18.100B: Real Analysis

Jesse Yang

Fall 2020

§1 September 1, 2020

Although this is an analysis course, we will start with some of the foundations, i.e. the real numbers, and as a subset of those, the rational numbers $\mathbb{Q}$.

Definition (The Rational Numbers)

The **rational numbers** $\mathbb{Q}$ are fractions in the form $\frac{p}{q}$ where both p and q are integers, and $q \neq 0$.

Fact 1

The rational numbers have operations such that you may add and multiply rational numbers together:

$$\frac{p_1}{q_1} + \frac{p_2}{q_2} \qquad \frac{p_1}{q_1} \cdot \frac{p_2}{q_2}$$

The rational number $\dfrac{p_1}{q_1} = \dfrac{ap_1}{aq_1}$ and thus does not have a unique representation. However, there is a representation where p_1, q_1 do not have a common divisor larger than 1.

Example

Is $\sqrt{2}$ a rational number?

Proof. Suppose for the sake of contradiction that $\sqrt{2} = \frac{p}{q}$ for relatively prime integers p and q (no common divisor larger than 1). Taking the square on both sides, then

$$2 = \frac{p^2}{q^2} \implies 2q^2 = p^2$$

which implies that p must be even, letting us write $p = 2p_1$ for some integer p_1. Then, we can specify

$$2q^2 = p^2 = (2p_1)^2 = 4p_1^2$$

meaning $q^2 = 2p_1^2$, which implies that q must be even. But since p and q are both even, they have a common divisor, namely 2, leading to a contradiction. Therefore, $\sqrt{2}$ is not a rational number. $\square$

Example

$\mathbb{Q}$ is not complete, i.e. it has gaps. Let $A = \{p \mid p \in \mathbb{Q}^+ \text{ and } p^2 < 2\}$ and $B = \{p \mid p \in \mathbb{Q}^+ \text{ and } p^2 > 2\}$. This is equivalent to showing that A *contains no largest number* and B *contains no smallest number.*

Proof. Let $p \in \mathbb{Q}$ where $p > 0$. Then, since $p \neq \sqrt{2}$ ($\sqrt{2}$ is not rational), we know that either $p \in A$ or $p \in B$. The problem becomes equivalent to showing that for every $p \in A$, we can find a rational $q \in A$ such that $p < q$, and for every $p \in B$, we can find a rational $q \in B$ such that $q < p$. In particular, we let

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

so then

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

Case. If $p \in A$ then $p^2 - 2 < 0$, and our two equations show that $q > p$ and $q^2 < 2$, meaning $q \in A$ is larger than p.

Case. If $p \in B$ then $p^2 - 2 > 0$, and our two equations show that $0 < q < p$ and $q^2 > 2$, meaning $q \in B$ is less than p.

$\square$

This property indicates that the rational number system has certain gaps, which the real number system $\mathbb{R}$ must fill. To illuminate the structure of $\mathbb{R}$, we discuss the notion of *ordered sets* and *fields*.

Definition (Ordered Sets)

Let A be a set. An **order** on A is a relation, denoted by $<$, which satisfies the following properties for $a, b, c \in A$:

(i) Either $a < b$, $b < a$, or $a = b$

(ii) If $a, b, c \in A$, if $a < b$ and $b < c$, then $a < c$

An **ordered set** is a set A in which an order is defined.

Definition (Upper Bound)

Given A an ordered set and $B \subseteq A$, B is said to be **bounded from above** if $\exists\, a \in A$ such that

$$\forall\, b \in B,\ b \leq a$$

We say that such an a is an **upper bound** of B.

Definition (Lower Bound)

Given A an ordered set and $B \subseteq A$, B is said to be **bounded from below** if $\exists\, a \in A$ such that

$$\forall\, b \in B,\ a \leq b$$

We say that such an a is a **lower bound** for B.

Definition (Supremum)

Given A an ordered set and $B \subseteq A$, we say that B has a **least upper bound** if $\exists\, a \in A$ such that

$$\forall\, b \in B,\ b \leq a \text{ and if } c < a,\ \text{c is } \textit{not} \text{ an upper bound}$$

Then, a is also denoted as the **supremum** of B and $a = \sup B$

Definition (Infimum)

Given A an ordered set and $B \subseteq A$, we say that B has a **greatest lower bound** if $\exists\, a \in A$ such that

$$\forall\, b \in B,\ a \leq b \text{ and if } c > a,\ c \text{ is } \textit{not} \text{ a lower bound}$$

Then, a is also denoted as the **infimum** of B and $a = \inf B$

Fact 2 (Least-Upper-Bound Property)

Given A an ordered set, it is said to have the **least-upper-bound property** if $\forall$ nonempty $B \subseteq A$, B is bounded from above, meaning $\exists \sup B \in A$.

Fact 3 (Greatest-Lower-Bound Property)

Similarly, A an ordered set has the **greatest-lower-bound property** if $\forall$ nonempty $B \subseteq A$, B is bounded from below, meaning $\exists \inf B \in A$.

Example

Given $a, b \in \mathbb{Q}$, $a < b \iff b - a > 0$. Considering $A = \mathbb{Q}$ and $B = \{a \in \mathbb{Q} \mid a < \sqrt{2}\}$, B is bounded from above, but it doesn't have a least upper bound.

Theorem 4 (Greatest Lower Bound, Given Least Upper Bound)

Suppose A is an ordered set with the least-upper-bound property, $B \subseteq A$, B is nonempty, and B is bounded below. Let L be the set of all lower bounds of B. Then $\exists\, a = \sup L \subseteq A$ and $a = \inf B$. In particular, $\exists \inf B \in A$.

Proof. Since B is bounded below, L must be nonempty, and in particular, consists of exactly those $y \in A$ which satisfy the inequality $y \leq x$ for every $x \in B$. Thus, every $x \in B$ serves as an upper bound of L. Thus L is bounded above, and our hypothesis about A implies that L has a supremum in A; denote it as a.

If $c < a$, by the definition of supremum, c is not an upper bound of L, and therefore $c \notin B$. Thus, it follows that $a \leq x$ for every $x \in B$, and therefore $a \in L$. Furthermore, since a is an upper bound of L, if $a < b$, then $b \notin L$. Equivalently, since $a \in L$ and $b > a$ is $\notin L$, a is a lower bound of B but b is not, meaning that $a = \inf B$. $\square$

Definition (Field)

A **field** is a set $\mathbb{F}$ with two operations, "+" and "$\cdot$" For $a, b \in \mathbb{F}$,

1. $a + b \in \mathbb{F}$
2. $a \cdot b \in \mathbb{F}$

The "addition" operation has four properties:

1. $a + b = b + a$ (Commutative)
2. $(a + b) + c = a + (b + c)$ (Associative)
3. $\exists\, 0 \in \mathbb{F}$ such that $\forall\, a \in \mathbb{F}$, $0 + a = a$ (Neutral element)
4. $\forall\, a \in \mathbb{F}$, $\exists - a \in \mathbb{F}$ such that $a + (-a) = 0$ (Inverse element)

The "multiplication" operation has four properties:

1. $a \cdot b = b \cdot a$ (Commutative)
2. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ (Associative)
3. $\exists\, 1 \in \mathbb{F}$ such that $\forall\, a \in \mathbb{F}$, $1 \cdot a = a$ (Neutral element)
4. $\forall\, a \neq 0 \in \mathbb{F}$, $\exists\, \frac{1}{a} \in \mathbb{F}$ such that $a \cdot \frac{1}{a} = 1$ (Inverse element)

Finally, each field has a final axiom, the Distributive Law:

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

Proposition 5

Given a field $\mathbb{F}$ with operations "+" and "$\cdot$" $\forall\, x \in \mathbb{F}$,

$$0 \cdot x = 0$$

Proof. $0 \cdot x + 0 \cdot x = (0 + 0) \cdot x$, and letting $a = 0 \cdot x$, we have $a + a = a$. Adding $-a$ to both sides,

$$a + a + (-a) = a + (-a) \implies a = 0$$

$\square$

Proposition 6
$\forall\, x \in \mathbb{F}$,

$$(-1) \cdot x = -x$$

Proof.
$$(-1) \cdot x + x = (-1) \cdot x + 1 \cdot x = ((-1) + 1) \cdot x = 0 \cdot x = 0$$

Lemma 7 (Unique Additive Inverse)
If $x + y = 0$ then $y = -x$

Proof. Adding $-x$ to both sides of $x + y = 0$ yields $-x + (x + y) = -x$ which means that $(-x + x) + y = -x$ by the associative property, so $0 + y = -x$ and thus $y = -x$ is the unique additive inverse. □

Therefore, $(-1) \cdot x$ is the additive inverse of x so $(-1) \cdot x = -x$. □

§2 September 3, 2020

Proposition 8
Given $x \neq 0$ and $xy = xz$, then $y = z$.

Proof. We may multiply by $\frac{1}{x}$ on each side:

$$\begin{aligned}\frac{1}{x}(x \cdot y) &= \frac{1}{x}(x \cdot z) \\ (\frac{1}{x} \cdot x)y &= (\frac{1}{x} \cdot x)z \\ y &= z\end{aligned}$$

□

Proposition 9
Given $x \neq 0$

$$\frac{1}{\frac{1}{x}} = x$$

Notice that this is trivially similar and parallel to the proof that the additive inverse of an element's inverse is itself.

Proposition 10

Given $x \neq 0$ and $y \neq 0$, then $xy \neq 0$.

Proof. Suppose for the sake of contradiction, for $x, y \neq 0$, we have $xy = 0$. Multiplying both sides by $\frac{1}{x}$ and $\frac{1}{y}$ yields

$$\frac{1}{x} \cdot x \cdot y \cdot \frac{1}{y} = \frac{1}{x} \cdot 0 \cdot \frac{1}{y}$$
$$1 = 0$$

which gives a contradiction. □

Proposition 11

$$(-x)y = -xy = x(-y)$$

Proof. Consider $(-x)y + xy = (-x + x)y = 0 \cdot y = 0$, making use of the distributive property for the first step. This indicates that $(-x)y = -(xy)$. Similarly, we can prove the second half of the equality. □

Proposition 12

$$(-x)(-y) = xy$$

Proof.

$$(-x)(-y) = -(x(-y)) = -(-(xy)) = xy$$

□

Definition (Ordered Field)

An **ordered field** is an ordered set that is also a field. We also have the following two properties for elements $x, y, z \in \mathbb{F}$ the field:

1. if $x > 0$ and $y > 0$, then $xy > 0$.
2. if $y < z$ then $x + y < x + z$.

Next, we consider some properties of ordered fields.

Proposition 13

If $x > 0$, then $-x < 0$.

Proof. Considering $x > 0$, we can add $-x$ on both sides and preserve the ordering:

$$x + (-x) > 0 + (-x)$$
$$0 > -x$$

$\square$

Proposition 14

If $x > 0$ and $y < z$, then $xy < xz$.

Proof. If $y < z$, then $y + (-y) < z + (-y)$, meaning $0 < z - y$. Therefore,

$$0 < x(z - y) = xz + x(-y) = xz - xy$$
$$xy < xz$$

$\square$

Proposition 15

If $x \neq 0$, then $x^2 > 0$.

Proof. Suppose that $x > 0$. Then based on an axiom of an ordered field, x has a product with x that is also positive (> 0), namely

$$0 < x \cdot x = x^2$$

Suppose instead that $x < 0$ which implies that $-x > 0$, meaning $(-x)^2 > 0$. More generally

$$(-x) \cdot (-x) = x \cdot x = x^2$$

$\square$

Proposition 16

If $0 < x < y$, then $0 < \dfrac{1}{y} < \dfrac{1}{x}$.

Proof. Considering the inequality $0 < x < y$, we may multiply by $\frac{1}{x} \cdot \frac{1}{y}$ (assuming that this product is positive):

$$0 < x \cdot \frac{1}{x} \cdot \frac{1}{y} = 1 \cdot \frac{1}{y} = \frac{1}{y}$$

from the first inequality and

$$\frac{1}{x} \cdot \frac{1}{y} \cdot x < \frac{1}{x} \cdot \frac{1}{y} \cdot y = \frac{1}{x}$$

from the second inequality. Abusing the associative property, we have that

$$0 < \frac{1}{y} < \frac{1}{x}$$

$\square$

Fact 17

We claim without proof that there exists an ordered field $\mathbb{R}$ with the least upper bound property that contains $\mathbb{Q}$.

Proposition 18 (Archimedean Property)

Consider $x \in \mathbb{R}$ where $x > 0$ and consider $y \in \mathbb{R}$. Then, there exists $n \in \mathbb{Z}^+$ such that

$$nx > y$$

Proof. Suppose for the sake of contradiction that for all $x, y \in \mathbb{R}$, $\forall\, n \in \mathbb{Z}^+$, $nx \leq y$. Then,

$$\{nx \mid n > 0\} \text{ is a set bounded from above.}$$

let a be the least upper bound of this set. Then,

$$a - x < a$$

meaning $a - x$ cannot be an upper bound. Therefore, there exists an n such that

$$nx > a - x$$
$$nx + x > a$$
$$(n+1)x > a$$

But then $(n+1) \in \mathbb{Z}^+$ yields a contradiction to the fact that a is a least upper bound of the set. The **Archimedean property** states that nx is not bounded from above if $x > 0$. $\square$

As a direct consequence

Corollary

Given $0 < x$, we claim there is a rational number between 0 and x.

Proof. Given $0 < x$, we have $0 < \frac{1}{x}$. Then, based off the Archimedean property, there is an n such that

$$\frac{1}{x} < n \cdot 1$$

Then

$$0 < \frac{1}{x} < n$$

and using Proposition 16,

$$0 < \frac{1}{n} < x$$

. □

Corollary

More generally, for $x < y$, there exists an n such that

$$x < x + \frac{1}{n} < y$$

Proof. Because $x < y$, we must have $0 < y - x$, and thus there exists an n such that $0 < \frac{1}{n} < y - x$ based off the previous corollary. We can manipulate to obtain the result we desire. □

Proposition 19

$$\sqrt{2} \in \mathbb{R}$$

Proof. Consider the set

$$B = \{x \mid x > 0, x^2 \leq 2\}$$

Notice that $1 \in B$ because $0 < 1 < 2$. Therefore B is nonempty and bounded from above, since 2 acts as an upper bound (because $2 \cdot 2 = 4 = 2 + 2 > 2$). Therefore, B has a least upper bound, a. We claim that $a > 0$ and $a^2 = 2$.

