseteq B_X(x, r) \subseteq U_{\gamma_0}.$$

But this contradicts that by construction we said that none of the $B_X(\tilde{x}_n, 1/k_n)$ are contain in any of the U_γ's.

End of proof of claim

We want to show that there is a finite subcover $\{U_k\}_{k=1}^m \subseteq \{U_\gamma\}_{\gamma \in \Gamma}$ of K. We prove this by contradiction: suppose not! Let $\delta > 0$ be as in the above claim (i.e. a Lebesgue number of $\{U_\gamma\}_{\gamma \in \Gamma}$). Take any $x_1 \in K$. By the above claim, there exists $\gamma_1 \in \Gamma$ such that

$$B_X(x_1; \delta) \subseteq U_{\gamma_1}.$$

Since we assumed that there is no finite subcover of K, U_{γ_1} cannot cover K and so there exists $x_2 \in K$ outside of U_{γ_1} (and hence outside of $B_X(x_1;\delta)$). By the above claim we again have that there exists $\gamma_2 \in \Gamma$ so that $B_X(x_2;\delta) \subseteq U_{\gamma_2}$ and so

$$B_X(x_1;\delta) \cup B_X(x_2;\delta) \subseteq U_{\gamma_1} \cup U_{\gamma_2}$$

By reasoning as before, there exists an $x_3 \in K$ outside of $U_{\gamma_1} \cup U_{\gamma_2}$ (and hence outside of $B_X(x_1;\delta) \cup B_X(x_2;\delta)$). Proceed inductively to get a sequence $\{x_k \in K\}$ in K which satisfies

$$B_X(x_1;\delta) \cup \dots \cup B_X\left(x_j;\delta\right) \subseteq U_{\gamma_1} \cup \dots \cup U_j$$

and x_{j+1} is outside of $U_{\gamma_1} \cup \dots \cup U_{\gamma_j}$ (and hence outside of $B_X(x_1;\delta) \cup \dots \cup B_X\left(x_j,\delta\right)$). Notice that by construction each x_k is at least a distance of δ from all of the other x_k's.

Now, since we assumed that K is sequentially compact. The sequence $\{x_k\}$ has a subsequence $\{\tilde{x}_n\}$ that has a limit $x \in K : x = \lim_{n\to\infty} \tilde{x}_n$. But by construction, each $\tilde{x}_n$ is at least a distance of δ from each other, and hence can't by Cauchy, and hence can't have a limit. Contradiction!

■

- We now prove the theorem that justifies the intuition that compact sets are used to control continuous functions. It says that continuous functions take compact sets to compact sets.

- **Theorem:** If $f : X \to Y$ is a continuous function and $K \subseteq X$ is compact, then the image of K under f:

$$f[K] = \{f(x) : x \in K\}$$

is also compact.

Proof: We will prove that $f[K]$ is sequentially compact. Take any sequence $\{y_k \in f[K]\}$ in $f[K]$. We must show that it has a subsequence $\{\tilde{y}_n \in f[K]\}$ with a limit $y \in f[K]$. For each y_k there exists $x_k \in K$ such that $y_k = f(x_k)$. Consider the sequence $\{x_k \in K\}$ in K. Since K is sequentially compact, there is a subsequence $\{\tilde{x}_n \in K\}$ in K with limit $x \in K : x = \lim_{n\to\infty} \tilde{x}_n$. Since f is continuous $f(x) = \lim_{n\to\infty} f(\tilde{x}_n)$. So set $\{\tilde{y}_n = f(\tilde{x}_n)\}$ and $y = f(x)$, which is a subsequence of $\{y_k\}$ and limit in K that we were looking for.

■

- Next we classify compact subsets of $\mathbb{R}^n$. We need a few preliminary lemmas:

- **Lemma:** A set $A \subseteq X$ is closed if and only if $A = \bar{A}$.

  Proof: Was proved in discussion section, or follows from Problem 3 on HW 8. ■

- **Lemma:** If $\{x_k\}$ has a limit x, then any subsequence will have the same limit.

- **Proof:** Homework. ■

- The following is the famous **Heine-Borel Theorem** that classifies compact subsets of $\mathbb{R}^n$.

- **Theorem:** A nonempty subset $K \subseteq \mathbb{R}^n$ is compact if and only if it is bounded and closed. Rigorously, "bounded" means that K is contained in some big ball centered at zero: $\exists R > 0$, $K \subseteq B_{\mathbb{R}^n}(0; R)$.

  Proof: Suppose K is compact. First we show that it is bounded. Suppose not! Then

  $$\forall R > 0 \ \ \exists x \in K \ \ |x| \geq R.$$

  Construct the sequence $\{x_k \in K\}$ in K as follows. Pick any $x_1 \in K$. By the above, there is $x_2 \in K$ such that $|x_2| \geq |x_1| + 1$.[3] Then there exists $x_3 \in K$ such that $|x_3| > |x_2| + 1$. Proceed inductively to get $\{x_k\}$ which satisfies $|x_{k+1}| > |x_k| + 1$. Note that if $k < j$

  $$|x_j - x_k| \geq |x_j| - |x_k| \geq |x_{k+1}| - |x_k| \geq 1.$$

  So each x_k is at least a distance of 1 away from the other x_k's. Since K is sequentially complete, there exists a subsequence $\{\tilde{x}_n\}$ with limit $x \in K$: $x = \lim_{n\to\infty} \tilde{x}_n$. But that is impossible since the above shows that each $\tilde{x}_n$ is at least a distance of 1 away from the other $\tilde{x}_n$'s, and hence isn't Cauchy, and hence can't have a limit. Hence K must be bounded.

  Next we show that K is closed. We will prove that $K = \overline{K}$. The inclusion $K \subseteq \overline{K}$ is immediate, so we'll prove $K \supseteq \overline{K}$. Take any $x \in \partial K$. We need to show that $x \in K$. Since x is in the closure $\overline{K}$, there exists a sequence $\{x_k \in K\}$ in K so that $x = \lim_{k\to\infty} x_k$. Since K is sequentially compact, there exists a subsequence $\{\tilde{x}_n\}$ which has a limit $\tilde{x} \in K$: $\tilde{x} = \lim_{n\to\infty} \tilde{x}_n$. But by the previous lemma, $x = \tilde{x}$ and so $x \in K$.

  Now suppose that K is closed and bounded. We will prove that it is sequentially compact (and hence compact). Take any sequence $\{z_k\}$ in K, we will show that there is a subsequence $\{\tilde{z}_n\}$ that has a limit $z \in K$. We will do this in $\mathbb{R}^2$ for simplicity, from which it should be clear how to do this in $\mathbb{R}^n$. We use the following notation for a box:

  $$[a, b] \times [c, d] = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, c \leq y \leq d\}$$

  Because K is bounded, we can take a large enough box B_1

  $$B_1 = [-R, R] \times [-R, R] = \{(x, y) \in \mathbb{R}^2 : -R \leq x \leq R, -R \leq y \leq R\}$$

  so that it will contain K (i.e. make R large enough). Break this box up into four equal pieces, which are also boxes:

  $$[-R, 0] \times [-R, 0], [0, R] \times [-R, 0], [-R, 0] \times [0, R], [0, R] \times [0, R]$$

  Since there are infinitely many z_k's in K and hence in B_1, at least one of these smaller boxes will necessarily also have infinitely many z_k's. Let B_2 be such a box. Then proceed similarly, break B_2 into four equal boxes/pieces as above, and let B_3 be one of these boxes that also has infinitely many z_k's. Proceed inductively. This way you get a sequence of boxes $B_1 \supseteq B_2 \supseteq B_3 \supseteq \cdots$ where the size of each box is getting smaller and smaller, precisely B_k is $R/2^{k-2}$ by $R/2^{k-2}$.

[3] "1" here is an arbitrary choice. Any number bigger than zero would work.

Since each B_k has infinitely many z_k's, we can inductively choose a subsequence $\{\tilde{z}_n\}$ of $\{z_k\}$ so that $\tilde{z}_1 \in B_1$, then $\tilde{z}_2 \in B_2$, then $\tilde{z}_3 \in B_3$, etc. The subsequence is Cauchy because for any $\varepsilon > 0$ let N be so that $R/2^{N-1} < \varepsilon/2$ and so for any $n > N$

$$\text{both } z_N, z_n \in B_N \text{ and so } |z_N - z_n| = \sqrt{((z_N)_x - (z_n)_x)^2 + \left((z_N)_y - (z_n)_y\right)^2} \leq \sqrt{\left(\frac{\varepsilon}{2}\right)^2 + \left(\frac{\varepsilon}{2}\right)^2}$$
$$= \varepsilon$$

Since $\mathbb{R}^2$ is complete, $\{\tilde{x}_n\}$ has a limit x. Since each $\tilde{x}_n \in K$, we have that $x \in \overline{K}$. Since K is closed and hence $K = \overline{K}$, we have that $x \in K$. Hence this is the $\{\tilde{x}_n\}$ and x that we were seeking.

■

- Now we can prove the **Extreme Value Theorem** which provides the theoretical justification for finding maximums and minimums of continuous functions over compact sets, such as what you did for closed and bounded intervals in calculus.

- **Theorem:** Suppose that $f : X \to \mathbb{R}$ is continuous and that $K \subseteq X$ is compact. Then there exists points $x_{\max} \in K$ and $x_{\min} \in K$ such that

$$f(x_{\max}) = \sup\{f(x) : x \in K\} \quad \text{and} \quad f(x_{\min}) = \inf\{f(x) : x \in K\}.$$

Proof: We'll start with proving that $x_{\max}$ exists. By a previous theorem, we know that the image $f[K]$ is a compact subset of $\mathbb{R}$. By the previous theorem, this means that $f[K]$ is closed and bounded. Bounded implies that the supremum

$$\sup\{f[K]\} = \sup\{f(x) : x \in K\}$$

exists.

Now, we claim that $\sup\{f[K]\} \in \partial(f[K])$. Take any $r > 0$, we will show that $B_{\mathbb{R}}(\sup\{f[K]\}; r) = (\sup\{f[K]\} - r, \sup\{f[K]\} + r)$ contains points inside and outside of $f[K]$. Points between $\sup\{f[K]\}$ and $\sup\{f[K]\} + r$ cannot be in $f[K]$ because $\sup\{f[K]\}$ is an upper bound. Next, there must be points in $f[K]$ between $\sup\{f[K]\} - r$ and $\sup\{f[K]\}$ or else $\sup\{f[K]\} - r$ would be a small upper bound, while $\sup\{f[K]\}$ is supposed to be the *smallest* upper bound. So indeed $\sup\{f[K]\} \in \partial(f[K])$.

Since we said that $f[K]$ is closed, $\partial(f[K]) \subseteq \overline{f[K]} = f[K]$ and so $\sup\{f[K]\} \in f[K]$. So there exists $x \in K$ so that $f(x) = \sup\{f[K]\}$. Hence x is the $x_{\max}$ that we wanted.

You can do a similar argument to prove that $x_{\min}$ exists. Or you can observe that $x_{\min}$ of f is $x_{\max}$ of $-f$. Choose your favorite approach!

■

- It turns out that one can generalize the classification of compact sets in $\mathbb{R}^n$ (i.e. being closed and bounded) to more general metric spaces, but one has to be careful. The following theorem provides the precise details. We won't cover it in detail and you're not allowed to use it on the homework or tests.

- **Theorem:** Suppose that (X, d_X) is a complete metric space and that $K \subseteq X$. Then K is compact if and only if it is closed and totally bounded ("totally bounded" means that for any radius $r > 0$, you can cover K by a finite number of balls of the form $B_X(x, r)$).

- **Remark:** The proof of the above theorem is essentially the same as the classification of compact sets $\mathbb{R}^n$, where the balls play the role of the boxes. Note the crucial assumption that X is a complete metric space. One way to see that this is necessary assumption is that the square $(-\pi, \pi) \times (-\pi, \pi)$ is both closed and bounded in $\mathbb{Q}^2$, but is not compact (this is not hard to show: it's on the level of an exercise).

Differentiation

- For a quantity that depends on time, the average rate of change is defined as

$$\frac{\text{change in the quantity}}{\text{time elapsed}}.$$

This is powerful, because if for instance the quantity in question is distance, knowing the average rate of change and the time elapsed allows you to compute the distance traveled. This, however, doesn't give you insight on what happens on smaller scales. The major contribution of differential calculus is to use the notion of limit to put the concept of "change" on the instantaneous level, which gives rise to the well-known concept of "instantaneous rate of change." This concept, turns out, also plays a central role in other subjects such as differential geometry.

- **Definition:** Suppose that $U \subseteq \mathbb{R}$ is an open subset and that $f : U \to \mathbb{R}$ is a function. For any point $a \in U$, the **derivative of $\boldsymbol{f}$ at $\boldsymbol{a}$** is defined as either of the following equivalent statements

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a},$$

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

if the limit exists. If the limit exists, we say that "f is differentiable at a." If f is differentiable at every point in U, we say that "f is differentiable (everywhere)" and we denote its derivative as a new function $f' : U \to \mathbb{R}$.

- **Remark:** The reason we require U to be open is that if we take $a \in U$, we have a full interval around a where we can freely move right and left to take values of f. This helps with the interpretation from calculus that the derivative gives the slope of the tangent line to the graph of f.

- **Remark:** The second derivative is the derivative of the first derivative, the third derivative is the derivative of the second derivative, and so on… For instance, we say that f is three-times differentiable if its third derivative exists everywhere. We denote the n^{th} derivative of f by $f^{(n)}$. If all of the derivatives of f exists, we say it's "infinitely differentiable." //

- **Remark:** One can also define **left-hand derivative** f_-' and **right-hand derivative** f_+' as

$$f_\pm'(a) = \lim_{x\to a^\pm}\frac{f(x)-f(a)}{x-a} = \lim_{x\to a^\pm}\frac{f(a+h)-f(a)}{h}.$$

It follows from a previous result that f is differentiable at a if and only if both $f_\pm'(a)$ exist and are equal, in which case $f'(a) = f_+'(a) = f_-'(a)$. The usefulness of this left and right derivatives is that you can extend the notion of derivative to closed intervals $[\alpha, \beta]$ by defining the derivative of f at α as the right-hand derivative and the derivative of f at β as the left-hand derivative. //

- **Proposition:** Suppose that $U \subseteq \mathbb{R}$ is open and consider the functions $f, g : \mathbb{R} \to \mathbb{R}$ where $f(x) = x$ and g is the constant function $g(x) = c$ (i.e. $c \in \mathbb{R}$ is some real number). Then both f and g are differentiable and $f'(x) = 1$ and $g'(x) = 0$.

Proof:

$$f'(a) = \lim_{x\to a}\frac{f(x)-f(a)}{x-a} = \lim_{x\to a}\frac{x-a}{x-a} = \lim_{x\to a} 1 = 1.$$

$$g'(a) = \lim_{x\to a}\frac{g(x)-g(a)}{x-a} = \lim_{x\to a}\frac{c-c}{x-a} = \lim_{x\to a} 0 = 0.$$

■

- **Example:** The function $f(x) = |x|$ is not differentiable at zero because the limit from the two sides are not equal:

$$f_+'(a) = \lim_{x\to 0^+}\frac{f(x)-f(0)}{x-0} = \lim_{x\to 0^+}\frac{|x|-0}{x-0} = \lim_{x\to 0^+} 1 = 1,$$

$$f_-'(a) = \lim_{x\to 0^-}\frac{f(x)-f(0)}{x-0} = \lim_{x\to 0^-}\frac{|x|-0}{x-0} = \lim_{x\to 0^-} -1 = -1.$$

//

- We mention a lemma on limits.

- **Lemma:** Suppose that $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ functions such that both $\lim_{x\to a} f(x)$ and $\lim_{x\to a} g(x)$ exist. Then

$$\lim_{x\to a}[f(x) \pm g(x)] = \lim_{x\to a}[f(x)] + \lim_{x\to a}[g(x)], \qquad \lim_{x\to a}[f(x)g(x)] = \lim_{x\to a}[f(x)]\lim_{x\to a}[g(x)],$$

$$\lim_{x\to a}[f(x)/g(x)] = \lim_{x\to a}[f(x)] \Big/ \lim_{x\to a}[g(x)],$$

where in the last statement we in addition require that $\lim_{x\to a} g(x) \neq 0$.

- **Proof:** Its proof is exactly the same as the analog version for limits of sequences that we discussed earlier (this lemma and proof also holds if you change "$U \subseteq \mathbb{R}^n$" to a metric space (X, d_X)).

  ∎

- **Theorem:** Suppose that $U \subseteq \mathbb{R}$ is open and that $f : U \to \mathbb{R}$ is differentiable at $a \in U$. Then f is continuous at $a \in U$. Obviously if it is differentiable everywhere (i.e. at all $a \in U$), then it is continuous everywhere.

  Proof: We will show that $f(a) = \lim_{x\to a} f(x)$ by showing that $\lim_{x\to a}[f(x) - f(a)] = 0$:

$$\lim_{x\to a}[f(x) - f(a)] = \lim_{x\to a} \frac{f(x) - f(a)}{x - a}(x - a) = \lim_{x\to a} \frac{f(x) - f(a)}{x - a} \lim_{x\to a}(x - a) = f'(a) \cdot 0 = 0.$$

  ∎

- **Remark:** The above theorem and proof work if $U = [a, b]$ is a closed interval and you use one-sided derivatives. //

- **Remark:** It follows immediately from the above theorem that if $U \subseteq \mathbb{R}$ is open and $f : U \to \mathbb{R}$ is n times differentiable, then $f, f', \ldots, f^{(n-1)} : U \to \mathbb{R}$ are continuous because the continuity of $f^{(n)}$ implies the continuity of $f^{(n-1)}$, continuity of $f^{(n-1)}$ implies the continuity of $f^{(n-2)}$, etc. //

- Next we prove the usual calculus rules surrounding derivatives (i.e. product rule, chain rule, etc.)

- **Theorem (Sum/Product Rule):** Suppose that $U \subseteq \mathbb{R}$ is open and that $f, g : U \to \mathbb{R}$ are differentiable at $a \in U$. Then $f + g$ and $f \cdot g$ are differentiable at a and

$$(f + g)'(a) = f'(a) + g'(a) \qquad \text{and} \qquad (f \cdot g)'(a) = f'(a)g(a) + f(a)g'(a).$$

  Obviously if they are differentiable everywhere (i.e. at all $a \in U$), then this holds everywhere.

  Proof: We will do $f \cdot g$, the proof for $f + g$ is left as an exercise. We have that

$$\begin{aligned}
\lim_{x\to a} \frac{f(x)g(x) - f(a)g(a)}{x - a} &= \lim_{x\to a} \frac{f(x)g(x) - f(a)g(x) + f(a)g(x) - f(a)g(a)}{x - a} \\
&= \lim_{x\to a} \frac{[f(x) - f(a)]g(x) + f(a)[g(x) - g(a)]}{x - a} \\
&= \lim_{x\to a}\left[\frac{f(x) - f(a)}{x - a} g(x)\right] + \lim_{x\to a}\left[f(x)\frac{g(x) - g(a)}{x - a}\right]
\end{aligned}$$

  Since f and g are continuous at a by the previous theorem, $f(a) = \lim_{x\to a} f(x)$ and $g(a) = \lim_{x\to a} g(x)$. So breaking up the above two limits and taking limits gives

$$= \lim_{x\to a} \frac{f(x) - f(a)}{x - a} \lim_{x\to a} g(x) + \lim_{x\to a} f(x) \lim_{x\to a} \frac{g(x) - g(a)}{x - a} = f'(a)g(a) + f(a)g'(a).$$

  ∎

- To prove the chain rule, we need a few lemmas which you'll prove on the homework.
- **Lemma:** Suppose that $f : U \to \mathbb{R}$ and $g : V \to \mathbb{R}$ are such that range $f \subseteq$ dom g (so it always makes sense to write $g\big(f(x)\big)$). If both $\lim_{x\to a} f(x)$ and $\lim_{y\to\lim_{x\to a} f(x)} g(y)$ exist, then