Clearly $a > 0$ because $1 \in B$ and $a \geq 1 > 0$. To prove that $a^2 = 2$, we proceed in two steps, namely by proving that $a^2 \not< 2$ and $a^2 \not> 2$.

Case. Assume that $a^2 < 2$ and write $b = a + \frac{1}{n}$ for some integer n (chosen later on by the Archimedean property), so

$$b^2 = \left(a + \frac{1}{n}\right)^2 = a^2 + \frac{1}{n^2} + \frac{2}{n} \cdot a = a^2 + \frac{1}{n}\left(\frac{1}{n} + 2a\right) \leq a^2 + \frac{1}{n}(1 + 2a)$$

since $\frac{1}{n} + 2a \leq 1 + 2a$. Because $a^2 < 2$, we can choose n sufficiently large (by the Archimedean property) so that

$$\frac{1}{n}(1 + 2a) < 2 - a^2$$

But then $b^2 < 2$, so $b \in B$, but because we set $b = a + \frac{1}{n}$ and assumed a was an upper bound of B, this leads to a contradiction. Thus, we must have $a^2 \geq 2$

Case. Assume now that $a^2 > 2$. We can choose $b = a - \frac{1}{n}$ (for integer n chosen by the Archimedean property). Then

$$b^2 = \left(a - \frac{1}{n}\right)^2 = a^2 + \frac{1}{n^2} - \frac{2}{n} \cdot a = a^2 - \frac{1}{n}\left(2a - \frac{1}{n}\right) > a^2 - \frac{1}{n} \cdot 2a > 2$$

for n sufficiently large by the Archimedean property. But then, $b = a - \frac{1}{n} < a$ but $b^2 > 2$, so b is also an upper bound for B and $b < a$, meaning that a is not the least upper bound, leading to a contradiction. Therefore, we must have $a^2 \leq 2$.

But because $a^2 \geq 2$ and $a^2 \leq 2$, we must have $a^2 = 2$.

$\square$

§3 September 8, 2020

Definition (Extended Reals)

The **extended reals** form an ordered set that is not a field:

$$R \cup \{-\infty, \infty\}$$

where $x < \infty$ and $-\infty < x$ for all $x \in \mathbb{R}$, and by transitivity, $-\infty < \infty$.

If $x \in \mathbb{R}$, then $x + \infty = \infty$, and similarly, $x - \infty = -\infty$. For $x \in \mathbb{R}\backslash\{0\}$, $x \cdot \infty = \infty$ for $x > 0$ and $x \cdot \infty = -\infty$ for $x < 0$. Similarly, $\frac{x}{\infty} = \frac{x}{-\infty} = 0$, and $\frac{\infty}{x} = \infty$ for $x > 0$, $\frac{\infty}{x} = -\infty$ for $x < 0$.

Definition (Decimals)

For $x \in \mathbb{R}$, with $0 < x$, there exists an integer n such that $n \leq x < n+1$ from the Archimedean property. If we were to consider π, we know that $3 \leq \pi < 4$. Continuning, we have

$$3 + \frac{k}{10} \leq \pi < 3 + \frac{k+1}{10}$$

where k is the first decimal (and can be derived similarly to the Archimedean property. Continuing in this manner allows us to consider a decimal value.

Definition (Sequences)

A **sequence** of real numbers is a map from the positive integers to the reals:

$$f : \{1, 2, 3, \dots\} \to \mathbb{R}$$

often written as $f(n) = a_n$.

As examples, consider

$$a_n = (-1)^n \text{ (the alternating sequence); } a_1 = -1, a_2 = 1, a_3 = -1, \dots$$

$$a_n = \frac{1}{n}$$

$$a_n = \begin{cases} 0 & \text{if } n \text{ is even} \\ 1 & \text{if } n \text{ is odd} \end{cases}$$

Definition (Subsequences)

A **subsequence** of a sequence a_n is the sequence a_{n_k}. If we think about a sequence

$$f : \mathbb{Z}_+ \to \mathbb{R}$$

a subsequence is another sequence:

$$f \circ g : \mathbb{Z}_+ \to \mathbb{R}$$

where

$$g : \mathbb{Z}_+ \to \mathbb{Z}_+$$

is a strictly increasing map.

As an example,

$$a_n = \frac{1}{n}$$

is a sequence $\{1, \frac{1}{2}, \frac{1}{3}, \dots\}$. A subsequence $a_{n_k} = \frac{1}{2k}$ is the sequence $\{\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \dots\}$. This is a subsequence because $g(k) = 2k$ is a strictly increasing map from $\mathbb{Z}_+ \to \mathbb{Z}_+$.

On the other hand, consider $\{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{5}, \frac{1}{3}, \frac{1}{6}, \dots\}$ which is not a subsequence due to the $\frac{1}{3}$ term.

Definition (Convergence)

Suppose that a_n is a sequence of real numbers. We say that a_n **converges** to a i.e. $a_n \to a$ if for all $\varepsilon > 0$, there exists N such that

$$n \geq N \implies |a_n - a| < \varepsilon$$

For example, consider $a_n = (-1)^n$. Then there does not exist any a such that $a_n \to a$ because terms alternate between -1 and 1 and thus the sequence does not converge.

On the other hand, considering $a_n = \frac{1}{n}$, $a_n \to 0$ because given $\varepsilon > 0$, we can find by the Archimedean property N such that

$$0 < \frac{1}{N} < \varepsilon$$

and if $N \leq n$, then

$$0 < \frac{1}{n} \leq \frac{1}{N} < \varepsilon$$

so $|a_n - 0| = \left|\frac{1}{n}\right| < \varepsilon$ for $n \geq N$, meaning the sequence a_n converges to 0.

Example

Consider the sequence given by

$$\begin{aligned} a_0 &= 0 \\ a_1 &= 0.9 \\ a_2 &= 0.99 \\ a_3 &= 0.999 \\ a_n &= 0.\underbrace{99\dots9}_{n} \end{aligned}$$

Then $a_n \to 1$.

Proposition 20

Consider the basic properties of convergence:

1. If $a_n \to a$ and $b_n \to b$, and $a_n + b_n = c_n$, then $c_n \to a + b$.
2. If $a_n \to a$ and $b_n \to b$ then $a_n b_n \to ab$.
3. If $a_n \to a$ and $a \neq 0$, then $\frac{1}{a_n} \to \frac{1}{a}$.

Proof. If $a_n \to a$ and $b_n \to b$ with $c_n = a_n + b_n$ then $c_n \to a + b$. Consider

$$|a + b - c_n| = |a + b - (a_n + b_n)| = |a - a_n + b - b_n| \leq |a - a_n| + |b - b_n|$$

Knowing that $a_n \to a$ and $b_n \to b$, there exists N_1 so that

$$n \geq N_1 \implies |a_n - a| < \frac{\varepsilon}{2}$$

and likewise, there exists N_2 so that

$$n \geq N_2 \implies |b_n - b| < \frac{\varepsilon}{2}$$

for all $\varepsilon > 0$. Then consider $N = \max\{N_1, N_2\}$. Then,

$$n \geq N \implies |a - a_n| + |b - b_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for all $\varepsilon > 0$. □

Proof. If $a_n \to a$ and $b_n \to b$ then $a_n b_n \to ab$. Consider

$$\begin{aligned} |ab - a_n b_n| &= |ab - a_n b + a_n b - a_n b_n| \leq |ab - a_n b| + |a_n b - a_n b_n| \\ &= |b|\,|a - a_n| + |a_n|\,|b - b_n| \end{aligned}$$

for n sufficiently large. There exists N such that

$$n \geq N_1 \implies |a_n - a| \leq \frac{\varepsilon}{2(|b| + 1)}$$

meaning that

$$n \geq N_1 \implies |b|\,|a_n - a| < \frac{\varepsilon}{2}$$

so that

$$|a_n b_n - ab| \leq |a_n|\,|b_n - b| + \frac{\varepsilon}{2}$$

We have

$$|a_n| \leq |a - a_n| + |a| \leq \varepsilon + |a|$$

and assuming that $0 < \varepsilon < 1$,

$$|a_n| \leq 1 + |a|$$

Then,

$$|ab - a_n b_n| < (1 + |a|)\,|b_n - b| + \frac{\varepsilon}{2}$$

Using that $b_n \to b$, there exists N_2 such that $n \geq N_2$ implies

$$|b_n - b| < \frac{\varepsilon}{2(|a| + 1) + 1}$$

Therefore, for $n \geq \max(\{N_1, N_2\})$,

$$|a_n b_n - ab| < \varepsilon$$

□

Proof. If $a_n \to a$, $\frac{1}{a_n} \to \frac{1}{a}$ for $a \neq 0$. Consider

$$\left|\frac{1}{a_n} - \frac{1}{a}\right| = \left|\frac{a - a_n}{a_n a}\right| = \frac{|a - a_n|}{|a_n|\,|a|}$$

and because $a_n \to a$ with $|a_n| \to |a| \neq 0$, for n sufficiently large, $|a_n| > \frac{a}{2}$, meaning

$$\frac{|a - a_n|}{|a_n|\,|a|} \leq \frac{2\,|a - a_n|}{|a|^2}$$

Since $a_n \to a$, there exists an N such that $n \geq N$ implies $|a - a_n| < \frac{\varepsilon |a|^2}{2}$ for all ε. But then for $n \geq N$, we have

$$\frac{|a - a_n|}{|a_n|\,|a|} < \varepsilon$$

□

§4 September 10, 2020

Definition (Monotone Sequence)

At least one of the following properties must hold for a **monotone sequence**.

1. $a_n \leq a_{n+1}$ for all n
2. $a_{n+1} \leq a_n$ for all n

The first case is monotone **nondecreasing** and the second case is monotone **nonincreasing**.

Theorem 21 (Monotone Convergence Theorem)

Any bounded monotone sequence is convergent.

Proof. Assume that a_n is monotone nondecreasing (the two cases are symmetric), so $a_n \leq a_{n+1}$ for all n. Define $a = \sup a_n < \infty$, because the sequence is assumed to be bounded. Given $\varepsilon > 0$,

$$a - \varepsilon < a$$

so $a - \varepsilon$ is not an upper bound for $\{a_n\}$. But since it is not an upper bound, there exists an N such that $a_N > a - \varepsilon$. But then,

$$n \geq N \implies a \geq a_n \geq a_N > a - \varepsilon$$

because the sequence is nondecreasing, meaning

$$0 \leq a - a_n < \varepsilon \implies |a - a_n| < \varepsilon$$

□

Suppose we try to get $\sqrt{2}$ as a limit of a monotone nondecreasing sequence.

For example, we could let

$$a_1 = 1, a_2 = 1.4, a_3 = 1.41, a_4 = 1.414, \ldots$$

Example

Consider three sequences

$$a_n \leq c_n \leq b_n$$

with c_n sandwiched in between a_n and b_n. If c_n is sandwiched and both $a_n \to a$, $b_n \to a$, then $c_n \to a$.

Proof. Given $\varepsilon > 0$, $\exists\, N_1$ such that $n \geq N_1 \implies |a_n - a| < \varepsilon \implies a - \varepsilon < a_n$.
Similarly, $\exists\, N_2$ such that $n \geq N_2 \implies |b_n - a| < \varepsilon \implies b_n < a + \varepsilon$.
Then if we let $N = \max\{N_1, N_2\}$,

$$n \geq N \implies a - \varepsilon < a_n \leq c_n \leq b_n < a + \varepsilon$$

meaning

$$|c_n - a| < \varepsilon$$

□

Definition (Cauchy Sequence)
a_n is a **Cauchy sequence** if given $\varepsilon > 0$, there exists N such that

$$n, m \geq N \implies |a_n - a_m| < \varepsilon$$

Theorem
Any Cauchy sequence is convergent.

Theorem 22 (Contracting Mapping Theorem)
Given a map $T : \mathbb{R} \to \mathbb{R}$, it is a **contracting map** if there exists $0 < \gamma < 1$ such that for $x, y \in \mathbb{R}$,

$$|T(x) - T(y)| \leq \gamma\, |x - y|$$

Any contracting map has a **fixed point** i.e. x such that $T(x) = x$.

Consider the following example of a contracting map:

$$T(x) = \frac{1}{2}x + 1$$

We show that T is a contracting map. Notice that

$$|T(x) - T(y)| = \left|(\frac{1}{2}x + 1) - (\frac{1}{2}y + 1)\right| = \left|\frac{1}{2}x - \frac{1}{2}y\right| = \frac{1}{2}\,|x - y|$$

which is indeed a contracting map for $\gamma = \frac{1}{2}$. Furthermore, the fixed point is given by

$$\frac{1}{2}x + 1 = x \implies x = 2$$

More generally, consider $T : \mathbb{R} \to \mathbb{R}$ with $0 < \gamma < 1$ such that

$$|T(x) - T(y)| \le \gamma\, |x - y|$$

Consider the sequence $a_1 = T(x_0)$, $a_2 = T(T(x_0))$, $\ldots$, $a_n = T^n(x_0)$. We claim that $\{T^n(x_0)\}$ is a Cauchy sequence. Once this is established, then $a_n = T^n(x_0) \to a$ and then we show that $T(a) = a$ is a fixed point.

For $a_n = T^n(x_0)$, we consider

$$\begin{aligned} |a_{n+1} - a_n| = \left|T^{n+1}(x_0) - T^n(x_0)\right| = |T^n(T(x_0)) - T^n(x_0)| \\ = \left|T(T^{n-1}(T(x_0))) - T(T^{n-1}(x_0))\right| \le \gamma \left|T^{n-1}(T(x_0)) - T^{n-1}(x_0)\right| \\ \le \gamma^n\, |T(x_0) - x_0| \end{aligned}$$

after iterating n times. Then

$$\begin{aligned} |a_m - a_n| &= |(a_m - a_{m-1}) + (a_{m-1} - a_{m-2}) + \cdots + (a_{n+1} - a_n)| \\ &\le |(a_m - a_{m-1})| + |(a_{m-1} - a_{m-2})| + \cdots + |(a_{n+1} - a_n)| \\ &\le (\gamma^{m-1} + \gamma^{m-2} + \cdots + \gamma^n)\, |T(x_o) - x_0| \end{aligned}$$

so for T a contracting map, $T^n(x_0)$ is a Cauchy sequence, meaning $T^n(x_0) \to a$ for some a. Next, we attempt to show that a is the only fixed point of T. Consider

$$|T(a) - a| = \left|T(a) - T^{n+1}(x_0) + T^{n+1}(x_0) - T^n(x_0) + T^n(x_0) - a\right|$$

so

$$|T(a) - a| \le \left|T(a) - T^{n+1}(x_0)\right| + \left|T^{n+1}(x_0) - T^n(x_0)\right| + |T^n(x_0) - a|$$

For arbitrarily large n we can make the second two terms less than $\frac{\varepsilon}{3}$ using the fact that $T^n(x_0)$ forms a Cauchy sequence. The first term

$$\left|T(a) - T^{n+1}(x_0)\right| \le \gamma\, |a - T^n(x_0)|$$

by the contracting map, and we can make this less than $\frac{\varepsilon}{3}$ for arbitrarily large n since $T^n(x_0)$ is a Cauchy sequence. Suppose that y_1, y_2 are fixed points of T. Then

$$|T(y_1) - T(y_2)| = |y_1 - y_2|$$

but also

$$|T(y_1) - T(y_2)| \le \gamma\, |y_1 - y_2|$$

and $\gamma < 1$ so it must be true that $|y_1 - y_2| = 0$, or equivalently, $y_1 = y_2$, and thus there is only one fixed point.

§5 September 15, 2020

Theorem 23 (Monotonic Subsequence)
Every sequence contains a monotonic subsequence.

Proof. We construct a nonincreasing subsequence, if possible, by calling the mth element x_m of the sequence $\{x_n\}$ a *turn-back point* if all later elements are less than or equal to it, i.e. $x_m \geq x_n$ for all $n > m$. If there is an infinite subsequence of turn-back points, our nonincreasing subsequence is

$$x_{m_1} \geq x_{m_2} \geq x_{m_3} \geq \dots$$

However, if there are only finitely many turn-back points, suppose that x_M is the last turn-back point, i.e. any element x_n for $n > M$ is not a turn-back point. Then, for some $m > n$, $x_m > x_n$. This allows us to choose $m_1 > M+1$ such that $x_{m_1} > x_{M+1}$, $m_2 > m_1$ such that $x_{m_2} > x_{m_1}$, and so on to obtain an increasing subsequence

$$x_{M+1} < x_{m_1} < x_{m_2} < \dots$$

Thus, every sequence must contain a monotonic subsequence. □

Theorem 24 (Bolzano-Weierstrass Theorem)
Any bounded sequence has a convergent subsequence.

Proof using monotonic subsequence. By Theorem 23, every sequence contains a monotonic subsequence. In this case, that subsequence would be both monotonic and bounded, and thus convergent by the Monotone Convergence Theorem (Theorem 21). □

Proof using Cauchy Convergence Criterion. Consider a bounded sequence a_n which lies between A and B. Dividing this region up into a left half and a right half, there are infinitely many that lie in the left or the right (without loss of generality, assume they lie in the left). We let $b_1 = a_1$ and let b_2 be the next a_n that is in this half with infinite elements.

Whenever we are selecting the next element of our subsequence b_n, we halve the infinite interval and pick the next a_n which is on the side with infinite elements (there must be at least one). Notice then, that $|b_2 - b_1| < |B - A|$, $|b_3 - b_2| < |B - A|/2$, and in general:

$$|b_{n+1} - b_n| < \frac{|B - A|}{2^{n-1}}$$

From here, we can extrapolate to determine that b_n is a Cauchy sequence and thus must be convergent by the Cauchy Convergence Criterion. □

Theorem 25 (Cauchy Convergence Criterion)

In $\mathbb{R}$, a sequence a_n is a Cauchy sequence if and only if it is convergent.

Proof. Assume first that a_n is convergent. Then $a_n \to a$, and for all $\varepsilon > 0$, there exists N such that $n \geq N$ implies

$$|a_n - a| < \frac{\varepsilon}{2}$$

Then, if $n, m \geq N$,

$$|a_n - a_m| \leq |a_n - a| + |a - a_m| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Therefore, if a_n is convergent, then it is a Cauchy sequence.

Consider, instead, if a_n is a Cauchy sequence. First observe that a Cauchy sequence is bounded. Namely, there exists N such that $n, m \geq N$ implies

$$|a_n - a_m| < 1$$

so in particular,

$$|a_n - a_N| < 1$$

for $n \geq N$, so all but $a_1, \ldots, a_{N-1}$ (finitely many elements) are within 1 of a_N, meaning that the a_n must be bounded. We define a sequence $b_n = \sup a_i$ where $i \geq n$, i.e. the supremum of $\{a_i \mid i \geq n\}$, which is essentially the tail of the sequence. A supremum exists because this is a bounded set, so $b_n \in \mathbb{R}$. Then,

$$b_{n+1} = \sup_{i \geq n+1} a_i \leq \sup_{i \geq n} a_i = b_n$$

so for the sequence $b_n = \sup\limits_{i \geq n} a_i$, $b_{n+1} \leq b_n$, meaning that the sequence is monotone nonincreasing. Since the sequence is also bounded (because all a_n lie within a bounded set), by the Monotone Convergence Theorem, $b_n \to a$ for some a.

We claim that $a_n \to a$ as well. Given $\varepsilon > 0$, there exists N_1 such that for $n, m \geq N_1$,

$$|a_n - a_m| < \frac{\varepsilon}{2}$$

and for $n \geq N_2$,

$$|a - b_n| < \frac{\varepsilon}{2}$$

from which $|a - a_i| < \frac{\varepsilon}{2}$ for some $i \geq n \geq N_2$, since b_n is monotonically nonincreasing towards a ($b_n \geq a$) and *least* upper bounds a_i ($b_n \geq a_i$) for $i \geq n$, i.e. there ought to be an a_i between a and b_n. Furthermore, we can choose such an $i \geq \max\{N_1, N_2\}$ so then

$$|a - a_n| \leq |a - a_i| + |a_i - a_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

meaning our Cauchy sequence a_n ought to converge. $\square$

Proof using Bolzano-Weierstrass. We proceed in the same manner as the proof described in class, by first showing that a convergent sequence is a Cauchy sequence and then showing that a Cauchy sequence is bounded. This allows us to apply the Bolzano-Weierstrass theorem (Theorem 24) to the bounded Cauchy sequence $\{a_n\}$ to obtain a convergent subsequence $\{a_{n_k}\}$ which converges to some value a. We let $\varepsilon > 0$ and choose N so that

$$|a_n - a_m| < \frac{\varepsilon}{2}$$

for all $m, n \geq N$ (since $\{a_n\}$ is a Cauchy sequence). Then, we choose K so that

$$|a_{n_k} - a| < \frac{\varepsilon}{2}$$

for $k \geq K$ (since $\{a_{n_k}\}$ is a convergent subsequence). For any value $n \geq N$ and some $k \geq K$ with $n_k \geq N$, we set $m = n_k$ and notice

$$|a_n - L| \leq |a_n - a_{n_k}| + |a_{n_k} - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

thus indiciating that $\{a_n\}$ indeed converges.

$\square$

Definition (Series)

Considering a sequence a_n, then

$$s_n = a_1 + a_2 + \cdots + a_n$$

is a sequence obtained by summing a_i:

$$s_n = \sum_{i=1}^{n} a_i$$

is a **series** which converges if the sequence s_n converges.

An example is the geometric series, where $a_n = \gamma^n$ and

$$s_n = \sum_{i=0}^{n} a_i = 1 + \gamma + \gamma^2 + \cdots + \gamma^n$$

Then,

$$(1 - \gamma)s_n = (1 - \gamma)(1 + \gamma + \gamma^2 + \cdots + \gamma^n) = 1 - \gamma^{n+1}$$

meaning if $\gamma \neq 1$,

$$s_n = \frac{1 - \gamma^{n+1}}{1 - \gamma}$$

Then s_n for $n \to \infty$ converges to $\frac{1}{1-\gamma}$ for $|\gamma| < 1$ and diverges for $|\gamma| \geq 1$.

Definition (Absolute Convergence)

A series
$$s_n = \sum_{i=1}^{n} a_i$$
is **absolutely convergent** if
$$\sum_{i=1}^{n} |a_i|$$
is convergent.

Saying that $s_n = \sum_{i=1}^{n} a_i$ is convergent is the same as saying s_n is a Cauchy sequence, i.e. for $\varepsilon > 0$ and n, m sufficiently large, $|s_n - s_m| < \varepsilon$. Then, for $n > m$
$$s_n - s_m = a_{m+1} + \cdots + a_n$$
meaning
$$|s_n - s_m| = |a_{m+1} + \cdots + a_n| \leq |a_{m+1}| + \cdots + |a_n|$$
while
$$\tilde{s_n} = \sum_{i=1}^{n} |a_i|$$
implies that
$$|s_n - s_m| \leq |a_{m+1}| + \cdots + |a_n| = |\tilde{s_n} - \tilde{s_m}|\,.$$
Thus, the property of absolute convergence is stronger than convergence alone. As an example, the sequence
$$s_n = \sum_{i=1}^{n} \frac{(-1)^i}{i},$$
the alternating harmonic series, is convergent but *not* absolutely convergent.

§6 September 17, 2020

Test (Comparison Test)

Given two series
$$\sum_{n=1}^{\infty} a_n \text{ and } \sum_{n=1}^{\infty} b_n$$
with $0 \leq a_n \leq b_n$ for all n, if $\sum_{n=1}^{\infty} b_n$ is convergent, then $\sum_{n=1}^{\infty} a_n$ is convergent. Furthermore, if $\sum_{n=1}^{\infty} a_n$ is divergent, then $\sum_{n=1}^{\infty} b_n$ is divergent.

Given a sequence $s_n = \sum_{n=1}^{\infty} a_n$, where $a_n \geq 0$, s_n is an increasing sequence, meaning it is convergent if and only if it is bounded. In particular, for

$$s_n = \sum_{n=1}^{\infty} a_n \text{ and } \tilde{s_n} = \sum_{n=1}^{\infty} b_n$$

we must have $s_n \leq \tilde{s_n}$. Thus, $\sum_{n=1}^{\infty} a_n$ is convergent if $\sum_{n=1}^{\infty} b_n$ is convergent. Furthermore, $\sum_{n=1}^{\infty} b_n$ is divergent if $\sum_{n=1}^{\infty} a_n$ is divergent.

Observation: Notice that it is actually enough to assume that $0 \leq a_n \leq b_n$ or $1 \leq \frac{b_n}{a_n}$ $(0 \leq a_n)$ for $n \geq N$ (not necessarily all n).

Observation: It is also enough to assume that $0 \leq a_n \leq cb_n$ or $\frac{1}{c} \leq \frac{b_n}{a_n}$ for all $n \geq N$ for some constant $c > 0$.

Example

Consider

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

In this case, we consider $a_n = \frac{1}{n} 2^{-n} \leq 2^{-n} = b_n$. Then $0 \leq a_n \leq b_n$ and because

$$\sum_{n=1}^{\infty} b_n$$

converges (as a geometric series), the comparison test reveals that $\sum_{n=1}^{\infty} a_n$ is also convergent.

Definition (Harmonic Series)

The **harmonic series** is given by

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

and is divergent.

Consider

$$S_k = \sum_{i=1}^{k} \frac{1}{i}$$

Clearly $S_1 = 1$. We claim that $S_{2^n-1} \geq \frac{n}{2}$ holds for all n (and thus the series diverges). We prove this inductively.

Proof. If $n = 1$ then $2^n - 1 = 2 - 1 = 1$ and $S_1 = 1 \geq \frac{1}{2}$.

Assume for our inductive hypothesis that $S_{2^n-1} \geq \frac{n}{2}$ for n. Then,

$$S_{2^{n+1}-1} = S_{2^n-1} + \left(\frac{1}{2^n} + \frac{1}{2^n+1} + \frac{1}{2^n+2} + \cdots + \frac{1}{2^{n+1}-1}\right)$$

$$\geq \frac{n}{2} + \left(\underbrace{\frac{1}{2^n} + \frac{1}{2^n+1} + \cdots + \frac{1}{2^{n+1}-1}}_{2^n \text{ terms}}\right) \geq \frac{n}{2} + 2^n \cdot \frac{1}{2^{n+1}} = \frac{n+1}{2}$$

so, by induction, we ought to have $S_{2^n-1} \geq \frac{n}{2}$ for all n. □

Example

Consider the series

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

Then, $a_n = \dfrac{n}{n^2+1} = \dfrac{1}{n+\frac{1}{n}} \geq \dfrac{1}{2n} \geq 0$ for $n \geq 1$. Since $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2n}$ is divergent (scalar of the harmonic series), our original series must be divergent as well.

Example

The series

$$\sum_{n=1}^{\infty} n^k \cdot 2^{-n}$$

is convergent for all k.

Test (Ratio Test)

Given a series

$$\sum_{n=1}^{\infty} a_n$$

where $a_n \neq 0$, the series converges when

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} \leq c < 1$$

and diverges when

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} \geq c > 1$$

If

$$\frac{|a_{n+1}|}{|a_n|} \leq c < 1$$

for all n, then

$$|a_{n+1}| \leq c\,|a_n|$$

for all n, meaning

$$|a_2| \leq c\,|a_1| \text{ and } |a_3| \leq c\,|a_2| \leq c^2\,|a_1| \text{ and... } |a_{n+1}| \leq c^n\,|a_1|$$

but because $0 < c < 1$, we can think of $|a_1|$ as being a constant, meaning

$$\sum a_n \text{ is bounded above by a convergent geometric series, and is thus convergent.}$$

With similar reasoning to the Comparison Test, we need not have all values n satisfy

$$\frac{|a_{n+1}|}{|a_n|} \leq c < 1$$

but only for n sufficiently large i.e. $n \to \infty$. Symmetrically, if

$$\frac{|a_{n+1}|}{|a_n|} \geq c > 1$$

for n sufficiently large, then the series diverges.