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

- **Lemma:** Suppose we have $g : V \to \mathbb{R}$ and a sequence $\{y_k\}$ in V. Suppose that $b = \lim_{k\to\infty} y_k$ exists, b is in V, and that each $y_k \neq b$. Suppose also that $\lim_{y\to\lim_{k\to\infty} y_k} g(y)$ also exists. Then

$$\lim_{k\to\infty} g(y_k) = \lim_{y\to\lim_{k\to\infty} y_k} g(y).$$

- **Theorem (Chain Rule):** Suppose that $U, V \subseteq \mathbb{R}$ are open and that $f : U \to \mathbb{R}$ and $g : V \to \mathbb{R}$ are such that range $f \subseteq$ dom g. Suppose f is differentiable at $a \in U$ and that g is differentiable at $f(a)$. Then the composition $g \circ f = g\big(f(\cdot)\big) : U \to \mathbb{R}$ is differentiable at a and

$$(g \circ f)'(a) = g'\big(f(a)\big)f'(a).$$

Proof: Suppose first that there exists a small interval $(a-\delta, a+\delta)$ around a contained in U such that $f(x) \neq f(a)$ for $x \in (a-\delta, a+\delta)$ and $x \neq a$. Then

$$(4) \qquad \lim_{x\to a} \frac{g\big(f(x)\big) - g\big(f(a)\big)}{x-a} = \lim_{x\to a} \frac{g\big(f(x)\big) - g\big(f(a)\big)}{f(x)-f(a)} \cdot \frac{f(x)-f(a)}{x-a}$$

$$= \lim_{x\to a} \text{Plug in } f(x) \text{ into } y \text{ in } \frac{g(y)-g(a)}{y-a} \lim_{x\to a} \frac{f(x)-f(a)}{x-a} = g'\big(f(a)\big)f'(a)$$

(we used the second to last lemma for the blue items). Next suppose such a small interval $(a-\delta, a+\delta)$ around a does not exist. In other words

$$\forall\delta > 0 \quad \exists x \in (a-\delta, a+\delta) : x \neq a, \;\; f(x) = f(a).$$

By a technique we've done many times before (e.g. setting $\delta = 1/k$), we can construct a sequence $\{x_k\}$ in U so that $a = \lim_{x\to\infty} x_k$, each $f(x_k) = f(a)$, and each $x_k \neq a$. By the previous lemma

$$f'(a) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} = \lim_{k\to\infty} \frac{f(x_k)-f(a)}{x_k - a} = \lim_{x\to a} 0 = 0.$$

So we just have to show that $(g\circ f)'(a) = 0$. We will show this by showing that the quantity inside the first limit in (4) is arbitrarily small. Take any $\varepsilon > 0$. There exists $\delta_1, \delta_2, \delta_3 > 0$ such that

$$(5) \qquad \forall x \in U : x \neq a \text{ and } |x-a| < \delta_1, \quad \left|\frac{f(x)-f(a)}{x-a}\right| < \varepsilon$$

$$(6) \qquad \forall y \in V : y \neq f(a) \text{ and } |y - f(a)| < \delta_2, \quad \left|\frac{g(y)-g\big(f(a)\big)}{y-a} - g'\big(f(a)\big)\right| < 1$$

(7) $\forall x \in U : x \neq a$ and $|x-a| < \delta_3,\ \ |f(x)-f(a)| < \delta_2$

Note that the last item in (6) and the triangle inequality imply that

$$\left|\frac{g(y)-g\big(f(a)\big)}{y-a}\right| \leq \left|\frac{g(y)-g\big(f(a)\big)}{y-a} - g'\big(f(a)\big)\right| + \left|g'\big(f(a)\big)\right| < 1 + \left|g'\big(f(a)\big)\right|.$$

Then for any $x \in U : x \neq f(a)$ and $|x-a| < \min\{\delta_1, \delta_3\}$

$$\left|\frac{g\big(f(x)\big)-g\big(f(a)\big)}{x-a} - 0\right| = \begin{cases} \dfrac{g\big(f(x)\big)-g\big(f(a)\big)}{f(x)-f(a)} \cdot \dfrac{f(x)-f(a)}{x-a} & \text{if } f(x) \neq f(a) \\ 0 & \text{if } f(x) = f(a) \end{cases}$$

$$\leq \begin{cases} \left[1+\left|g'\big(f(a)\big)\right|\right]\varepsilon & \text{if } f(x) \neq f(a) \\ \left[1+\left|g'\big(f(a)\big)\right|\right]\varepsilon & \text{if } f(x) = f(a) \end{cases} = \left[1+\left|g'\big(f(a)\big)\right|\right]\varepsilon.$$

Oops… from here we see that had we divided our original ε by $\left[1+\left|g'\big(f(a)\big)\right|\right]$, we would get that this is less than ε. Hence, as mentioned above, this proves the theorem.

■

- **Lemma:** The derivative of the function $1/x$ (defined over $\mathbb{R} \setminus \{0\}$) is $-1/x^2$.

  Proof: Homework. ■

- **Theorem:** Suppose that $U \subseteq \mathbb{R}$ is open and that $f, g : U \to \mathbb{R}$ are differentiable at $a \in U$ and $g(a) \neq 0$. Then f/g is differentiable at a and

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{\big(g(a)\big)^2}.$$

  Proof: Let $h_D : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ be the function $h_D(x) = 1/x$. Then by the product rule, chain rule, and the previous lemma (here I omit writing "(a)")

$$\left(\frac{f}{g}\right)' = f \cdot (h_D \circ g) = f' \cdot (h_D \circ g) + f \cdot (h_D \circ g)' = \frac{f'}{g} + f \cdot \left(-\frac{1}{g^2} \cdot g'\right) = \frac{f'g - fg'}{g^2}.$$

■

- Next we prove a theorem that justifies finding maximums and minimums by taking the derivative of a function and setting it to zero.

- **Definition:** Suppose that U is open and that $f : U \to \mathbb{R}$ is differentiable. We say that a is a **local maximum** of f if there exists a small interval $(a-\delta, a+\delta)$ contained in U so that

$$f(x) \leq f(a) \quad \forall x \in (a-\delta, a+\delta)$$

The point a being a **local minimum** is defined similarly but change "≤" to "≥."

Similarly, a is called a **global maximum** if

$$f(x) \leq f(a) \quad \forall x \in U$$

The point a being a **global minimum** is defined similarly but change "≤" to "≥."

Obviously a global maximum/minimum is a local maximum/minimum, but the other direction is not true.

- **Lemma:** If $f, g : U \to \mathbb{R}$ are such that $f \leq g$ and both $\lim_{x \to a^{\pm}} f(x)$ and $\lim_{x \to a^{\pm}} g(x)$ exist, then $\lim_{x \to a^{\pm}} f(x) \leq \lim_{x \to a^{\pm}} g(x)$ (you can remove the "±" as well).

  Proof: Similar to the proof of the analogous theorem you proved on the homework for sequences. ∎

- **Lemma:** Suppose that U is open and that $f : U \to \mathbb{R}$ is a function. Suppose that f is differentiable at $a \in U$ and that a is a local maximum or local minimum. Then $f'(a) = 0$. Since global maximums/minimums are also local maximums/minimums, this also works for global maximums/minimums.

  Proof: We'll do the local maximum case: the other case is done similarly. We have that

$$f'_+(a) = \lim_{x \to a^+} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a^+} \frac{\left(\text{something smaller than } f(a)\right) - f(a)}{\text{something positive}} \leq \lim_{x \to a^+} 0 = 0,$$

$$f'_-(a) = \lim_{x \to a^-} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a^-} \frac{\left(\text{something smaller than } f(a)\right) - f(a)}{\text{something negative}} \geq \lim_{x \to a^-} 0 = 0.$$

  Since $f'(a) = f'_\pm(a)$, we have that $f'(a) = 0$.

  ∎

- Now we build towards the Mean Value Theorem.

- **Rolle's Theorem:** Suppose that $f : [a, b] \to \mathbb{R}$ is continuous, differentiable on (a, b) and $f(a) = f(b)$. Then there exists a $c \in (a, b)$ such that $f'(c) = 0$.

  Proof: By the extreme value theorem, f has a global maximum and global minimum. If $f(a) = f(b)$ is both the global maximum and minimum, then f is constant and so $f'(c) = 0$ for all $c \in (a, b)$. Otherwise, there exists a global maximum or minimum $c \in (a, b)$. By the previous lemma, $f'(c) = 0$.

  ∎

- **Mean Value Theorem I:** Suppose that $f : [a, b] \to \mathbb{R}$ is continuous and differentiable on (a, b). Then there exists a $c \in (a, b)$ such that

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

Proof: Let $l : [a, b] \to \mathbb{R}$ be the line

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

Observe that

$$l(a) = f(a) \quad \text{and} \quad l(b) = f(a) + \frac{f(b) - f(a)}{b - a}(b - a) = f(b) \quad \text{and} \quad l'(x) = \frac{f(b) - f(a)}{b - a}.$$

Consider the function $h : [a, b] \to \mathbb{R}$ given by $h(x) = f(x) - l(x)$. Then

$$h(a) = f(a) - l(a) = f(a) - f(a) = 0 \quad \text{and} \quad h(b) = f(b) - l(b) = f(b) - f(b) = 0$$

and thus $h(a) = h(b)$. So by Rolle's Theorem there exists $c \in (a, b)$ such that

$$0 = h'(c) = f'(c) - l'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}$$

and so

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

■

- **Lemma:** Suppose that $g : (a, b) \to \mathbb{R}$ is differentiable and $g' > 0$. Then g is strictly increasing (i.e. $x_1 < x_2$ implies $g(x_1) < g(x_2)$). Same thing holds if you change the assumption to "$g' < 0$" and the conclusion to "g is strictly decreasing."

  Proof: We'll do the case $g' > 0$, the other case $g' < 0$ is proved very similarly. Suppose not! Then there exists $x_1 < x_2$ such that $g(x_1) \geq g(x_2)$. By the Mean Value Theorem I, there exists $c \in (a, b)$ such that

$$g'(c) = \frac{g(x_2) - g(x_1)}{x_2 - x_1} \leq 0,$$

  contradiction!