Test (Alternating Series Test)

Given an alternating series

$$\sum_{n=1}^{\infty} a_n$$

i.e. $a_{n+1}a_n < 0$, with $|a_{n+1}| \leq |a_n|$ and $a_n \to 0$, then the series is convergent.

As a basic example, consider the alternating series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

where the sign of a_n is alternating. Then

$$|a_{n+1}| = \frac{1}{n+1} \leq \frac{1}{n} = |a_n|$$

with

$$a_n = \frac{(-1)^n}{n} \to 0$$

and thus the alternating series is convergent (although it is not absolutely convergent).

§7 September 22, 2020

Definition

Suppose a_n is a sequence and $b_n = \sup_{i \geq n} a_i$. Then $b_n \geq b_{n+1}$, forming a monotone nonincreasing sequence. Likewise, $c_n = \inf_{i \geq n} a_i$ satisfies $c_n \leq c_{n+1}$, forming a monotone nonincreasing sequence. Then

$$\lim_{n\to\infty} \sup a_n = \lim_{n\to\infty} b_n$$

and

$$\lim_{n\to\infty} \inf a_n = \lim_{n\to\infty} c_n$$

Test (Restatement of Ratio Test)

If

$$\lim_{n\to\infty} \sup \frac{|a_{n+1}|}{|a_n|} < 1$$

then the series converges.

The key idea here is to compare with the geometric series.

Test (Root Test)

If

$$\lim_{n\to\infty} \sup \sqrt[n]{|a_n|} < 1$$

then the series converges.

Both convergences in the root and ratio test are indicative of absolute convergence.

Example

Consider

$$\sum_{n=1}^{\infty} n2^{-n}$$

The ratio test yields

$$\frac{|a_{n+1}|}{|a_n|} = \frac{(n+1)2^{-(n+1)}}{n2^{-n}} = \frac{n+1}{n} \cdot \frac{1}{2} \to \frac{1}{2} < 1$$

so the series is convergent.

Example

Consider

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

The ratio test yields

$$\frac{|a_{n+1}|}{|a_n|} = \frac{2^{n+1}n!}{(n+1)!2^n} = \frac{2}{n+1} \to 0 < 1$$

so the series is convergent.

Definition (Power Series)

For a sequence a_n, for each fixed $x \in \mathbb{R}$,

$$\sum_{n=0}^{\infty} a_n x^n$$

is a **power series**.

Using the root test, we consider

$$\sqrt[n]{|a_n|\,|x|^n} = \sqrt[n]{|a_n|}\,|x|$$

Then, the series converges when

$$\lim_{n\to\infty} \sup |x| \sqrt[n]{|a_n|} < 1 \iff |x| < \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{|a_n|}}$$

where

$$R = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{|a_n|}}$$

gives the *radius of convergence*.

Example

Consider the power series

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$

This series converges when (disregarding absolute value)

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

which converges to 0 for all fixed x as $n \to \infty$. Therefore, the given power series is convergent for all x.

Definition (Continuous Functions)

Given a function

$$f : I \to \mathbb{R}$$

where I is an interval, f is said to be **continuous** at $x_o \in I$ if for all $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon$$

A function f is continuous if it is continuous at all points in its domain.

Fact

If $f : I \to R$ and $g : I \to R$ are continuous, then

$$h = f + g \text{ is continuous}$$
$$h = fg \text{ is continuous}$$

Given $\varepsilon > 0$, there exists $\delta_1 > 0$ such that

$$|x - x_0| < \delta_1 \implies |f(x) - f(x_0)| < \frac{\varepsilon}{2}$$

and there exists $\delta_2 > 0$ such that

$$|x - x_0| < \delta_2 \implies |g(x) - g(x_0)| < \frac{\varepsilon}{2}$$

and for $\delta = \min\{\delta_1, \delta_2\}$, then

$$|x - x_0| < \delta \implies |f(x) + g(x) - (f(x_0) + g(x_0))| \leq |f(x) - f(x_0)| + |g(x) - g(x_0)| < \varepsilon.$$

For $h = fg$, $|h(x) - h(x_0)|$ equals

$$|f(x)g(x) - f(x_0)g(x_0)| = |f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) - f(x_0)g(x_0)|$$

which is less than or equal to

$$|g(x)|\,|f(x) - f(x_0)| + |f(x_0)|\,|g(x) - g(x_0)|\,.$$

§8 September 24, 2020

Functions like $f(x) = c$ for some constant c or $f(x) = x$ are continuous. In general, all polynomials are continuous.

Theorem 26

If f and g are continuous functions $f, g : \mathbb{R} \to \mathbb{R}$. Assume that $f(x) = g(x)$ for $x \in \mathbb{Q}$. Then $f(x) = g(x)$ everywhere in $\mathbb{R}$.

Proof. We want to show that if $x_0 \in \mathbb{R}$,

$$f(x_0) = g(x_0).$$

For all x_0 and all x,

$$|f(x_0) - g(x_0)| \leq |f(x_0) - f(x)| + |f(x) - g(x)| + |g(x) - g(x_0)|$$

In particular, assume $x \in \mathbb{Q}$, meaning

$$|f(x_0) - g(x_0)| \leq |f(x_0) - f(x)| + |g(x) - g(x_0)|$$

For a given $\varepsilon > 0$, using that f, g are continuous, there exists $\delta_1, \delta_2 > 0$ such that

$$|x - x_0| < \delta_1 \implies |f(x) - f(x_0)| < \frac{\varepsilon}{2}$$

$$|x - x_0| < \delta_2 \implies |g(x) - g(x_0)| < \frac{\varepsilon}{2}$$

So if $\delta = \min\{\delta_1, \delta_2\}$ (and there must exist a rational $x \in \mathbb{Q}$ within δ of x_0), then

$$|x - x_0| < \delta \implies |f(x_0) - g(x_0)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

□

Lemma

Given

$$E(x) = \sum_{i=0}^{\infty} \frac{x^i}{i!} \text{ and } E(y) = \sum_{j=0}^{\infty} \frac{y^j}{j!}$$

then

$$E(x+y) = E(x)E(y)$$

Proof. We claim that given $\varepsilon > 0$ there exists N_0 such that

$$N \geq N_0 \implies \left| E(x+y) - \sum_{i+j \leq N} \frac{x^i}{i!} \cdot \frac{y^j}{j!} \right| < \varepsilon$$

where

$$\sum_{i+j \leq N} \frac{x^i}{i!} \cdot \frac{y^j}{j!} = \sum_{i=0}^{N} \sum_{j=0}^{N-i} \frac{x^i}{i!} \cdot \frac{y^j}{j!} = \sum_{n=0}^{N} \sum_{i=0}^{n} \frac{x^i}{i!} \cdot \frac{y^{n-i}}{(n-i)!}$$

$$\sum_{n=0}^{N} \frac{1}{n!} \sum_{i=0}^{n} \frac{n!}{i!(n-i)!} x^i y^{n-i} = \sum_{n=0}^{N} \frac{1}{n!} \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}$$

Consider the binomial formula

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}$$

so our expression is equivalent to

$$\sum_{i+j \leq N} \frac{x^i}{i!} \cdot \frac{y^j}{j!} = \sum_{n=0}^{N} \frac{(x+y)^n}{n!}$$

which is the first $N+1$ terms of the power expansion $E(x+y)$. Thus, our claim is true. We also prove that given $\varepsilon > 0$ and for N sufficiently large,

$$\left| E(x)E(y) - \sum_{i+j \leq N} \frac{x^i}{i!} \cdot \frac{y^j}{j!} \right| < \varepsilon$$

This is left as an exercise hehe oof. □

Example

Given

$$E(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

then

$$E(x+y) = E(x)E(y)$$

Assume n is a positive integer, and consider

$$E(n) = E(\underbrace{1 + \cdots + 1}_{n \text{ times}}) = E(1)^n$$

Notice that

$$E(\frac{1}{n}) \geq 0$$

and

$$\left(E(\frac{1}{n})\right)^n = E(1)$$

or equivalently

$$E(\frac{1}{n}) = \sqrt[n]{e}$$

For p, q positive integers, then

$$E(\frac{p}{q}) = \left(E(\frac{1}{q})\right)^p = e^{p/q}$$

so in general, for $x = \frac{p}{q}$ with p, q positive integers, we have

$$E(x) = e^x$$

For positive integers p,

$$1 = E(0) = E(p + (-p)) = E(p)E(-p) \implies E(-p) = e^{-p}$$

Then, for *all* integers p, q, it is true that

$$E(\frac{p}{q}) = e^{p/q}$$

$E(x)$ is continuous (to be proved later) and equals e^x for all $x \in \mathbb{Q}$, so they must be the same for all $x \in \mathbb{R}$.

§9 September 29, 2020

Proposition 27

Given continuous $f : I \to \mathbb{R}$, with a convergent sequence $x_n \to x$, then $f(x_n) \to f(x)$.

Proof. For all $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|y - x| < \delta \implies |f(y) - f(x)| < \varepsilon$$

Next, since $x_n \to x$, there exists N such that if $n \geq N$,

$$|x_n - x| < \delta$$

Hence, for $n \geq N$,

$$|f(x_n) - f(x)| < \varepsilon$$

□

As a consequence of this, we consider two theorems.

Theorem 28 (Extreme Value Theorem)
Given an interval $I = [a, b]$ $(a \leq x \leq b)$ with a continuous function $f : I \to \mathbb{R}$,

$$\sup f = \sup_{x \in I} f(x) \in \mathbb{R} \text{ and } \inf f = \inf_{x \in I} f(x) \in \mathbb{R}$$

Moreover, both are achieved, i.e. there exists $x, y \in [a, b]$ such that

$$f(x) = \sup f = \max f \text{ and } f(y) = \inf f = \min f$$

Proof. The Bolzano-Weierstrass Theorem says that any bounded sequence has a convergent subsequence. We know that

$$\sup f \in \mathbb{R} \cup \{\infty\}$$

We first show that $\sup f$ is a real number. Assume, for the sake of contradiction, that it is not. Then there exists a sequence x_n such that $f(x_n) > n$ for all n, and $x_n \in [a, b]$ so by the Bolzano-Weierstrass theorem, x_n has a subsequence x_{n_k} converging to $x \in [a, b]$. Since f is continuous,

$$f(x_{n_k}) \to f(x)$$

but $f(x_{n_k}) \to \infty$ while $f(x)$ is finite, leading to a contradiction. Therefore, $\sup f$ must be finite, i.e., $\sup f \in \mathbb{R}$. Next, we find $x \in [a, b]$ such that $f(x) = \sup f$. For all n, there exists x_n such that

$$\sup f \geq f(x_n) > \sup f - \frac{1}{n}$$

This implies the convergence

$$f(x_n) \to \sup f$$

On the other hand, by the Bolzano-Weierstrass theorem, there exists

$$x_{n_k} \to x \in [a, b]$$

and since f is continuous,

$$f(x_{n_k}) \to f(x)$$

and since convergence is unique, it must be true that

$$f(x) = \sup f.$$

The proof proceeds similarly for the inf case. $\square$

Theorem 29 (Intermediate Value Theorem)

Suppose $f : [a, b] \to \mathbb{R}$ is a continuous function. Then, if y lies between $f(a)$ and $f(b)$, there exists $x \in [a, b]$ such that

$$f(x) = y.$$

Proof. We assume that $f(a) < f(b)$. Consider some c such that

$$f(a) < c < f(b)$$

We want to find x_0 such that $f(x_0) = c$. We consider the set

$$A = \{x \mid f(z) \leq c \text{ for all } a \leq z \leq x\}$$

which is clearly non-empty. Then $a \in A \subseteq [a, b]$ so we can consider $\sup A$. Then, for $x < \sup A$, $f(x) \leq c$. We consider a sequence x_n that converges $x_n \to \sup A$ with each $x_n < \sup A$. Then, $f(x_n) \leq c$. Since $x_n \to \sup A$, then

$$f(x_n) \to f(\sup A) \leq c$$

Suppose, for the sake of contradiction, that $f(\sup A) < c$. Then we consider

$$\varepsilon = c - f(\sup A) > 0$$

meaning because f is continuous at $\sup A$, there exists some $\delta > 0$ such that

$$|x - \sup A| < \delta \implies |f(x) - f(\sup A)| < \varepsilon$$

But then $f(\sup A + \frac{\delta}{2}) < f(\sup A) + \varepsilon = c$, which contradicts that $\sup A$ is the least upper bound of A, since $\sup A + \frac{\delta}{2} \in A$. Therefore $f(\sup A) = c$ and $\sup A \in [a, b]$. $\square$

Definition (Metric Space)

A **metric space** is a set X together with a distance function:

$$d : X \times X \to [0, \infty)$$

For $x_1, x_2 \in X$, the **distance** $d(x_1, x_2)$ between x_1, x_2 has the following properties

(1) $d(x_1, x_2) \geq 0$ and $d(x_1, x_2) = 0$ if and only if $x_1 = x_2$

(2) $d(x_1, x_2) = d(x_2, x_1)$ (Symmetric)

(3) $d(x_1, x_3) \leq d(x_1, x_2) + d(x_2, x_3)$ (Triangle Inequality)

Examples of Metric Spaces

- $\mathbb{R}$ with $d(x_1, x_2) = |x_1 - x_2|$
- $\mathbb{R}^2$ with $d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$
- $\mathbb{R}^2$ with $d_B((x_1, y_1), (x_2, y_2)) = \max\{|x_1 - x_2|, |y_1 - y_2|\}$
- $C([0,1], \mathbb{R})$ (the set of continuous functions $f : [0,1] \to \mathbb{R}$) with

$$d(f, g) = \sup_{x \in [0,1]} |f(x) - g(x)|$$

§10 October 1, 2020

Definition (Convergence of a Metric Space)
Given a metric space (X, d) with $\{x_n\}$ a sequence ($x_n \in X$), x_n **converges** to x if:

$$d(x_n, x) \to 0$$

Suppose x_n converges to x and x_n also converges to y. We claim that $x = y$.

Proof. It suffices to show that $d(x, y) = 0$, which implies that $x = y$. We can choose n sufficiently large such that

$$d(x, y) \le d(x, x_n) + d(x_n, y) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for all $\varepsilon > 0$, meaning that $d(x, y) = 0$ and thus $x = y$. $\square$

Definition (Cauchy Sequence)
Given (X, d) a metric space with x_n a sequence in X, we say that $\{x_n\}$ is a **Cauchy sequence** if for all $\varepsilon > 0$, there exists N such that $n, m \ge N$ implies

$$d(x_n, x_m) < \varepsilon.$$

As a reminder, we showed that if x_n is a Cauchy sequence in $\mathbb{R}$, it converges to some x.

Definition (Cauchy Complete)
A metric space (X, d) is said to be **Cauchy complete** if any Cauchy sequence is convergent.

For example, $\mathbb{R}$ is Cauchy complete, $[a, b]$ is Cauchy complete, but $\mathbb{Q}$ is Cauchy *incomplete*, and $(0, 1)$ is Cauchy incomplete.

Definition

Given a set X with $A, B \subseteq X$,

$$A \cap B = \{x \in X \mid x \in A \text{ and } x \in B\}$$

$$A \cup B = \{x \in X \mid x \in A \text{ or } x \in B\}$$

We can similarly define these definition for multiple A_α. We can also express the complement of A in X:

$$X \backslash A = \{x \in X \mid x \notin A\}$$

Fact

$$X \backslash (\cap A_\alpha) = \cup (X \backslash A_\alpha)$$

Proof. Suppose that

$$x \in X \backslash (\cap A_\alpha)$$

This is equivalent to saying that for some α, $x \notin A_\alpha$, which is equivalent to saying that for some α, $x \in X \backslash A_\alpha$, meaning it must be in the union of the complements. □

Fact

$$X \backslash (\cup A_\alpha) = \cap (X \backslash A_\alpha)$$

Proof. Suppose that

$$x \in X \backslash (\cup A_\alpha)$$

This is equivalent to saying that x is not in any A_α; therefore it must be in the complement of each A_α. □

Definition (Open Subset)

Given a metric space (X, d) with $A \subseteq X$, A is said to be an **open** subset if for all $x \in A$, there exists $r > 0$ such that a tiny ball with radius r

$$B_r(x) = \{y \in X \mid d(x, y) < r\} \subseteq A$$

Both the empty set and X are open subsets.

Fact

Given open subsets $A_1, A_2 \subseteq X$ in the metric space (X, d),

$$A_1 \cup A_2 \text{ is also open}$$

$$A_1 \cap A_2 \text{ is also open}$$

For $x \in A_1 \cup A_2$, x is either in A_1 or A_2, so there is some ball around x that is contained in A_1 and/or A_2, meaning it must be in $A_1 \cup A_2$.

For $x \in A_1 \cap A_2$, there exists r_1, r_2 such that $B_{r_1}(x) \subseteq A_1$ and $B_{r_2}(x) \subseteq A_2$, so for $r = \min\{r_1, r_2\}$,

$$B_r(x) \subseteq B_{r_1}(x) \cap B_{r_2}(x) \subseteq A_1 \cap A_2$$

Fact

For a family of open subsets A_α, $\cup A_\alpha$ is open in general, and $\cap A_\alpha$ is open for *finitely* many A_α.

Definition (Closed Subset)

Given a metric space (X, d) with $A \subseteq X$, A is **closed** if $X \backslash A = \{y \in X \mid y \notin A\}$ is open.

Both X and the empty set are closed subsets (notice they are both open and closed).

Fact

Any ball (not including boundary, i.e., a strict inequality) is an open subset.

We can prove this with the triangle inequality.

§11 October 6, 2020

Fact

For A_i closed, $\cup A_i$ from 1 to n is closed.

$$X \backslash \cup A_i = \cap (X \backslash A_i)$$

and each term on the right is open, so our union must be closed. We claim that there is another way of defining closed sets:

Definition (Closed Subset (Additional))

Given a metric space (X, d), A is closed if and only if for all $\{x_n\}$ sequences $(x_n \in A)$ for all x_n that converge to x, the limit $x \in A$.

If A is closed and $x_n \in A$ is a sequence converging $x_n \to x$, then we show $x \in A$. If not, then there exists $x \in X \backslash A$ (open) and $x_n \to x$ where $x_n \in A$. Since $X \backslash A$ is open, there exists $r > 0$ such that $B_r(x) \subseteq X \backslash A$. On the other hand, $d(x, x_n) \to 0$, meaning for N sufficiently large, $x_n \in B_r(x)$ which is a contradiction, since $x_n \in A$.

For the converse, consider $A \subseteq X$ with $x_n \in A$ and $x_n \to x$ and $x \in A$. We want to show that A is closed, i.e. $X \backslash A$ is open. If $X \backslash A$ is not open, then there exists $x \in X \backslash A$ such that for all $r > 0$, $B_r(x)$ intersects A. We can pick $r = \frac{1}{n}$, meaning

$$x_n \in B_{\frac{1}{n}}(x) \cap A \neq \{\}$$

so

$$d(x, x_n) < \frac{1}{n} \implies x_n \in A \to x \in A$$

but then this is a contradiction, since $x \in A$ and $x \in X \backslash A$.

Definition (Compact)

Let (X, d) be a metric space. A subset $A \subseteq X$ is said to be **compact** if the following is the case:

- $A \subseteq \cup O_\alpha$ (O_α is a family of open subsets) $\implies$ finitely many of O_α cover A

As an example, $[a, b] \subseteq \mathbb{R}$ is compact. As another example, in $\mathbb{R}^2$ with Euclidean distance, $\overline{B_r(x)} = \{y \in \mathbb{R} \mid d(x, y) \leq r\}$ is compact.

Proposition 30

Suppose that (X, d) is a metric space with $x \in X$ and $r > 0$ with

$$B_r(x) = \{y \in X \mid d(x, y) < r\}$$

$$\overline{B_r(x)} = \{y \in X \mid d(x, y) \leq r\}$$

$B_r(x)$ is open (provable with the triangle inequality) and $\overline{B_r(x)}$ is closed.

Proof. Suppose $z \in X \backslash \overline{B_r(x)}$. Then,

$$d = d(z, x) > r \text{ and we can let } s = d - r.$$

Then, $B_s(z)$ lies completely outside $\overline{B_r(x)}$, so the complement of $\overline{B_r(x)}$ contains an open subset for each point in the complement. Then, we can take the union of infinite open subsets to form the complement, which must itself be open. $\square$

Definition (Bounded)

For a metric space (X, d), $A \subset X$ is said to be **bounded** if $A \subseteq B_R(x)$ for some R and some x with $0 < R < \infty$.

Proposition 31

If $A \subseteq X$ is compact, then A is closed.

Proof. Suppose for the sake of contradiction that A is not closed. Then there exists some sequence $x_n \in A$ with $x_n \to x \in X \backslash A$. Then consider

$$C_n = X \backslash \overline{B_{\frac{1}{n}}(x)}$$

where the given ball is closed. Then, C_n is open. Then,

$$\cup C_n = X \backslash \{x\}$$

so

$$A \subseteq \cup (X \backslash \overline{B_{\frac{1}{n}}(x)}) = \cup C_n$$

but then A is not compact, because any finite subset of $\cup C_n$ does not contain all elements of the sequence x_n. □

Proposition 32

Suppose (X, d) is a metric space and X is compact. Then any sequence in X has a Cauchy subsequence.

For $\varepsilon > 0$, we can take

$$X = \cup_{x \in X} B_\varepsilon(x)$$

and if X is compact, finitely many of these cover X. Thus, for one of our finite balls, there are infinitely many x_n in this ball. We can consider that there are infinite elements in the balls

$$B_{\varepsilon/2}(x_1) \supset B_{\varepsilon/4}(x_2) \supset B_{\varepsilon/8}(x_3) \supset \cdots$$

and choose elements in each to form our subsequence which much be a Cauchy subsequence.

Proposition 33

Suppose X is compact and A_α is a family of closed sets with $\cap A_\alpha = \emptyset$. Then, there is a finite set A_i such that $\cap A_i = \emptyset$.

Since A_α is closed, each $X \backslash A_\alpha$ is open, and

$$X \backslash (\cap A_\alpha) = \cup (X \backslash A_\alpha)$$

where our left hand side is simply X in this case. Then since X is compact, finitely many A_i will satisfy

$$X = \cup (X \backslash A_i)$$

where

$$\cup (X \backslash A_i) = X \backslash \cap A_i$$

so

$$\cap A_i = \emptyset$$

§12 October 8, 2020

Fact

Given a metric space (X, d) with $x \in X$ and $r > 0$, $y, z \in B_r(x) \implies d(y, z) < 2r$.

Theorem 34

Suppose (X, d) is a compact metric space and x_n is a sequence in X. Then, x_n has a convergent subsequence.

Proof. We find a Cauchy subsequence x_{n_k} and prove that this subsequence has a limit.

We can cover X by balls of radius 1, so

$$X = \cup_{x \in X} B_1(x)$$

and compactness tells us that finitely many of these balls cover X, composed of $B_1(y_i)$. Then, we pass this to a subsequence

$$x_{n_k} \in B_1(y_i)$$

for some y_i and k. Suppose that $x_n, x_m \in B_1(y)$. Then $d(x_n, x_m) < 2$. Now, consider covering X by balls of radius $\frac{1}{2}$, so

$$X = \cup_{x \in X} B_{\frac{1}{2}}(x)$$

meaning we can write X as a union of finitely many balls of radius $\frac{1}{2}$. As above, we pick $x_n \in B_1(y)$ for all n (some ball with infinite elements). We may assume that in one of our balls of radius $\frac{1}{2}$ (which cover this ball of radius 1), there are infinitely many x_ns lying in the ball, and this ball lies within a ball of radius 1. We let x_{n_2} be the first element of x_n lying in this ball of radius $\frac{1}{2}$ (and consider the remainder of are elements as being chosen from within this ball alone). We continue picking smaller balls and picking the next element in the sequence in this manner.

In general, for our subsequence, for $n, m \geq N$, we have $x_n, x_m \in B_{2^{-N+1}}(z)$ and $d(x_n, x_m) < 2^{-N+2}$. Given $\varepsilon > 0$, we can choose N so that $d(x_n, x_m) < \varepsilon$.

As the primary idea, each value in the subsequence, we take $\overline{B_{\varepsilon_k}(x_{n_k})} \supseteq \overline{B_{\varepsilon_{k+1}}(x_{n_{k+1}})}$ where $\varepsilon_k \to 0$.

Since we can find a subsequence with the property that for $n, m \geq N$, $x_n, x_m \in B_{2^{-N+1}}(z_N)$, or equivalently, $x_n, x_m \in \overline{B_{2^{-N+1}}(z_N)}$ (closed), which we denote as A_N. We consider

$$\cap_N A_N \neq \emptyset$$

and in fact consists of a single point which is the limit of our subsequence. Suppose that

$$y, z \in \cap_N A_N$$

meaning $y, z \in A_N$ so $d(y, z) \leq 2^{-N+2}$ for all, meaning we must have $d(y, z) = 0$. Therefore, there is at most one point $z \in \cap_N A_N$. If it exists, for $n \geq N$, we can consider $x_n \in A_N$ and $z \in A_N$. Then,

$$d(x_n, z) \leq 2^{-N+2}$$

for all $n \geq N$, which gives a definition of convergence of x_n to z. All the remains is to show that $\cap_N A_N$ is nonempty. Suppose, for the sake of contradiction that

$$\cap_N A_N = \emptyset$$

meaning that finitely many

$$A_1 \cap \cdots \cap A_M = \emptyset$$

However, for $n \geq M$, our subsequence x_n will lie inside each of these balls (based on how they were constructed), contradicting the fact that the intersection of our finite sets is $\emptyset$. $\square$

Lemma

Let (X, d) be a compact metric space and $A \subseteq X$ be a closed subset. Then, A is compact.

Proof. We express A as a subset of a union of open subsets:

$$A \subseteq \cup_\alpha O_\alpha$$

and $O_\alpha, X \backslash A$ forms a new family that is also open. Then,

$$X \subseteq (\cup O_\alpha) \cup (X \backslash A)$$

meaning finitely many of the open subsets cover X, meaning finitely many of the O_α cover A. $\square$

Theorem 35

The interval $[a, b] \subseteq \mathbb{R}$ consisting of x such that $a \leq x \leq b$ is compact. Furthermore $[a_1, b_1] \times [a_2, b_2] \subseteq \mathbb{R}^2$.

Corollary

If $A \subseteq \mathbb{R}^n$ is closed and bounded, then A is compact.

Definition (Continuity Between Metric Spaces)

Suppose we have two metric spaces, (X, d_X) and (Y, d_Y), with a map

$$f : X \to Y$$

We say that f is **continuous** at x_0 if for all $\varepsilon > 0$, there exists $\delta > 0$ such that

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

§13 October 15, 2020

Lemma

Given $I = [a, b] \subseteq \mathbb{R}$ with the usual distance metric in $\mathbb{R}$, I is compact.