■

- **Mean Value Theorem II:** Suppose that $f, g : [a, b] \to \mathbb{R}$ is continuous and differentiable on (a, b) and that g' is not zero on (a, b). Then there exists a $c \in (a, b)$ such that

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

Proof: Homework to read: Theorem 2.9 in Advanced Calculus Second Edition by Gerald B. Folland (click on the link to see the book on the author's website) It should be an easy read: so you are responsible for it. ■

- **L'Hôpital's Rule I:** Suppose $U \subseteq \mathbb{R}$ is open and that $f, g : U \to \mathbb{R}$ are differentiable. Suppose that for some $a \in U$

$$\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0$$

Suppose also that g' is not zero on some small interval $(a - \delta, a + \delta) \subseteq U$ except possibly at $x = a$ and that the limit

$$\lim_{x \to a} \frac{f'(x)}{g'(x)}$$

exists. Then g is also not zero on $(a - \delta, a + \delta)$ except at $x = a$ and

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

Proof: Since f and g are differentiable, they are continuous and so

$$f(a) = \lim_{x \to a} f(x) = 0 \quad \text{and} \quad g(a) = \lim_{x \to a} g(x) = 0.$$

We now focus only on $x \in (a - \delta, a + \delta)$. For such x, we cannot have $g(x) = 0$ or else by Rolles' theorem there would be a c between a and x such that $g'(c) = 0$ and we said that g' is not zero on $(a - \delta, a + \delta)$ except possibly at $x = a$. So indeed g is also not zero on $(a - \delta, a + \delta)$ except at $x = a$. We will show that

$$\lim_{x \to a^\pm} \frac{f(x)}{g(x)} = \lim_{x \to a^\pm} \frac{f'(x)}{g'(x)}$$

from which the theorem will follow by the equality of the limits from both sides. Let's start with $x \to a^+$. Take any $x \in (a - \delta, a + \delta)$ such that $x > a$. Then there exists some $c_x \in (a, x)$ such that

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

Notice that since c_x is stuck in between a and x, $c_x \to a^+$ as $x \to a^+$ and so this will force

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = \lim_{x \to a^+} \frac{f'(c_x)}{g'(c_x)} = \lim_{x \to a^+} \frac{f'(x)}{g'(x)}.$$

The case $x \to a^-$ is done similarly.

■

- **Homework:** Read the statements of Theorem 2.10 and Theorem 2.11 in Advanced Calculus Second Edition by Gerald B. Folland (click on the link to see the book on the author's website). If you wish, you may read the proofs, but you are not responsible for the proofs.

Integration

- **Notation:** Suppose that $\{y_k\}_{k=1}^{m}$ is a set of numbers. We define the **sum notation** Σ:

$$\sum_{k=1}^{m} y_k = y_1 + y_2 + \cdots + y_m.$$

Of course, this can be done for other objects as well that come with the addition operation, such as when y_k are vectors, functions, etc.

- **Examples:** If we take $\{y_k\}_{k=1}^{5}$ where each $y_k = k^2$, then

$$\sum_{k=1}^{5} y_k = y_1 + y_2 + \cdots + y_5 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 55$$

People often don't make any mention of y_k and simply write

$$\sum_{k=1}^{5} k^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 55.$$

- We now embark on constructing the integral. Intuitively it is an infinite sum of objects that are defined on the instantaneous scale. One use of this is to invert differentiation, such as computing total distance traveled from knowledge of instantaneous rate of change. Recall from calculus that integrals also give areas under curves and hence have application in geometry (or in higher dimensions they give volumes). In fact, areas under curves is the idea behind how we will define integrals.

  We will define the **Riemann integral**, which integrates bounded function $f : [a, b] \to \mathbb{R}$ defined over a compact interval $[a, b]$. Integration is computed by approximations obtained by breaking up the axis of the independent variable (i.e. $[a, b]$) into small intervals and then making the length of the intervals smaller and smaller. The following definition is the starting point for this.

- **Definition:** A **partition** of an interval $[a, b] \subseteq \mathbb{R}$ is a collection of points

$$P = \{x_0, x_1, \dots, x_J\}$$

that have the property that $a = x_0 < x_1 < \cdots < x_J = b$. Another partition P' is called a **refinement** of P if $P \subseteq P'$ (i.e. $P' = \{x_0', x_1', \dots, x_{J'}'\}$ is a set of points that contains the points of $P = \{x_0, x_1, \dots, x_J\}$ and has even more points in it!)

- Now we define upper and lower Riemann sums which approximate the true integral from above and below respectively.

- **Definition:** Suppose $f : [a, b] \to \mathbb{R}$ is bounded and that P is a partition of $[a, b]$. The **upper Riemann sum** of f with partition $P = \{x_0, \dots, x_J\}$ is defined as

$$(8) \qquad S_P f = \sum_{j=0}^{J-1} \sup_{x \in [x_j, x_{j+1}]} \{f(x)\} \left(x_{j+1} - x_j\right)$$

$$= \sup_{x \in [x_0, x_1]} \{f(x)\} (x_1 - x_0) + \sup_{x \in [x_1, x_2]} \{f(x)\} (x_2 - x_1) + \cdots + \sup_{x \in [x_{J-1}, x_J]} \{f(x)\} \left(x_J - x_{J-1}\right).$$

The **lower Riemann sum** of f with partition P is defined similarly but using the infimum:

$$(9) \qquad s_P f = \sum_{j=0}^{J-1} \inf_{x \in [x_j, x_{j+1}]} \{f(x)\} \left(x_{j+1} - x_j\right).$$

Observe that since each

$$\inf_{x \in [x_j, x_{j+1}]} \{f(x)\} \leq \sup_{x \in [x_j, x_{j+1}]} \{f(x)\},$$

we immediately get that the lower Riemann sum is smaller than the upper Riemann sum: $s_P f \leq S_P f$.

- Intuitively speaking, the following lemma says that taking smaller intervals in the partition improves the approximation of the upper and lower Riemann sums to the true integral.

- **Lemma:** Suppose that $f : [a, b] \to \mathbb{R}$ is bounded. Suppose that P is a partition of $[a, b]$ and that P' is a refinement of it (i.e. $P \subseteq P'$). Then

$$S_P f \geq S_{P'} f \qquad \text{and} \qquad s_P f \leq s_{P'} f.$$

Proof: We will do $S_P f \geq S_{P'} f$, the proof for $s_P f \leq s_{P'} f$ is very similar. Let $P = \left\{x_0, \ldots, x_J\right\}$ and $P' = \left\{x_1', \ldots x_{J'}'\right\}$. Take any interval $\left[x_j, x_{j+1}\right]$ in the partition P. Take all of the $x_k', \ldots, x_m'$ in P' such that

$$x_j = x_k' < x_{k+1}' < \cdots < x_{m-1}' < x_m' = x_{j+1}.$$

Intuitively, we broke up $\left[x_j, x_{j+1}\right]$ into subintervals of the form $[x_r', x_{r+1}']$. We now analyze every term in the sum (8):

$$\sup_{x \in [x_j, x_{j+1}]} \{f(x)\} \left(x_{j+1} - x_j\right) = \sup_{x \in [x_j, x_{j+1}]} \{f(x)\} (x_m' - x_k')$$

$$= \sup_{x \in [x_j, x_{j+1}]} \{f(x)\} [(x_m' - x_{m-1}') + (x_{m-1}' - x_{m-2}') + \cdots + (x_{k+1}' - x_k')]$$

$$= \sup_{x \in [x_j, x_{j+1}]} \{f(x)\} (x_m' - x_{m-1}') + \cdots + \sup_{x \in [x_j, x_{j+1}]} \{f(x)\} (x_{k+1}' - x_k').$$

Since each $[x_r', x_{r+1}']$ subinterval considered here is contained in $\left[x_j, x_{j+1}\right]$, we have that

$$\sup_{x \in [x_j, x_{j+1}]} \{f(x)\} \geq \sup_{x \in [x_r', x_{r+1}']} \{f(x)\}.$$

Hence from the previous calculation we get that

$$\sup_{x\in[x_j,x_{j+1}]}\{f(x)\}\left(x_{j+1}-x_j\right)$$
$$\geq \sup_{x\in[x'_{m-1},x'_m]}\{f(x)\}\,(x'_m-x'_{m-1})+\cdots+\sup_{x\in[x'_k,x'_{k+1}]}\{f(x)\}\,(x'_{k+1}-x'_k).$$

Observe that adding up the left-side for all $j\in\{0,\ldots,J-1\}$ gives $S_P f$ while the right-hand side will add up to $S_{P'}f$. Hence indeed $S_P f\geq S_{P'}f$.

■

- The next lemma says that lower Riemann sums are smaller than upper Riemann sums regardless of the partition you choose for either.
- **Lemma:** Suppose that $f:[a,b]\to\mathbb{R}$ is bounded and that P and Q are partitions of $[a,b]$. Then $s_P f\leq S_Q f$.

  Proof: Consider the partition $P\cup Q$ of $[a,b]$. Then by the previous lemma,

$$s_P f\leq s_{P\cup Q}f\leq S_{P\cup Q}f\leq S_Q f.$$

■

- **Definition:** Suppose that $f:[a,b]\to\mathbb{R}$ is bounded. The **upper** and **lower Riemann integrals** are defined as

$$\underline{I}_a^b f=\sup_P\{s_P f\}\qquad\text{and}\qquad \overline{I}_a^b f=\inf_Q\{S_Q f\}.$$

  Note that since $s_P f\leq S_Q f$ by the previous lemma and "sup" and "inf" preserve "$\leq$" (this takes a little thought to see why), we get that $\underline{I}_a^b f\leq\overline{I}_a^b f$ always. If $\overline{I}_a^b f=\underline{I}_a^b f$, then we say that f is **Riemann integrable** (on $[a,b]$) and we define the Riemann integral of f as

$$\int_a^b f(x)dx=\overline{I}_a^b f=\underline{I}_a^b f.$$

  Since we don't consider other types of integrals in this course for a while (e.g. the Lebesgue integral), for now we sometimes don't write "Riemann" in "Riemann integral."

- The idea behind the above definition is that intuitively speaking the area under the graph of f (i.e. the integral $\int_a^b f$) must be less than or equal to all upper Riemann sums and bigger than or equal to all lower Riemann sums. Thus the integral $\int_a^b f$ must be stuck between $\underline{I}_a^b f$ and $\overline{I}_a^b f$. If the two are not equal: $\underline{I}_a^b f<\overline{I}_a^b f$, then we don't know what $\int_a^b f$ should be and so we can't define it. If the two are equal: $\underline{I}_a^b f=\overline{I}_a^b f$, then we define $\int_a^b f$ to be that number.