Proof. Assume for the sake of contradiction that $I \subseteq \cup O_\alpha$ for an open family O_α and there is no finite subcovering, i.e., I is not compact. Without loss of generality, we can let $a = 0$ and $b = 1$. We have

$$[0, \frac{1}{2}] \subseteq [0, 1] \subseteq \cup O_\alpha$$

$$[\frac{1}{2}, 1] \subseteq [0, 1] \subseteq \cup O_\alpha$$

and one of these two intervals does not have a finite subcovering (otherwise taking the union of both's finite subcoverings would be a finite subcovering of the original interval). Without loss of generality, we may assume that $[0, \frac{1}{2}]$ does not have a finite subcovering. Similarly,

$$[0, \frac{1}{4}] \subseteq [0, \frac{1}{2}] \subseteq \cup O_\alpha$$

$$[\frac{1}{4}, \frac{1}{2}] \subseteq [0, \frac{1}{2}] \subseteq \cup O_\alpha$$

and we may assume that $[\frac{1}{4}, \frac{1}{2}]$ does not have a finite subcovering. Repeating this process, we determine an interval which does not have a finite subcovering, with left endpoint of the interval x_n increasing and right endpoint of the interval y_n decreasing. Furthermore,

$$y_{n+1} - x_{n+1} = \frac{1}{2}(y_n - x_n)$$

and since x_n, y_n are bounded x_n and y_n both converge to some value between 0 and 1, and $x_n \leq y_n$. However, since $y_n - x_n$ converges to 0, x_n and y_n converge to the same value. But then,

$$x_\infty \in I \subseteq \cup O_\alpha$$

meaning there exists some α' such that $x_\infty \in O_{\alpha'}$. But $O_{\alpha'}$ is open, so

$$x_\infty \in (z_1, z_2) \subset O_{\alpha'}$$

but then for sufficiently large n, $[x_n, y_n] \in O_{\alpha'}$, a single (and thus finite) subcovering, but this is a contradiction. Therefore, I must be compact. □

Lemma

For $a_i < b_i$, and $i = 1, \ldots, n$,

$$A = [a_1, b_1] \times \cdots \times [a_n, b_n] \subseteq \mathbb{R}^n$$

with the usual distance metric in $\mathbb{R}^n$ is compact.

We can proceed similarly by contradiction, dividing each interval into 2^n n-dimensional cubes, and letting one of the cube intervals not have a finite subcovering.

Theorem 36

Considering ($\mathbb{R}^n$, Euclidean metric) with $A \subseteq \mathbb{R}^n$ being closed and bounded, then A is compact.

Proof. If A is bounded, then $A \subseteq [-R, R] \times \cdots \times [-R, R]$ for R sufficiently large. This n-dimensional cube is compact, so A is a closed subset of a compact set and hence A is compact (using a Lemma from last week). □

Lemma

Given metric spaces $(X, d_X), (Y, d_Y), (Z, d_Z)$ with continuous maps $f : X \to Y$ and $g : Y \to Z$ then

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

is also continuous.

Proof. From the continuity of g, given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$d_Y(y_0, y) < \delta \implies d_Z(g(y_0), g(y)) < \varepsilon$$

where we can let $y_0 = f(x_0)$. Next using that f is continuous at x_0, there exists $\mu > 0$ such that

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

Combining these facts, for all $\varepsilon > 0$, there exists $\mu > 0$ such that

$$d_X(x, x_0) < \mu \implies d_Z(g(f(x)), g(f(x_0))) < \varepsilon.$$

□

Definition (Differentiable/Derivative)

Given a function $f : I \to \mathbb{R}$, with $x_0 \in I$, we say that f is **differentiable** at x_0 with **derivative** a if for all $\varepsilon > 0$, there exists $\delta > 0$ such that

$$0 < |x - x_0| < \delta \implies \left| \frac{f(x) - f(x_0)}{x - x_0} - a \right| < \varepsilon$$

Example

Consider the case where $f(x) = c$ for some constant c for all x.

Then, for all $x \neq x_0$,

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{c - c}{x - x_0} = 0.$$

Thus, $f(x) = c$ is differentiable at all poitns with derivative 0.

Example

Consider the case where $f(x) = x$.

Then, for all $x \neq x_0$,

$$\left| \frac{f(x) - f(x_0)}{x - x_0} \right| = \frac{x - x_0}{x - x_0} = 1$$

so $f(x) = x$ is differentiable at all points with derivative 1.

Lemma (Leibniz/Product, Quotient, Chain Rules)

Given differentiable functions f and g, their sums and products are also differentiable with derivatives

$$(f+g)' = f' + g'$$

$$(fg)' = f'g + fg'$$

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

$$(f(g(x)))' = f'(g(x)) \cdot g'(x)$$

Example

All polynomials are differentiable.

We can write

$$x^n = x \cdot x \cdots x$$

and then repeatedly apply Leibniz's rule. We can then extend this to all polynomials by combining monomials.

§14 October 27, 2020

Fact

For a function $f : I \to \mathbb{R}$ with $x_0 \in I$, we say that f is differentiable at x_0 with derivative

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

if the limit exists.

Example

Consider

$$f(x) = \begin{cases} x \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

First, notice that f is continuous at 0, as $\left|x \sin\left(\frac{1}{x}\right)\right| \leq |x|$, which approaches 0 as x does. Then,

$$\frac{f(x) - f(0)}{x - 0} = \frac{f(x)}{x} = \frac{x \sin\left(\frac{1}{x}\right)}{x} = \sin\left(\frac{1}{x}\right).$$

Then, as x approaches 0, $\sin\left(\frac{1}{x}\right)$ continues to oscillate, so the difference quotient does not have a limit, and thus $f'(0)$ does not exist.

Example

Using the fact that $x \sin\left(\frac{1}{x}\right)$ is continuous at $x = 0$, we can see that

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

is differentiable at $x = 0$.

Lemma

For $f : (a, b) \to \mathbb{R}$, if f is differentiable at $x_0 \in (a, b)$, then f is continuous at x_0.

Proof. Using our definition for differentiability, for all ε,

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| < \varepsilon$$

meaning

$$\left| f(x) - f(x_0) - f'(x_0)(x - x_0) \right| < \varepsilon \left| x - x_0 \right|$$

so

$$\left| f(x) - f(x_0) \right| - \left| f'(x_0)(x - x_0) \right| < \varepsilon \left| x - x_0 \right|$$

and hence

$$\left| f(x) - f(x_0) \right| < \left(\left| f'(x_0) \right| + \varepsilon \right) \left| x - x_0 \right|$$

meaning

$$\left| f(x) - f(x_0) \right| < \left(\left| f'(x_0) \right| + 1 \right) \left| x - x_0 \right|$$

provided $|x - x_0| < \mu$. Then, we can make the right-hand side less than ε for some $|x - x_0| < \delta$, since $|f'(x_0)| + 1$ is constant. ☐

Lemma

Assuming f and g are differentiable with $f, g : (a, b) \to \mathbb{R}$ and c is constant, then

$$(f + g)' = f' + g'$$

$$(cf)' = cf'$$

$$(fg)' = f'g + fg'$$

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

Proof. Consider

$$\frac{(f+g)(x)-(f+g)(x_0)}{x-x_0}=\frac{f(x)-f(x_0)}{x-x_0}+\frac{g(x)-g(x_0)}{x-x_0}\to f'(x_0)+g'(x_0).$$

Furthermore,

$$\frac{(cf)(x)-(cf)(x_0)}{x-x_0}=c\cdot\frac{f(x)-f(x_0)}{x-x_0}\to cf'(x_0).$$

Also,

$$\frac{(fg)(x)-(fg)(x_0)}{x-x_0}=\frac{f(x)g(x)-f(x_0)g(x)+f(x_0)g(x)-f(x_0)g(x_0)}{x-x_0}$$

$$=\frac{f(x)-f(x_0)}{x-x_0}\cdot g(x)+f(x_0)\cdot\frac{g(x)-g(x_0)}{x-x_0}\to f'(x_0)g(x_0)+f(x_0)g'(x_0).$$

□

Example

We can evaluate the derivative of

$$f(x)=a_0+a_1x+a_2x^2+a_3x^3+\cdots$$

We have

$$f'(x)=(a_0)'+(a_1x)'+(a_2x^2)'+(a_3x^3)'=a_1+2a_2x+3a_3x^2+\cdots$$

In general,

$$(x^n)'=nx^{n-1}.$$

Lemma (Chain Rule)

For differentiable functions $f:(a,b)\to(c,d)$ and $g:(c,d)\to\mathbb{R}$, then

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

Lemma

Consider $f:(a,b)\to\mathbb{R}$ and assume that f has its maximum/minimum at $x_0\in(a,b)$. If f is differentiable at x_0, then $f'(x_0)=0$.

Proof. Consider the difference quotient

$$\frac{f(x) - f(x_0)}{x - x_0}$$

We consider the case where f has its maximum at x_0. If $x > x_0$, the difference quotient is ≤ 0, and if $x < x_0$, the difference quotient is ≥ 0. Thus, if f is differentiable at x_0, the difference quotient must approach 0, meaning $f'(x_0) = 0$. The minimum case follows analogously. □

Theorem 37 (Rolle's Theorem)

For differentiable $f : [a, b] \to \mathbb{R}$ with $f(a) = f(b)$, then there exists $x_0 \in (a, b)$ such that $f'(x_0) = 0$.

Proof. Either f is constant, in which case all possible x_0 yield $f'(x_0) = 0$; otherwise f has either a maximum or minimum in (a, b), and our lemma above yields that $f'(x_0) = 0$ at this maximum/minimum. □

Theorem 38 (Mean Value Theorem)

For differentiable $f : [a, b] \to \mathbb{R}$, there exists $x_0 \in (a, b)$ such that

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$

Proof. Define a function $g : [a, b] \to \mathbb{R}$ with

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

where

$$g(a) = f(a) \text{ and } g(b) = f(a).$$

By Rolle's Theorem, there exists $x_0 \in (a, b)$ such that $g'(x_0) = 0$. In particular,

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

meaning

$$g'(x_0) = 0 = f'(x_0) - \frac{f(b) - f(a)}{b - a}$$

and therefore,

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}.$$

□

§15 October 29, 2020

Theorem 39 (L'Hospital's Law)

Given differentiable functions $f, g : [a, b] \to \mathbb{R}$, then there exists some point $c \in [a, b]$ such that
$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

Proof. The Mean Value Theorem yields that
$$\frac{f(b) - f(a)}{b - a} = f'(c) \text{ and } \frac{g(b) - g(a)}{b - a} = g'(\tilde{c})$$
for some $c, \tilde{c} \in [a, b]$. Dividing the first two equations yields
$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(\tilde{c})}$$
but we in fact want the stronger statement that we can choose $c = \tilde{c}$. Consider the function
$$h(x) = [f(b) - f(a)]g(x) - [g(b) - g(a)]f(x).$$
Simply evaluating,
$$h(a) = [f(b) - f(a)]g(a) - [g(b) - g(a)]f(a) = f(b)g(a) - g(b)f(a)$$
and similarly,
$$h(b) = [f(b) - f(a)]g(b) - [g(b) - g(a)]f(b) = g(a)f(b) - f(a)g(b).$$
Hence, $h(a) = h(b)$, and by Rolle's theorem, there exists $c \in (a, b)$ such that $h'(c) = 0$, and in particular,
$$h'(c) = [f(b) - f(a)]g'(c) - [g(b) - g(a)]f'(c)$$
and therefore,
$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$
□

Theorem 40 (L'Hospital's Rule)

Given differentiable functions $f, g : [a, b] \to \mathbb{R}$, suppose that $f(a) = g(a) = 0$ and $g'(a) \neq 0$. Then,
$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}.$$

Proof. There exists some $c \in (a, x)$ such that

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

However, the expression on the left-hand side equals

$$\frac{f(x)}{g(x)}$$

and as x approaches a, c approaches a because c is strictly between x and a, and therefore,

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

$\square$

Fact (Taylor Expansion (Simple Version))
Given a differentiable function $f : [a, b] \to \mathbb{R}$, manipulating the Mean Value Theorem yields for some $c \in (a, b)$,

$$f(b) = f(a) + f'(c)(b - a).$$

We can consider the equation for any $x \in [a, b]$ by considering the function $f; [a, x] \to \mathbb{R}$, which yields for some $a < c < x$,

$$f(x) = f(a) + f'(c)(x - a).$$

Fact (Taylor Expansion (Second Degree Polynomial))
Given a differentiable function $f : [a, b] \to \mathbb{R}$ where f' is also differentiable, the Taylor polynomial is given by

$$f(a) + f'(a)(b - a) + \frac{M}{2}(b - a)^2$$

where M is a constant to be determined.

Define a function

$$h(x) = f(x) - f(a) - f'(a)(x - a) - \frac{M}{2}(x - a)^2$$

and choose M such that

$$0 = h(b) = f(b) - f(a) - f'(a)(b - a) - \frac{M}{2}(b - a)^2.$$

We can also easily see that $h(a) = f(a) - f(a) = 0$. Applying Rolle's theorem, there exists some $c \in (a, b)$ such that $h'(c) = 0$. Furthermore,

$$h'(x) = f'(x) - f'(a) - M(x - a)$$

so therefore we have both $h'(c) = 0$ and $h'(a) = 0$. Thus there exists $\tilde{c} \in (a, c)$ such that

$$h''(\tilde{c}) = 0$$

by a second application of Rolle's Theorem. From our equation for $h'(x)$, taking the derivative yields

$$h''(x) = f''(x) - M$$

and $h''(\tilde{c}) = 0$, so $f''(\tilde{c}) = M$. Then, substituting into the equation where $h(b) = 0$, we see that

$$f(b) = f(a) + f'(a)(b - a) + \frac{f''(\tilde{c})}{2}(b - a)^2$$

and in a particularly small interval, $f''(\tilde{c}) \approx f''(a)$.

Fact (Taylor Expansion (Generalized))

Given differentiable function $f : [a, b] \to \mathbb{R}$ with existing derivatives $f', f'', \ldots, f^{(n+1)}$, then there exists $c \in (a, b)$ (from MVT) such that

$$f(b) = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(b - a)^i + \frac{f^{(n+1)}(c)}{(n + 1)!}(b - a)^{n+1}.$$

§16 November 3, 2020

Fact (Infinite Taylor Series)

It is not always true that

$$f(x) = \sum_{i=0}^{\infty} \frac{f^{(i)}(a)}{i!}(x - a)^i$$

Example

Consider $f(x) = e^x$ where $f'(x) = f(x)$ and $f(0) = 1$, and hence $f^{(k)}(0) = 1$. Then, the Taylor series is given by

$$\sum_{i=0}^{\infty} \frac{x^i}{i!}$$

The remainder term (final term from generalized Taylor Expansion) is given by

$$\frac{f^{(k)}(c)}{k!}x^k = \frac{f(c)}{k!}x^k = \frac{e^c}{k!}x^k = R_k(x).$$

Then, since $0 < c < x$,

$$|R_k(x)| \leq \frac{e^x x^k}{k!}$$

which approaches 0 as k approaches ∞. In this case, the function is equal to its infinite Taylor series.