- **Notation:** If $b < a$, then we define

$$\int_b^a f(x)dx = -\int_a^b f(x)dx.$$

  This is simply a convention, that has proved to be useful.

- Next we explore standard properties of integrals.

- **Theorem:**

  a) Suppose that $a < b < c$ and that $f : [a,c] \to \mathbb{R}$ is Riemann integrable on $[a,b]$ and $[b,c]$ separately. Then f is integrable on all of $[a,c]$ and

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

  b) Suppose that $f, g : [a,b] \to \mathbb{R}$ are Riemann integrable. Then $f + g : [a,b] \to \mathbb{R}$ is Riemann integrable and

$$\int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

Proof: We begin with a). In our proof of a), let P's denote partitions of $[a,b]$, Q's denote partitions of $[b,c]$, and R's denote partitions of $[a,c]$. Observe that for any P and Q, $P \cup Q$ is a partition of $[a,c]$ and

$$s_P f + s_Q f = s_{P\cup Q} f \leq \sup_R s_R f. \tag{10}$$

We claim that this implies that

$$\sup_P s_P f + \sup_Q s_Q f \leq \sup_R s_R f. \tag{11}$$

Essentially we threw $\sup_P$ and $\sup_Q$ onto $s_P f$ and $s_Q f$ respectively in (10). Let's show how to justify this. From (10) we have that

$$s_P f \leq \sup_R s_R f - s_Q f.$$

So the right-hand side is an upper bound of the left-hand side for all P. Since $\sup_P s_P f$ is the least such upper bound we get that $\sup_P s_P f \leq \sup_R s_R f - s_Q f$ and so

$$\sup_P s_P f + s_Q f \leq \sup_R s_R f.$$

Throwing on $\sup_Q$ onto $s_Q f$ is done similarly, and hence we indeed get (11). Note that (11) is another way of writing

$$\underline{I}_a^b f + \underline{I}_b^c f \leq \underline{I}_a^c f.$$

This same argument but using upper Riemann sums and infimums gives $\overline{I}_a^c f \leq \overline{I}_a^b f + \overline{I}_b^c f$. Thus, since $\underline{I}_a^c f \leq \overline{I}_a^c f$ we get that

$$(12) \qquad \underline{I}_a^b f + \underline{I}_b^c f \leq \underline{I}_a^c f \leq \overline{I}_a^c f \leq \overline{I}_a^b f + \overline{I}_b^c f.$$

Since we said that f is Riemann integrable on $[a, b]$ and $[b, c]$, we have that

$$\underline{I}_a^b f = \overline{I}_a^b f = \int_a^b f(x)dx \quad \text{and} \quad \underline{I}_b^c f = \overline{I}_b^c f = \int_b^c f(x)dx$$

which by (12) forces

$$(13) \qquad \underline{I}_a^c f = \overline{I}_a^c f = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

Hence indeed f is Riemann integrable on $[a, c]$ and $\int_a^c f(x)dx$ is equal to (13).

Next let's prove b). Take any partition P of $[a, b]$. Let $P = \{x_0, \dots, x_J\}$. Consider one of the intervals $[x_j, x_{j+1}]$. Observe that for any $\tilde{x} \in [x_j, x_{j+1}]$

$$\inf_{x\in[x_j,x_{j+1}]}\{f(x)\} + \inf_{x\in[x_j,x_{j+1}]}\{g(x)\} \leq f(\tilde{x}) + g(\tilde{x}).$$

Thus the left-hand side is a lower bound for all values of the form on the right-hand side, and so

$$\inf_{x\in[x_j,x_{j+1}]}\{f(x)\} + \inf_{x\in[x_j,x_{j+1}]}\{g(x)\} \leq \inf_{\tilde{x}\in[x_j,x_{j+1}]}\{f(\tilde{x}) + g(\tilde{x})\}.$$

We can replace $\tilde{x}$ with x on the right-hand side since it doesn't matter what letter we use. Thus

$$\sum_{j=0}^{J-1}\left[\inf_{x\in[x_j,x_{j+1}]}\{f(x)\} + \inf_{x\in[x_j,x_{j+1}]}\{g(x)\}\right](x_{j+1} - x_j) \leq \sum_{j=0}^{J-1}\inf_{x\in[x_j,x_{j+1}]}\{f(x) + g(x)\}(x_{j+1} - x_j)$$

which is the same thing as

$$s_P f + s_P g \leq s_P(f + g).$$

We can't simply throw $\sup_P$ onto both sides because it's not clear how $\sup_P$ breaks over the "+" sign on the left-hand side. So consider another partition Q of $[a, b]$ and observe that the above proves that (here R is any partition of $[a, b]$)

$$s_P f + s_Q g \leq s_{P\cup Q} f + s_{P\cup Q} g \leq s_{P\cup Q}\,(f+g) \leq \sup_R s_R(f+g).$$

Now we can throw on $\sup_P$ and $\sup_Q$ onto the left-hand side as before to get

$$\sup_P s_P f + \sup_Q s_Q g \leq \sup_R s_R(f+g),$$

which is equivalent to

$$\underline{I}_a^b f + \underline{I}_a^b g \leq \underline{I}_a^b (f+g).$$

Repeating the above arguments but with upper Riemann sums and using "inf", one similarly gets $\overline{I}_a^b(f+g) \leq \overline{I}_a^b f + \overline{I}_b^b g$. Since $\underline{I}_a^b(f+g) \leq \overline{I}_a^b(f+g)$ we have that

$$\underline{I}_a^b f + \underline{I}_a^b g \leq \underline{I}_a^b(f+g) \leq \overline{I}_a^b(f+g) \leq \overline{I}_a^b f + \overline{I}_b^b g.$$

As before, since f and g are Riemann integrable on $[a,b]$ we have that

$$\underline{I}_a^b f = \overline{I}_a^b f = \int_a^b f(x)dx \quad \text{and} \quad \underline{I}_a^b g = \overline{I}_a^b g = \int_a^b g(x)dx$$

which by the previous inequality forces

$$\underline{I}_a^b(f+g) = \overline{I}_a^b(f+g) = \int_a^b f(x)dx + \int_b^c f(x)dx. \tag{14}$$

Thus indeed $f+g$ is Riemann integrable on $[a,b]$ and $\int_a^b [f(x)+g(x)]dx$ is equal to (14).

■

- **Remark:** There are a few more standard properties of integrals that you should already know from calculus, and I will most likely ask you to prove in the homework. The are written out in Theorem 4.9 in *Advanced Calculus 2nd Ed* by Gerald Folland:

  https://sites.math.washington.edu/~folland/Homepage/AdvCalc24.pdf

  You are responsible for reading the statement of the theorem.

- **Remark:** Using the convention that $\int_a^b f(x)dx = -\int_b^a f(x)dx$, it's easy to see that part a) of the above theorem holds for any triple of numbers a, b, c (i.e. not necessarily $a < b < c$). This is why this convention was chosen.

- We've defined Riemann integrable functions, but we haven't given any examples of such functions. To do this, we will show that all continuous functions are Riemann integrable. To prove this, we need a stronger notion of continuity:

- **Definition:** Suppose (X, d_X) and (Y, d_Y) are metric spaces. We say that $f : X \to Y$ is **uniformly continuous** if

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall x \in X \;\; \forall y \in X : d_X(x, y) < \delta, \;\; d_Y\big(f(x), f(y)\big) < \varepsilon.$$

Theorem: Suppose (X, d_X) and (Y, d_Y) are metric spaces and that X is furthermore compact. Then any continuous function $f : X \to Y$ is also uniformly continuous.

Proof: Take any continuous $f : X \to Y$. We need to show that it is uniformly continuous. Suppose not! Then

$$\exists \varepsilon > 0 \;\; \forall \delta > 0 \;\; \exists x \in X \;\; \exists y \in Y : d_X(x, y) < \delta, \;\; d_Y\big(f(x), f(y)\big) \geq \varepsilon.$$

Fix such a $\varepsilon > 0$, and for each $\delta_k = 1/k$ let $x_k, y_k \in X$ be such that $d_X(x_k, y_k) < \delta_k = 1/k$ and $d_Y\big(f(x), f(y)\big) \geq \varepsilon$. Since X is compact, there exists a subsequence $\{\tilde{x}_n\}$ of $\{x_k\}$ that converges to some point $x \in X$. Take the corresponding subsequence $\{\tilde{y}_n\}$ of $\{y_k\}$ that uses the same sub-indexing as $\{\tilde{x}_n\}$. We claim that x is also the limit of $\{\tilde{y}_n\}$. To see why, let k_n be the corresponding integer such that $d_X(\tilde{x}_n - \tilde{y}_n) < 1/k_n$ and observe that

$$d_X(x, \tilde{y}_n) \leq d_X(x, \tilde{x}_n) + d_X(\tilde{x}_n, \tilde{y}_n) \leq d_X(x, \tilde{x}_n) + \frac{1}{k_n} \to 0 \quad \text{as} \quad n \to \infty.$$

So indeed $x = \lim_{n\to\infty} \tilde{y}_n$ as well. Now, we have by the (sequential) continuity of d_Y and f

$$\lim_{n\to\infty} d_Y\big(f(\tilde{x}_n), f(\tilde{y}_n)\big) = d_Y\big(f(x), f(x)\big) = 0$$

But at the same time each $d_Y\big(f(\tilde{x}_n), f(\tilde{y}_n)\big) \geq \varepsilon$ by construction. Hence contradiction!

■

- We finally prove the existence of integrable functions:

- **Theorem:** Suppose that $f : [a, b] \to \mathbb{R}$ is continuous. Then f is Riemann integrable.

Proof: The function f if bounded by the extreme value theorem. Next, we will show that

$$\overline{I}_a^b f = \inf_P S_P f = \sup_Q s_Q f = \underline{I}_a^b f$$

by showing that for any $\varepsilon > 0$

$$\left| \inf_P S_P f - \sup_Q s_Q f \right| \leq \varepsilon.$$

Note that we can remove the absolute values here since $\inf_P S_P f \geq \sup_Q s_Q f$ automatically. By the previous theorem, f is uniformly continuous and so

$$\exists \delta > 0 \;\; \forall x \in [a, b] \;\; \forall y \in [a, b] : |x - y| < \delta, \;\; |f(x) - f(y)| < \frac{\varepsilon}{b - a}.$$

Let $P = \{x_0, \dots, x_J\}$ be a partition of $[a, b]$ so that the length of each $[x_j, x_{j+1}]$ subinterval is less than δ. Fix a subinterval $[x_j, x_{j+1}]$. For any $x, y \in [x_j, x_{j+1}]$ we have that $|x - y| < \delta$ by construction and so

$$f(x) - f(y) < |f(x) - f(y)| < \frac{\varepsilon}{b-a}.$$

Let's rewrite this as

$$f(x) + \big(-f(y)\big) < \frac{\varepsilon}{b-a}$$

As we did in the second to last theorem, we can throw $\sup_{x\in[x_j,x_{j+1}]}$ and $\sup_{y\in[x_j,x_{j+1}]}$ onto the two quantities on the left-hand side to get

$$\sup_{x\in[x_j,x_{j+1}]}\{f(x)\} + \sup_{y\in[x_j,x_{j+1}]}\{-f(y)\} \le \frac{\varepsilon}{b-a}$$

Using the homework result that $\sup\{-\,\cdot\} = -\inf\{\,\cdot\,\}$, we get that

$$\sup_{x\in[x_j,x_{j+1}]}\{f(x)\} - \inf_{y\in[x_j,x_{j+1}]}\{f(y)\} \le \frac{\varepsilon}{b-a}.$$

Multiplying this equation through by $(x_{j+1} - x_j)$ and then summing in j gives

$$\sum_{j=0}^{J-1} \sup_{x\in[x_j,x_{j+1}]}\{f(x)\}\,(x_{j+1} - x_j) - \sum_{j=0}^{J-1} \inf_{y\in[x_j,x_{j+1}]}\{f(y)\}\,(x_{j+1} - x_j) \le \sum_{j=0}^{J-1} \frac{\varepsilon}{b-a}(x_{j+1} - x_j),$$

$$\Longrightarrow \quad S_P f - s_P f \le \frac{\varepsilon}{b-a}(b-a) = \varepsilon.$$

Now, we have that

$$(15) \qquad \inf_Q S_Q f \le S_P f \quad \text{and} \quad \sup_Q s_Q f \ge s_P f \quad (\text{hence } -\sup_Q s_Q f \le -s_P f).$$

So the previous inequality gives

$$\inf_Q S_Q f - \sup_Q s_Q f \le \varepsilon.$$

As argued before, this proves the theorem.

■

- An important corollary is the following.
- **Corollary:** Suppose that $f : [a, b] \to \mathbb{R}$ is bounded and continuous except at finitely many points. Then f is Riemann integrable.

Proof: Let $B > 0$ be such that $|f| < B$ everywhere, which exists since f is bounded. Like in the previous proof, we will show that for any $\varepsilon > 0$

$$\inf_P S_P f - \sup_Q s_Q f \leq \varepsilon.$$

Let $\{y_1, \dots, y_m\}$ be the points where f is discontinuous. Take any $\delta < \varepsilon$ (this δ has nothing to do with the δ in the previous proof). If we make $\delta > 0$ small enough, we can write

$$[a, b] \setminus \bigcup_{k=1}^{m} (y_k - \delta, y_k + \delta) = [x_0, x_0'] \cup [x_1, x_1'] \cup \dots \cup [x_m, x_m']$$

for some disjoint intervals $[x_j, x_j']$ (i.e. they don't intersect). The function f is continuous on each $[x_j, x_j']$, hence integrable, and hence we can choose partitions P_j and Q_j of $[x_j, x_j']$ so that

$$S_{P_j} f < \int_{x_j}^{x_j'} f + \varepsilon \qquad \text{and} \qquad s_{Q_j} f > \int_{x_j}^{x_j'} f - \varepsilon.$$

Since $P_j \cup Q_j$ are a refinement of both P_j and Q_j, by a previous lemma $S_{P_j \cup Q_j} f \leq S_{P_j} f$ and $s_{P_j \cup Q_j} f \geq S_{Q_j} f$ and so furthermore

$$S_{P_j \cup Q_j} f < \int_{x_j}^{x_j'} f + \varepsilon \qquad \text{and} \qquad s_{P_j \cup Q_j} f > \int_{x_j}^{x_j'} f - \varepsilon.$$

Consider the partition $R = P_0 \cup \dots \cup P_m \cup Q_0 \cup \dots \cup Q_m$ of $[a, b]$ whose upper Riemann sum of f will satisfy

$$S_R f = S_{P_0 \cup Q_0} f + \dots + S_{P_m \cup Q_m} f + \sup_{x \in [y_1 - \delta, y_1 + \delta]} \{f\}\, 2\delta + \dots + \sup_{x \in [y_m - \delta, y_m + \delta]} \{f\}\, 2\delta$$

$$< \left(\int_{x_0}^{x_0'} f + \varepsilon \right) + \cdots \left(\int_{x_m}^{x_m'} f + \varepsilon \right) + B2\varepsilon + \dots + B2\varepsilon$$

$$= \int_{x_0}^{x_0'} f + \dots + \int_{x_m}^{x_m'} f + [(m + 1) + m2B]\varepsilon.$$

One can similarly show that

$$s_R f > \int_{x_0}^{x_0'} f + \dots + \int_{x_m}^{x_m'} f - [(m + 1) + m2B]\varepsilon.$$

Hence

$$S_R f - s_R f < 2[(m+1) + m2B]\varepsilon$$

As near the end of the previous proof, using (15) we get that

$$(16) \qquad \inf_Q S_Q f - \sup_Q s_Q f \leq 2[(m+1) + m2B]\varepsilon.$$

Oh no: we didn't get the right-hand side to be ε! To fix this, go back in this proof and divide the ε's in the right places by $2[(m+1)+m2B]$ to get the right-hand side of (16) to be ε. As argued above, this proves the theorem.

■

- Now we get to the theorem with the dramatic name! To prove it we need the following lemma:
- **Lemma:** Take the constant function $h : [a,b] \to \mathbb{R}$ given by $h(x) = c$. Then $\int_a^b h(x)dx = c(b-a)$ (this should be obvious from calculus, we are simply proving this rigorously).

  Proof: Since h is constantly c, for any partition $P = \{x_0, \dots, x_J\}$ of $[a,b]$ the infimum and supremum of h over any interval $[x_j, x_{j+1}]$ is also c. Hence both $S_P f$ and $s_P f$ are equal to

$$\sum_{j=0}^{J-1} c(x_{j+1} - x_j) = c\sum_{j=0}^{J-1}(x_{j+1} - x_j) = c(b-a).$$

  Thus both $\underline{I}_a^b h = \overline{I}_a^b h = c(b-a)$, proving the lemma.

- **Fundamental Theorem of Calculus:**
  1. Suppose $f : [a,b] \to \mathbb{R}$ is an integrable function. Then the function $F : [a,b] \to \mathbb{R}$ defined by

$$F(x) = \int_a^x f(y)dy$$

     is continuous. Furthermore

$$(17) \qquad F'(x) = f(x) \quad \text{for } x \in [a,b] \text{ where } f \text{ is continuous}$$

     (if $x = a$ or $x = b$, use $F'_+(x)$ and $F'_-(x)$ respectively instead).

  2. Suppose $F : [a,b] \to \mathbb{R}$ is continuous and $f = F'$ exists except possibly at finitely many points $y_1, \dots, y_m$. Suppose also that $f : [a,b] \to \mathbb{R}$ is integrable (define f at the y_i's to be anything – it doesn't matter). Then

$$\int_a^b f(y)dy = F(b) - F(a).$$

Remark: Think of the derivative statements in part 1) and part 2) respectively as:

$$\frac{d}{dx}\left(\int_a^x f(y)dy\right) = f(x) \qquad \text{and} \qquad F(b) - F(a) = \int_a^b F'(y)dy$$

under the conditions stated in the theorem.

Proof: We start with 1). Let's first show that F is continuous. Fix any $x \in [a,b]$. We will show that

$$F(x) = \lim_{y \to x} F(y) \tag{18}$$

(use $y \to x^{\pm}$ if x is an endpoint). Since f is integrable, it is bounded and so there exists a constant $B > 0$ such that $-B < f < B$ everywhere. For $y \geq x$ we have that

$$-B \cdot (y - x) = \int_x^y -Bdz \leq \underbrace{\int_x^y f(z)dz}_{F(y)-F(x)} \leq \int_x^y Bdz = B \cdot (y - x)$$

By the squeeze theorem $\lim_{y \to x^+}[F(y) - F(x)] = 0$. One similarly shows $\lim_{y \to x^-}[F(y) - F(x)] = 0$ and hence (18) holds.

Next let's show (17). Fix $x \in [a,b]$ where f is continuous. Let us assume that $x \in (a,b)$ since $x = a$ and $x = b$ are handled similarly using the right- and left-hand derivatives. First consider only $h > 0$. We have that

$$\inf_{y \in [x,x+h]}\{f(y)\} = \frac{1}{h} \inf_{y \in [x,x+h]}\{f(y)\}\, h = \frac{1}{h}\int_x^{x+h} \inf_{y \in [x,x+h]}\{f(y)\}\, dz$$

$$\leq \underbrace{\frac{1}{h}\int_x^{x+h} f(z)dz}_{\frac{1}{h}(F(x+h)-F(x))}$$

$$\leq \frac{1}{h}\int_x^{x+h} \sup_{y \in [x,x+h]}\{f(y)\}\, dz = \frac{1}{h}\sup_{y \in [x,x+h]}\{f(y)\}\, h = \sup_{y \in [x,x+h]}\{f(y)\}$$

It's not hard to show (I plan to assign it as homework – it's not a hard exercise) that f being continuous at x implies

$$f(x) = \lim_{h\to 0^+} \inf_{y\in[x,x+h]}\{f(y)\} = \lim_{h\to 0^+} \sup_{y\in[x,x+h]}\{f(y)\}.$$

Hence by the previous inequality and the squeeze theorem

$$f(x) = \lim_{h\to 0^+}\frac{1}{h}\big(F(x+h) - F(x)\big) = F'(x).$$

The limit with $h \to 0^-$ is proved similarly, and hence we've proved (17).