Example

Consider the function $f : \mathbb{R} \to [0, \infty)$ where

$$f(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ e^{-\frac{1}{x^2}} & \text{for } x > 0 \end{cases}$$

It is clear that f is infinitely differentiable except at 0, but also $f^{(k)}(0) = 0$ for all k. The Taylor series is

$$\sum_{i=0}^{\infty} \frac{f^{(i)}(0)}{i!}x^i = 0$$

but $f(x) > 0$ for $x > 0$, so the function is not equal to its infinite Taylor series. Equivalently, $R_k(x) \not\to 0$ for $k \to \infty$.

Definition (Riemann Integral)

Considering a function $f : [a, b] \to \mathbb{R}$, we can determine the area "under" its curve by expressing it as less than a sum of rectangles and greater than another sum of rectangles. Each rectangle between x_i and x_{i+1} is between $m_i(x_{i+1} - x_i)$ and $M_i(x_{i+1} - x_i)$ where

$$m_i = \inf_{x \in [x_i, x_{i+1}]} f(x) \text{ and } M_i = \sup_{x \in [x_i, x_{i+1}]} f(x).$$

We can sum over intervals/rectangles making up the interval $[a, b]$. For a partition P, given by $a = x_0 < x_1 < \cdots < x_n = b$, let

$$L_P = \sum m_i(x_{i+1} - x_i) \text{ and } U_P = \sum M_i(x_{i+1} - x_i)$$

As n (the number of intervals) approaches ∞, we let

$$\underline{\int_a^b} = \lim_{n\to\infty} L_P \text{ and } \overline{\int_a^b} = \lim_{n\to\infty} U_P.$$

A function is *integrable* if

$$\underline{\int_a^b} = \overline{\int_a^b}.$$

Definition

A function $f : [a, b] \to \mathbb{R}$ is **uniformly continuous** if for all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

§17 November 5, 2020

Example

Suppose $f(x) = x^2$, where $f : \mathbb{R} \to \mathbb{R}$. Then,

$$|f(x) - f(y)| = \left|x^2 - y^2\right| = |x + y| \, |x - y| \, .$$

There does not exist $\delta > 0$ such that for all x, y within δ,

$$|x + y| \, \delta < \varepsilon.$$

For this example, f is not uniformly continuous.

Example

Consider $f(x) = \frac{1}{x}$, where $f : (0, 1] \to \mathbb{R}$. Then,

$$|f(x) - f(y)| = \left|\frac{1}{x} - \frac{1}{y}\right| = \left|\frac{y - x}{xy}\right|$$

Because of the xy in the denominator, f is not uniformly continuous.

Theorem 41

Consider $f : [a, b] \to \mathbb{R}$. If f is continuous, then f is uniformly continuous.

Proof. We argue by contradiction. If not, there exists $\varepsilon > 0$ such that for all $\delta > 0$ there exists x_δ, y_δ such that

$$|x_\delta - y_\delta| < \delta \text{ but } |f(x_\delta) - f(y_\delta)| \geq \varepsilon.$$

In particular, let $\delta = 2^{-n}$ amd $x_n, y_n \in [a, b]$ such that $|x_n - y_n| < 2^{-n}$. However, $|f(x_n) - f(y_n)| \geq \varepsilon$. There exists a subsequence $x_{n_k} \to x$ and $y_{n_k} \to y$, converging to the same value. Then,

$$|f(x_{n_k}) - f(y_{n_k})| \geq \varepsilon$$

is a contradiction for large enough n. □

Theorem 42

If $f : [a, b] \to \mathbb{R}$ is continuous, then f is integrable.

Proof. Recall

$$L(f, P) = \sum_{i=0}^{n-1} \inf_{x \in [x_i, x_{i+1}]} f(x) \cdot (x_{i+1} - x_i)$$

and

$$U(f, P) = \sum_{i=0}^{n-1} \sup_{x \in [x_i, x_{i+1}]} f(x) \cdot (x_{i+1} - x_i)$$

We want to show that for all $\varepsilon > 0$ there exists a partition such that

$$U(f, P) < L(f, P) + \varepsilon.$$

Fix $\varepsilon_0 = \frac{\varepsilon}{b-a}$, where $b - a$ is the length of the interval. Since f is uniformly continuous, there exists $\delta > 0$ such that

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon_0.$$

Let P be a partition where

$$\max(x_{i+1} - x_i) < \delta.$$

Consider the function on this interval, $f : [x_i, x_{i+1}] \to \mathbb{R}$. Then,

$$\inf_{x \in [x_i, x_{i+1}]} f(x) > \sup_{x \in [x_i, x_{i+1}]} f(x) - \varepsilon_0.$$

Equivalently,

$$\inf_{x \in [x_i, x_{i+1}]} f(x) + \varepsilon_0 > \sup_{x \in [x_i, x_{i+1}]} f(x).$$

Hence,

$$\sum_{i=0}^{n-1} \sup_{x \in [x_i, x_{i+1}]} f(x) \cdot (x_{i+1} - x_i) \leq \sum_{i=0}^{n-1} \left(\inf_{x \in [x_i, x_{i+1}]} f(x) + \varepsilon_0 \right) \cdot (x_{i+1} - x_i).$$

Therefore,

$$U(f, P) < L(f, P) + \varepsilon.$$

This means that $U(f, P) = L(f, P)$, and thus f is integrable. □

Fact

If $f : [a, b] \to \mathbb{R}$ is integrable and c is a constant, then cf is integrable with

$$\int_a^b cf = c \int_a^b f.$$

To show this, we can simply consider

$$L(cf, P) = c \cdot L(f, P) = c \cdot U(f, P) = U(cf, P).$$

Fact

If $f, g : [a, b] \to \mathbb{R}$ are both integrable, then $f + g$ is integrable with

$$\int_a^b (f + g) = \int_a^b f + \int_a^b g.$$

For a partition P, consider an interval $[x_i, x_{i+1}]$. Then,

$$\inf_{x \in [x_i, x_{i+1}]} f(x) + \inf_{x \in [x_i, x_{i+1}]} g(x) \le \inf_{x \in [x_i, x_{i+1}]} (f(x) + g(x)).$$

Similarly,

$$\sup_{x \in [x_i, x_{i+1}]} f(x) + \sup_{x \in [x_i, x_{i+1}]} g(x) \ge \sup_{x \in [x_i, x_{i+1}]} (f(x) + g(x)).$$

Consider partitions P_f and P_g such that

$$L(f, P_f) > U(f, P_f) - \frac{\varepsilon}{2} \text{ and } L(g, P_g) > U(g, P_g) - \frac{\varepsilon}{2}.$$

We can form a new partition P by taking all dividing points in P_f and P_g (so that inequalities are still satisfied) such that

$$L(f + g, P) > U(f + g, P) - \varepsilon.$$

Theorem 43 (Fundamental Theorem of Calculus)

Suppose $f : [a, b] \to \mathbb{R}$ is continuous. Suppose $x \in [a, b]$ and

$$F(x) = \int_a^x f(s)ds.$$

Then, F is differentiable on (a, b) and $F'(x) = f(x)$.

Proof. Assume first that $x > x_0$. Then,

$$F(x) = \int_a^x f = \int_a^{x_0} f + \int_{x_0}^x f.$$

So,

$$F(x) - F(x_0) = \int_{x_0}^x f.$$

Consider

$$\frac{F(x) - F(x_0)}{x - x_0} = \frac{\int_{x_0}^x f}{x - x_0}.$$

In particular,

$$\inf_{[x_0,x]} f \le f \le \sup_{[x_0,x]} f \implies \inf_{[x_0,x]} f \cdot (x - x_0) \le \int_{x_0}^x f \le \sup_{[x_0,x]} f \cdot (x - x_0).$$

Therefore,

$$\inf_{[x_0,x]} f \le \frac{F(x) - F(x_0)}{x - x_0} \le \sup_{[x_0,x]} f.$$

As $x \to x_0$, both

$$\inf_{[x_0,x]} f \to f(x_0) \text{ and } \sup_{[x_0,x]} f \to f(x_0).$$

Therefore, as $x \to x_0$,

$$\frac{F(x) - F(x_0)}{x - x_0} \to f(x_0).$$

□

Theorem 44 (Restatement of Fundamental Theorem of Calculus)
For $F : [a, b] \to \mathbb{R}$ where F is differentiable and $F' = f$ is integrable,

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

Proof. Consider, by telescoping of a partition,

$$F(b) - F(a) = \sum_{i=0}^{n-1} (F(x_{i+1}) - F(x_i)) = \sum_{i=0}^{n-1} f(c_i)(x_{i+1} - x_i).$$

Then,

$$L(f, P) \le F(b) - F(a) \le U(f, P).$$

□

§18 November 10, 2020

Lemma

Another helpful tool for computing integrals is

$$\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx$$

for $a < c < b$.

Proof. Suppose P_1 is a partition of $[a, c]$ and P_2 is a partition of $[c_b]$. Then $P = P_1 \cup P_2$ gives a partition of $[a, b]$. In particular,

$$L(f, P) = L(f, P_1) + L(f, P_2)$$

and

$$U(f, P_1) + U(f, P_2) = U(f, P).$$

Then, for all $\varepsilon > 0$,

$$U(f, P_1) < L(f, P_1) + \frac{\varepsilon}{2} \text{ and } U(f, P_2) < L(f, P_2) + \frac{\varepsilon}{2}$$

so

$$U(f, P) < L(f, P) + \varepsilon.$$

□

Definition (Arc Length)

Given a **curve** in the plane, $t \to (f(t), g(t))$ where f and g are functions, assuming f', g' exist and are continuous, the **arc length** is given by

$$\int_a^b \sqrt{(f'(t))^2 + (g'(t))^2}dt.$$

In particular, along the interval $[t_i, t_{i+1}]$, we can think about the arc length between $(f(t_i), g(t_i))$ and $(f(t_{i+1}), g(t_{i+1}))$, which can be approximated as

$$\sqrt{(f(t_{i+1}) - f(t_i))^2 + (g(t_{i+1}) - g(t_i))^2},$$

which we can write as

$$\sqrt{(f'(t_i)(t_{i+1} - t_i))^2 + (g'(t_i)(t_{i+1} - t_i))^2}$$

and simplify as

$$(t_{i+1} - t_i)\sqrt{(f'(t_i)^2 + g'(t_i)^2}.$$

Example

Consider the curve given by $t \to (f(t), g(t))$ with $f(t) = t$ and $g(t) = t^2$.

From $t = a$ to b, the arc length is given by

$$\int_a^b \sqrt{1 + 4t^2}dt.$$

Definition (Improper Integrals)

Taking some function $f : [a, \infty) \to \mathbb{R}$ (this also applies for intervals like $(-\infty, a]$) with finite value x between a and ∞ such that $f : [a, x] \to \mathbb{R}$ is integrable, if

$$\lim_{x \to \infty} \int_a^x f(y)dy \text{ exists,}$$

then we say that the improper integral exists, and in particular,

$$\int_a^\infty f(y)dy = \lim_{x \to \infty} \int_a^x f(y)dy.$$

Example

Consider $f(x) = \frac{1}{x}$ with $x \in (0, 1]$ and consider $0 < z \leq 1$.

Then, by the Fundamental Theorem of Calculus, we can evaluate

$$\int_z^1 f(x)dx = \ln(1) - \ln(z) = -\ln(z).$$

Since this value does not converge as $z \to 0$,

$$\int_0^1 f(x)dx$$

does not exist.

Example

We could also examine $f(x) = \frac{1}{x}$ on the interval $x \in [1, \infty)$.

For $1 < z < \infty$, then by the Fundamental Theorem of Calculus

$$\int_1^z f(x)dx = \int_1^z \frac{1}{x}dx = \ln(z) - \ln(1) = \ln(z)$$

which does not converge as $z \to \infty$, so

$$\int_1^\infty f(x)dx$$

does not exist.

Example

Next, consider $f(x) = \frac{1}{x^2}$ on the interval $[1, \infty)$.

For $1 < z < \infty$,

$$\int_1^z \frac{1}{x^2}dx = -\frac{1}{z} - (-\frac{1}{1}) = 1 - \frac{1}{z}$$

which converges to 1 as $z \to \infty$.

Definition (Cosine and Sine)

Starting at $(1, 0)$ on the unit circle, and traversing counterclockwise by an arc length of θ, the point on the unit circle is given by $(\cos\theta, \sin\theta)$.

We can also parameterize the curve of the unit circle with

$$s \to (s, \sqrt{1 - s^2}),$$

where $f(s) = s$ and $g(s) = \sqrt{1 - s^2}$. Then,

$$f'(s) = 1 \text{ and } g'(s) = -\frac{s}{\sqrt{1 - s^2}}.$$

Therefore,

$$(f')^2 + (g')^2 = 1 + \frac{s^2}{1 - s^2} = \frac{1}{1 - s^2}.$$

Then, the arc length can be given by

$$\int \frac{1}{\sqrt{1 - s^2}}ds.$$

Then, we can express

$$\int_{\cos\theta}^1 \frac{1}{\sqrt{1 - s^2}}ds = \theta$$

and θ in terms of $\sin\theta$ similarly. To calculate arc length from $(1, 0)$ to $(0, y)$, we can calculate

$$\int_0^y \sqrt{(f')^2 + (g')^2}ds = \int_0^y \frac{1}{\sqrt{1 - s^2}}ds.$$

In particular,

$$\int_0^{\sin\theta} \frac{1}{\sqrt{1 - s^2}}ds = \theta.$$

§19 November 12, 2020

Fact (Arcsin)

Interestingly, we can consider $\arcsin : (-1, 1) \to (-\frac{\pi}{2}, \frac{\pi}{2})$:

$$\arcsin(y) = \int_0^y \frac{1}{\sqrt{1-s^2}} ds$$

Consequently, by the Fundamental Theorem of Calculus,

$$\arcsin(y)' = \frac{1}{\sqrt{1-y^2}}.$$

We can also rewrite

$$\arcsin(y) = \frac{\pi}{2} - \int_0^{\sqrt{1-y^2}} \frac{1}{\sqrt{1-s^2}} ds$$

and use this formula to investigate what happens for $y \to \pm 1$, consequently showing that arcsin is continuous on $[-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$.