Next let's prove 2). Let P be a partition of $[a,b]$. Let $P' = P \cup \{y_1, \dots, y_m\}$. Let's write out $P' = \{x_0, \dots, x_J\}$. Then F is differentiable over each subinterval $\left(x_j, x_{j+1}\right)$ and continuous over $\left[x_j, x_{j+1}\right]$ (since we assumed that F is continuous everywhere). So by the mean-value theorem there exists $c_j \in \left[x_j, x_{j+1}\right]$ such that

$$\frac{F\left(x_{j+1}\right) - F\left(x_j\right)}{x_{j+1} - x_j} = f\left(c_j\right)$$

$$\Longrightarrow \quad \left[F\left(x_{j+1}\right) - F\left(x_j\right)\right] = f\left(c_j\right)\left(x_{j+1} - x_j\right) \le \sup_{x\in[x_j,x_{j+1}]}\{f(x)\}\left(x_{j+1} - x_j\right).$$

Summing in j gives

$$[F(x_1) - F(x_0)] + [F(x_2) - F(x_1)] + \cdots + \left[F\left(x_J\right) - F\left(x_{J-1}\right)\right]$$
$$= \sum_{j=0}^{J-1} \sup_{x\in[x_j,x_{j+1}]}\{f(x)\}\left(x_{j+1} - x_j\right)$$
$$\Longrightarrow \quad F(b) - F(a) = S_{P'}f \le S_P f.$$

Since P was chosen arbitrarily, this shows that

$$F(b) - F(a) \le \inf_P S_P f = \int_a^b f(z)dz.$$

Using infimums, one similarly shows that

$$\int_a^b f(z)dz = \sup_P s_P f \le F(b) - F(a).$$

Hence indeed $F(b) - F(a) = \int_a^b f(z)dz$.

■

Taylor's Theorem

- Differentiation shows that a function is well approximated by a tangent line. But a parabola will approximate the function even better because it can take into account the function's curvature. A cubic polynomial gives an even better approximation, and so on. Quantitively this is given by Taylor's theorem, which we now discuss. It is of fundamental importance because it allows to numerically compute standard functions such as the exponential and trigonometric functions. First we need a standard piece of notation.

- **Notation:** Suppose $U \subseteq \mathbb{R}$ is open. For $k \geq 0$, $C^k(U)$ denotes the set of function $f : U \to \mathbb{R}$ that have k derivatives and all k derivatives are continuous (note that since differentiable implies continuous, it's sufficient to simply check that the k^{th} derivative is continuous). By convention the zeroth derivative of f is simply f itself (i.e. $f^{(0)} = f$). $C^\infty(U)$ denotes the set of functions with infinitely many derivatives.

- **Taylor's Theorem:** Suppose that $I \subseteq \mathbb{R}$ is an open interval and that $f \in C^{k+1}(I)$. Fix any point $a \in I$. Then

$$f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(a)}{j!}(x-a)^j + E_k(x) \tag{19}$$

where the "error function"

$$E_k(x) = \left[\frac{1}{k!}\int_0^1 (1-t)^k f^{(k+1)}\big(a+t(x-a)\big)dt\right](x-a)^{k+1} \tag{20}$$

$$\Longrightarrow \quad |E_k(x)| \leq \frac{\sup\limits_{y \text{ between } a \text{ and } x} \left|f^{(k+1)}(y)\right|}{(k+1)!}|x-a|^{k+1}.$$

Proof: We start with (in the second equality we use the "u-substitution" $z = a + t(x-a)$)

$$f(x) - f(a) = \int_a^x f'(z)dz = \int_0^1 f'\big(a+t(x-a)\big)(x-a)dt.$$

$$\Longrightarrow \quad f(x) = f(a) + \left[\int_0^1 f'\big(a+t(x-a)\big)dt\right](x-a)$$

This proves the $k = 1$ case. Next, notice that $-(1-t)$ is an antiderivative of 1. So integration by parts gives

$$\int_0^1 1 \cdot f'\big(a+t(x-a)\big)dt$$

$$= -(1-t)f'\big(a+t(x-a)\big)\Big|_{t=0}^{t=1} - \int_0^1 -(1-t)f''\big(a+t(x-a)\big)(x-a)dt$$

$$= f'(a) + \left[\int_0^1 (1-t)f''\big(a+t(x-a)\big)dt\right](x-a)$$

Plugging this into the equation before gives

$$f(x) = f(a) + f'(a)(x-a) + \left[\int_0^1 (1-t)f''\big(a+t(x-a)\big)dt\right](x-a)^2.$$

This proves the $k=2$ case. A similar integration by parts calculation gives

$$\int_0^1 (1-t)f''\big(a+t(x-a)\big)dt = f''(a) + \frac{1}{2}\int_0^1 (1-t)^2 f''\big(a+t(x-a)\big)dt$$

which if we plug into the previous equation gives

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \frac{1}{2}\left[\int_0^1 (1-t)^2 f'''\big(a+t(x-a)\big)dt\right](x-a)^3.$$

This proves the $k=3$ case. Continuing like this inductively will prove (19) where the error function is explicitly given by (20) (we leave it to the reader to try out the proof by induction, which is carried out as illustrated above).

Next let's prove the bound on $E_k(x)$. We have that

$$\left|\int_0^1 (1-t)^k f^{(k+1)}\big(a+t(x-a)\big)dt\right| \le \int_0^1 \left|(1-t)^k f^{(k+1)}\big(a+t(x-a)\big)\right|dt$$

$$\le \sup_{y\in[a,x]}\left|f^{(k+1)}(y)\right|\int_0^1 (1-t)^k dt = \sup_{y\in[a,x]}\left|f^{(k+1)}(y)\right|\frac{1}{k+1}.$$

Plugging this into the equation for $E_k(x)$ proves the theorem.

■

- **Homework:** Please read page 85 of *Advanced Calculus 2nd Ed* (not the exercises on top) by Gerald Folland:

https://sites.math.washington.edu/~folland/Homepage/AdvCalc24.pdf

It should be an easy read since we did everything there above. Please also read the statements of Theorem 2.58 and Theorem 2.63. You are not responsible for the proofs of the latter two theorems.

Series

- **Definition:** Suppose we have an infinite (countable) sequence of numbers $\{y_0, y_1, y_2, \dots\}$. We define the **(infinite) series** as the formal expression

$$\sum_{j=0}^{\infty} y_j = y_0 + y_1 + \cdots$$

If the limit on the right-hand side of (21) below exists, then we say that the series **converges** (or "**is convergent**"), and we define the value of the series $\sum_{j=0}^{\infty} y_j$ to be equal to that limit:

$$(21) \qquad \sum_{j=0}^{\infty} y_j = \lim_{k\to\infty} \sum_{j=0}^{k} y_j.$$

If this limit does not exist, we say that the series **diverges** (or "**is divergent**"), and we don't define a value for the series $\sum_{j=0}^{\infty} y_k$. A special case of when $\sum_{j=0}^{\infty} y_k$ diverges is when the above limit is $\pm\infty$ in which case we write $\sum_{j=0}^{\infty} y_k = \pm\infty$.The y_j's are called the **terms** of the series and $\sum_{j=0}^{k} y_j$ are called the **partial sums**.

- **Remark:** A few remarks about the above definition
  1. The series doesn't necessarily have to start at $j = 0$ but could start at any other value (e.g. $j = -3$). We simply chose $j = 0$ for illustration.
  2. The y_j's don't necessarily have to be numbers, but could also be vectors, matrices, etc.

- In the following theorem, we illustrate the power of Taylor's theorem for computing the exponential, sine, and cosine. For the proof we will assume derivative rules for these mentioned functions.

- **Theorem:**

$$e^x = \sum_{j=0}^{\infty} \frac{1}{j!} x^j, \quad \cos(x) = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j)!} x^{2j}, \quad \sin(x) = \sum_{k=0}^{\infty} \frac{(-1)^j}{(2j+1)!} x^{2j+1}.$$

Proof: We will prove the formula for e^x since the proof for the other two are essentially the same. Fix $b \geq 0$, we will prove that $e^x = \sum_{k=0}^{\infty} (1/j!) x^j$ on $[-b, b]$ from which the theorem will follow since $b \geq 0$ is chosen arbitrarily. Since the derivative of e^x is e^x and $e^0 = 1$, if we set $a = 0$ and $f(x) = e^x$ in Taylor's theorem we get that all $f^{(j)}(a) = 1$ and

$$e^x = \sum_{j=0}^{k} \frac{1}{j!} x^j + E_k(x)$$

where $E_k(x)$ is as in Taylor's Theorem. To prove the theorem, it will suffice to show that $\lim_{k\to\infty} \left| e^x - \sum_{j=0}^{k} (1/j!) x^j \right| = 0$ for $x \in [-b, b]$. Notice that this limit is equal to

$$\lim_{k\to\infty} |E_k(x)| \leq \lim_{k\to\infty} \frac{\sup_{y \text{ between } a \text{ and } x} \overbrace{|e^y|}^{|f^{(k+1)}(y)|}}{(k+1)!} x^{k+1} \leq \lim_{k\to\infty} \frac{\sup_{y\in[-b,b]} |e^y|}{(k+1)!} x^{k+1} \leq \lim_{k\to\infty} \frac{e^b}{(k+1)!} b^{k+1}$$
$$= 0$$

where in the last step we've used that $\lim_{k\to\infty} b^{k+1}/(k+1)! = 0$ which, although is not immediate, is a quick exercise to show.

■

- **Example:** An example of a famous function $f \in C^\infty(\mathbb{R})$ that is not equal to its Taylor series is

$$f = \begin{cases} e^{-1/x} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}.$$

It's on the level of a homework problem to show that all derivatives of this function at zero are zero (you'll have to use L'Hôpital's rule), and hence its Taylor series is $\sum_{k=0}^{\infty} (0/k!) x^k = 0$. But obviously this Taylor series cannot be equal to f since f is not zero for $x > 0$.

Improper integrals

- We've only defined integrals for bounded functions on bounded intervals $[a, b]$. But very often we need to define integrals for functions with asymptotes or on infinitely long intervals. Such integrals are called **improper integrals** and are defined, naturally, using limits.

- **Improper Integrals of Type I:** Suppose that $f : [a, \infty) \to \mathbb{R}$ is such that f is integrable on $[a, b]$ for all $b > a$. Then we define the improper integral

$$\int_0^\infty f(x)dx = \lim_{b\to\infty} \int_a^b f(x)dx$$

More precisely, if the limit on the right-hand side exists then we say that $\int_0^\infty f(x)dx$ **converges** (or "**is convergent**") and we set it equal to that limit. If the limit does not exist, then we say that $\int_0^\infty f(x)dx$ **diverges** (or "**is divergent**") and we don't assign it a numerical value. A special case of when $\int_0^\infty f(x)dx$ diverges is when the above limit is $\pm\infty$ in which case we write $\int_0^\infty f(x)dx = \pm\infty$.

- **Improper Integrals of Type II:** Suppose that $f : (a,b] \to \mathbb{R}$ is such that f is integrable on $[c,b]$ for all $c > a$. Then we define the improper integral

$$\int_a^b f(x)dx = \lim_{c\to a^+} \int_c^b f(x)dx$$

More precisely, if the limit on the right-hand side exists then we say that $\int_a^b f(x)dx$ **converges** (or "**is convergent**") and we set it equal to that limit. If the limit does not exist, then we say that $\int_a^b f(x)dx$ **diverges** (or "**is divergent**") and we don't assign it a numerical value. A special case of when $\int_a^b f(x)dx$ diverges is when the above limit is $\pm\infty$ in which case we write $\int_a^b f(x)dx = \pm\infty$.

- **Remark:** One can analogously define improper integrals on intervals of the form $(-\infty, a]$ and $[a,b)$. A few more straightforward generalizations:

1. Let us illustrate what to do if you need to take improper integrals at two places. Suppose $f : (a,\infty)$ is integrable on every interval $[b,c]$ where $a < b < c < \infty$. Fix an a_0 such that $a < a_0 < \infty$ and we define

$$\int_a^\infty f(x)dx = \lim_{c\to a^+} \int_c^{a_0} f(x)dx + \lim_{c\to\infty} \int_{a_0}^c f(x)dx$$

where we require both limits to exist separately. It's a quick exercise to show that this does not depend on the choice of a_0. Integrals on intervals of the form $(-\infty, a)$, (a,b), and $(-\infty,\infty)$ are defined similarly.

2. Next we illustrate what happens if you need to take an improper integral at a point in the middle of an interval. In other words, suppose $a < b < c$ and that $f : [a,c] \setminus \{b\} \to \mathbb{R}$ is integrable on any closed interval contained in $[a,c] \setminus \{b\}$. Then we define

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

In other words, we require both improper integrals $\int_a^b f(x)dx$ and $\int_b^c f(x)dx$ to exist separately. You cannot simply integrate through the bad point "b" in one go!

Inverse Function Theorem

- We go back and prove a few tail ends that we have left.
- **Note:** In the following theorem only, a, b, c, and/or d can be $\pm\infty$.
- **Inverse Function Theorem:** Suppose that $f : (a,b) \to \mathbb{R}$ is differentiable, which implies that f is in fact of the form $f : (a,b) \to (c,d)$ and is surjective. Suppose furthermore that $f' > 0$ or $f' < 0$ everywhere. Then f has an inverse $f^{-1} : (c,d) \to (a,b)$ and this inverse is differentiable with derivative:

$$(f^{-1})'(y) = \frac{1}{f'\big(f^{-1}(y)\big)}. \tag{22}$$

Proof: Let's assume that $f' > 0$ since the proof is very similar in the case $f' < 0$. Observe that $f' > 0$ implies that f is strictly increasing. We have that f maps surjectively $f : (a,b) \to$ "Interval" because it is continuous and hence maps connected sets to connected sets, and hence its image is also an interval. That "interval" can be either of the form (c,d), $[c,d)$, $(c,d]$, or $[c,d]$. We claim that it must be of the form (c,d). We show why c can't be in the image of f, the case for d is similar. If c was in the image, then there would be an $x \in (a,b)$ such that $f(x) = c$. Then if we take any $x' \in (a,b)$ smaller than x, then $f(x') < c$ since f is strictly increasing. But we just said that the image of f cannot go below c. So c indeed can't be in the image of f. As mentioned, the argument for d is similar, and so we get that f is of the form $f : (a,b) \to (c,d)$ and is surjective.

Since f is strictly increasing, we have that the inverse exists: for every $y \in (c,d)$ define $f^{-1}(y)$ as $x = f^{-1}(y)$ where x is the unique $x \in (a,b)$ such that $y = f(x)$ (i.e. x exists since f is surjective onto (c,d) and is unique since f is strictly increasing). So let us prove (22) above.

Pick any $x_0 \in (a,b)$ and let $y_0 = f(x_0) \in (c,d)$. We need to show that

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}$$

In other words, we need to show that

$$\lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{1}{f'(x_0)}$$

$$\Leftrightarrow \forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall y \in (c,d) : |y - y_0| < \delta \text{ and } y \neq y_0, \left| \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} - \frac{1}{f'(x_0)} \right| < \varepsilon.$$

Take any $\varepsilon > 0$, we will show that such a $\delta > 0$ exists. Since $f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \neq 0$, we have that

$$\frac{1}{f'(x_0)} = \lim_{x \to x_0} \frac{1}{\frac{f(x) - f(x_0)}{x - x_0}} = \lim_{x \to x_0} \frac{x - x_0}{f(x) - f(x_0)}.$$

This means that

$$\exists \delta' > 0 \quad \forall x \in (a, b) : |x - x_0| < \delta' \text{ and } x \neq x_0, \quad \left| \frac{x - x_0}{f(x) - f(x_0)} - \frac{1}{f'(x_0)} \right| < \varepsilon$$

For reasons as before, f maps the interval $(x_0 - \delta', x_0 + \delta')$ surjectively onto an interval of the form $(y_0 - \delta, y_0 + \delta)$, and we claim that this is the $\delta > 0$ that we're seeking.

Indeed, take any $y \in (c, d) : |y - y_0| < \delta$ and $y \neq y_0$ (which means $y \in (y_0 - \delta, y_0 + \delta)$). Let $x \in (x_0 - \delta', x_0 + \delta')$ be such that $y = f(x)$. Then

$$\left| \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} - \frac{1}{f'(x_0)} \right| = \left| \frac{x - x_0}{f(x) - f(x_0)} - \frac{1}{f'(x_0)} \right| < \varepsilon.$$

So we're done!

■

- **Corollary:** Suppose $I \subseteq \mathbb{R}$ is an open interval, that $f \in C^k(I)$ where $k \geq 1$ (k could be ∞), and that f' is never zero. Then the inverse f^{-1} exists and is in $C^k(I)$ as well.
- **Proof:** The fact that $f \in C^k(I)$ for $k \geq 1$ implies that f' is continuous. Hence f' never being zero on the interval I implies that either always $f' > 0$ or always $f' < 0$ (by the intermediate value theorem applied to f' on I). So let us prove that $f^{-1} \in C^k(I)$ as well. From the expression

$$(f^{-1})'(y) = \frac{1}{f'\big(f^{-1}(y)\big)}$$

we see that $(f^{-1})'$ is continuous. If $k \geq 2$, we can take the derivative of this to get

$$(f^{-1})''(y) = \frac{-1}{\big[f'\big(f^{-1}(y)\big)\big]^2} f''\big(f^{-1}(y)\big) \cdot (f^{-1})'(y)$$

from which we see that $(f^{-1})''$ is also continuous. If $k \geq 2$, we can take the derivative of this again and again and keep going until we arrive at that $(f^{-1})^{(k)}$ is continuous and hence $f^{-1} \in C^k(I)$. If $k = \infty$, this process simply never stops.

■

- As an application, we mention a rigorous construction of the logarithm and exponential function.
- **Definition:** We <u>define</u> the **natural logarithm function** as

$$\ln(x) = \int_1^x \frac{1}{t} dt$$

(an interesting exercise is to use this to prove that $\ln(ab) = \ln(a) + \ln(b)$). By the fundamental theorem of calculus, the derivative of this is

$$\frac{d}{dx}\ln(x) = \frac{1}{x}.$$

In fact, since we can take infinitely many more derivatives here, this shows that $\ln(x) \in C^\infty(\mathbb{R})$. The **exponential function** e^x is defined as the inverse of $\ln(x)$. By the previous theorem

$$\frac{d}{dx}e^x = \frac{1}{\text{derivative of}\ln(x)\ \text{ evaluated at } e^x} = \frac{1}{\frac{1}{e^x}} = e^x.$$

By the previous corollary we also get that $e^x \in C^\infty(\mathbb{R})$

- **Note:** Defining sine and cosine is harder for the following reason: how does one define angle? You need to answer this in order to make sense of $\cos\theta$ and $\sin\theta$ where θ is an angle. Typically, it's rigorously defined using arclength along the unit circle. Arclength is something you'll learn about in Math 302.

Final Notes

- Interchanging derivatives and integrals.
- Graph Taylor polynomial approximations of cosine: it's very illustrative.
- No one's a dictionary, we learn math for the techniques. Big concepts are
  - Real numbers didn't fall from the heavens as a number line, they are defined using sets of sequences of rational numbers.
  - We have precise ways of talking about limits, that precisely encode what it means to get arbitrarily close. This allows to precisely build the foundation of all of analysis. A related subject if "sup and "inf"
  - Limits and continuity have been pushed to more general settings of metric spaces, one of whose fundamental examples are function spaces which revolutionized the theory of differential equations.
  - Big concepts from topology:
    - Balls
    - Open sets
    - Boundary
    - Connectedness and sequential compactness
    - Continuous function take connected/compact to connected/compact which gives the intermediate and extreme value theorems respectively
  - Derivatives as limits of secants lines

- Derivatives and integrals are like Jekyll and Hyde, connected by the Fundamental Theorem of Calculus
- Taylor polynomials gives numerical approximations for functions, and we have ways to say how small the error is.