Example

Suppose a function is given by an infinite Taylor series

$$f(x) = \sum_{n=0}^{\infty} a_n x^n.$$

Is f differentiable and what is its derivative?

If the sum had been finite, you could take the derivative of every element, so we would formally *expect*

$$f'(x) = \sum_{n=1}^{\infty} a_n n x^{n-1}.$$

For example, considering

$$f(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!},$$

we would *expect*

$$f'(x) = \sum_{n=1}^{\infty} \frac{n x^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{m=0}^{\infty} \frac{x^m}{m!} = f(x).$$

Notice that this is just the exponential $f(x) = e^x$.

Definition (Pointwise Convergence)

Let $f_n : [a,b] \to \mathbb{R}$ be a sequence of functions. We say that f_n converges to f **pointwise** if for all x_0,

$$f_n(x_0) \to f(x_0).$$

Definition (Uniform Convergence)

Let $f_n : [a,b] \to \mathbb{R}$ be a sequence of functions. We say that f_n converges to f **uniformly** if

$$\max_{[a,b]} |f_n - f| \to 0$$

as $n \to \infty$.

Lemma

If f_n converges to f uniformly, then f_n converges to f pointwise.
Warning: In general, the converse is *not* the case.

As an example of the converse *not* being true, consider on the interval $[0,1)$ the constant function $f = 0$ and the function $f_n(x) = x^n$. Observe that if x_0 is fixed (where $0 \leq x_0 < 1$), clearly

$$x_0^n \to 0,$$

so $f_n \to f$ pointwise. On the other hand,

$$|f_n(x) - f(x)| = |x^n|$$

so f_n will not converge uniformly to zero, because given N, we can find an x very close to 1 such that

$$x^N > \frac{1}{2},$$

namely for $x > \sqrt[N]{\frac{1}{2}}$.

Theorem 45

Suppose $f_n \to f$ pointwise and $f_n' \to g$ uniformly. Then, if the f_n' are continuous, g is continuous. Then, f is differentiable and $f' = g$.

Consider

$$\int_a^x f_n(s)' ds + f_n(a) = f_n(x),$$

and simultaneously,

$$\int_a^x g(s)ds + f(a) = G(x).$$

We claim that $f_n(x) \to G(x)$ uniformly. Notice that by the Fundamental Theorem of Calculus, G is differentiable with $G' = g$. So, if $f_n \to G$ uniformly, then also $f_n \to G$ pointwise, but limits are unique, and since $f_n \to f$ pointwise as well, we would have $f = G$. To show that $f_n(x) \to G(x)$ uniformly, consider

$$|f_n(x) - G(x)| = \left| \int_a^x (f_n'(s) - g(s))ds + f_n(a) - f(a) \right|,$$

which, by the triangle inequality, is $\leq$

$$\int_a^x \left|f_n'(s) - g(s)\right| ds + |f_n(a) - f(a)| \leq \int_a^b \left|f_n'(s) - g(s)\right| ds + |f_n(a) - f(a)|,$$

where $x \leq b$. This quantity is $\leq$

$$(b-a) \max_{[a,b]} \left|f_n' - g\right| + |f_n(a) - f(a)|$$

and both these quantities converge to 0, so their sum does as well. Thus, $f_n \to G$ uniformly. Consequently, $f_n \to G$ pointwise, meaning $f = G$, and consequently $f' = g$.

Fact (Weierstrass M-test)

Given a function $f_n : I \to \mathbb{R}$ with $|f_n| \leq M_n$ where

$$\sum_{n=0}^{\infty} M_n < \infty, \text{ i.e., the sum is summable,}$$

then

$$\sum_{k=0}^{n} f_k \to \sum_{k=0}^{\infty} f_k$$

converges uniformly.

Proof. Consider

$$\left|\sum_{k=0}^{n} f_k - \sum_{k=0}^{m} f_k\right| = \left|\sum_{k=m+1}^{n} f_k\right| \leq \sum_{k=m+1}^{n} |f_k| \leq \sum_{k=m+1}^{n} M_k$$

where, without loss of generality, $n > m$, and we can make this quantity arbitrarily small (allowing Cauchy convergence). In $\mathbb{R}$, Cauchy convergence implies convergence.

□

Example

Consider

$$f_n = \sum_{k=0}^{n} \frac{x^k}{k!}$$

where $f_n : [-L, L] \to \mathbb{R}$. We claim that $f_n \to e^x$.

Notice

$$|f_n(x)| \leq \frac{L^n}{n!} = M_n$$

and by the Ratio Test,

$$\sum_{n=0}^{\infty} M_n < \infty,$$

so by the Weierstrass M-test,

$$\sum_{k=0}^{n} f_k \to \sum_{k=0}^{\infty} f_k$$

converges uniformly.

Lemma

Consider the summation

$$\sum_{n=0}^{\infty} a_n x^n$$

where

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lambda > 0$$

and $R = \frac{1}{\lambda}$ is the *radius of convergence*. Then

$$\sum_{k=0}^{n} a_k x^k \to \sum_{k=0}^{\infty} a_k x^k$$

converges uniformly on intervals $[-L, L]$ for positive $L < R$.

Proof. Considering

$$\frac{a_{n+1}}{a_n} \to \lambda,$$

then if $|x| \leq L < R$,

$$\left| a_{n+1} x^{n+1} \right| \to L^{n+1} \lambda^{n+1} = \frac{L^{n+1}}{R^{n+1}} \leq \gamma^{n+1}$$

for sufficiently large n, where $0 < \gamma < 1$. Since the γ^{n+1} are summable, the series

$$\sum a_n x^n$$

converges uniformly by the Weierstrass M-test. $\square$

Example (Geometric Series Convergence (using Weierstrass M-test))
We could apply this to see that

$$\sum_{k=0}^{n} x^k \to \sum_{k=0}^{\infty} x^k$$

converges uniformly when $x \in [-L, L]$ for $0 < L < 1$ (exactly when a geometric series converges), since

$$\left|x^k\right| \leq L^k.$$

§20 November 17, 2020

Example

If we consider $f_n = \sum_{k=0}^{n} a_k x^k$ which converges uniformly with $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \lambda$, then

$$f_n' = \sum_{k=0}^{n-1} (k+1) a_{k+1} x^k = \sum_{k=0}^{n-1} b_k x^k$$

where $b_k = (k+1)a_{k+1}$.

The ratio

$$\left|\frac{b_{k+1}}{b_k}\right| = \left|\frac{(k+2)a_{k+2}}{(k+1)a_{k+1}}\right| = \left|\frac{k+2}{k+1}\right| \left|\frac{a_{k+2}}{a_{k+1}}\right|$$

approaches λ if the ratio $\left|\frac{a_{k+2}}{a_{k+1}}\right|$ does. Then, on the interval $[-L, L] \subset (-R, R)$,

$$f_n \to f = \sum_{k=0}^{\infty} a_k x^k$$

uniformly,

$$f_n' \to f' = \sum_{k=1}^{\infty} k a_k x^{k-1}$$

uniformly, and similarly, we could see that

$$f_n'' \to f'' = \sum_{k=2}^{\infty} k(k-1)a_k x^{k-2}$$

uniformly.

Fact

Note that there exist functions that are infinitely many times differentiable but are not given by a power series. For example, consider

$$f(t) = \begin{cases} e^{-\frac{1}{t^2}} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

which doesn't have a power series expression near the origin.

Theorem (Contracting Mapping Theorem)

Given a metric space (X, d), then $T : X \to X$ is said to be a contracting map if there exists $0 < \lambda < 1$ such that for all $x_1, x_2 \in X$,

$$d(T(x_1), T(x_2)) \leq \lambda d(x_1, x_2).$$

If T is a contracting map, there exists at most one fixed point, i.e. $x \in X$ such that $T(x) = x$ (proved by showing any two fixed points are the same).

In particular, if (X, d) is a complete metric space and $T : X \to X$ is a contracting map, T must have a fixed point. Specifically,

$$\lim_{n\to\infty} T^n(x) = y$$

must be a fixed point.

Theorem 46

We show the existence and uniqueness for a certain ODE (ordinary differential equation). Suppose $f : \mathbb{R} \to \mathbb{R}$ is differentiable with $|f'| \leq L < \infty$. Then, consider the differentiable function $y; [a, b] \to \mathbb{R}$ satisfying

$$y'(x) = f(y(x)) \text{ and } y(a) = c.$$

The two equations dictate that

$$y(x) = c + \int_a^x f(y(s))ds.$$

For continuous functions $h \in C([a, b])$, let $T : C([a, b]) \to C([a, b])$, so

$$T(h)(x) = c + \int_a^x f(h(s))ds$$

where $T(h) \in C([a, b])$. Recall that $C([a, b])$ is a metric space, and for $h_1, h_2 \in C([a, b])$,

$$d(h_1, h_2) = \sup_{x\in[a,b]} |h_1(x) - h_2(x)| .$$

We want to show that T is a contracting map; then there would be a unique fixed point

$$T(h)(x) = c + \int_a^x f(h(s))ds = h(x).$$

Then, by the Fundamental Theorem of Calculus, it must be true that

$$h' = f(h(s)) \text{ and } h(a) = c,$$

meaning h satisfies the original differential equation. To show that T is a contracting map, assume initially that

$$(b - a)L < 1.$$

By definition,

$$T(h)(x) = c + \int_a^x f(h(s))ds$$

so

$$\begin{aligned} d(T(h_1), T(h_2)) &= \sup_{x\in[a,b]} \left| \left(c + \int_a^x f(h(s)ds \right) - \left(c + \int_a^x f(h_2(s))ds \right) \right| \\ &= \sup_{x\in[a,b]} \left| \int_a^x (f(h_1(s)) - f(h_2(s)))\, ds \right| \end{aligned}$$

In particular, by the Mean Value Theorem,

$$|(f(h_1(s)) - f(h_2(s))| = \left|f'(z)(h_1(s) - h_2(s)\right| \le L\, |h_1(s) - h_2(s)| .$$

Then,

$$\begin{aligned} d(T(h_1), T(h_2)) &= \sup_{x\in[a,b]} \left| \int_a^x (f(h_1(s)) - f(h_2(s)))\, ds \right| \\ &\le \sup_{x\in[a,b]} \int_a^x |f(h_1(s)) - f(h_2(s))|\, ds \\ &\le \int_a^b |f(h_1(s)) - f(h_2(s))|\, ds \\ &\le L \int_a^b |h_1(s) - h_2(s)|\, ds \\ &\le L(b - a)d(h_1, h_2) \end{aligned}$$

which is exactly a contracting map.

§21 November 19, 2020

Example

Considering $F(x,y) = y^2 - x$, then for $F(x,y) = 0$, we can express

$$x = y^2 \text{ and } y = \pm\sqrt{x}.$$

Definition (Partial Derivatives)

Suppose we have a function $F : O \to \mathbb{R}$ with $(0,0) \in O$, an open subset of $\mathbb{R}^2$, and $F(x,y) = 0$. If we fix y_0, and $F(x, y_0)$ is differentiable with derivative F_x and if we fix x_0, and $F(x_0, y)$ is differentiable with derivative F_y, then F has partial derivatives F_x and F_y.

Example

Suppose that y was expressable as a function of x, $y = \varphi(x)$, where $F(x, \varphi(x)) = 0$ and $\varphi(x)$ is differentiable. We could consider $G(s) = F(s, g(s)) = 0$ with derivative

$$0 = G'(s) = F_x + F_y\varphi'(s)$$

by the chain rule. Consequently, it would be true that

$$\varphi'(s) = -\frac{F_x}{F_y}$$

assuming $F_y \neq 0$.

Theorem 47 (Implicit Function Theorem)

Suppose $F : O \to \mathbb{R}$ is a function satisfying $F(x,y) = 0$, where O is an open subset of $\mathbb{R}^2$ containing the origin. Furthermore, suppose the partial derivaives F_x, F_y exist and are continuous with $F_y(0,0) \neq 0$. Then, in a neighborhood of $(0,0)$, y can be expressed as a function $\varphi(x)$ and

$$\varphi'(x) = -\frac{F_x}{F_y}.$$

Proof. Suppose that $F(0,0) = 0$ and consider an open neighborhood containing the origin with $F_y(0,0) \neq 0$. Without loss of generality, assume $F_y(0,0) > 0$. All vertical lines have at most one point where $F = 0$, and for a sufficiently small neighborhood, there is exactly one such point per line, meaning for this neighborhood, we can express y as a function of x:

$$\{(x, y) \mid F(x, y) = 0\} = \{(x, \varphi(x))\}$$

for some function φ. To show that φ is continuous it is enough to show that

$$x_n \to x \implies \varphi(x_n) \to \varphi(x).$$

It suffices to show that if

$$x_n \to x \text{ and } \varphi(x_n) \to y$$

then $y = \varphi(x)$. Then $F(x_n, \varphi(x_n)) = 0$ and $(x_n, \varphi(x_n)) \to (x, y)$ so $F(x, y) = 0$. Next, we show that φ is differentiable. Consider

$$0 = F(x + h, \varphi(x + h)) - F(x, \varphi(x)) = G(1) - G(0)$$

where we define

$$G(s) = F(x + sh, \varphi(x) + s(\varphi(x + h) - \varphi(x))$$

which has derivative

$$G' = F_x h + F_y(\varphi(x + h) - \varphi(x)).$$

By the Mean Value Theorem, there exists some $z \in (0, 1)$ such that

$$G(1) - G(0) = G'(z).$$

Then, for this intermediate point, $G'(z) = G(1)-G(0) = F(x+h, \varphi(x+h))-F(x, \varphi(x)) = 0 - 0 = 0$. Hence,

$$0 = F_x h + F_y(\varphi(x + h) - \varphi(x))$$

for $s = z$. In particular, $G'(z)$ equals

$$F_x(x+zh, \varphi(x)+z(\varphi(x+h)-\varphi(x)))h+F_y(x+zh, \varphi(x)+z(\varphi(x+h)-\varphi(x)))(\varphi(x+h)-\varphi(x))$$

We can see

$$\varphi(x + h) - \varphi(x) = -\frac{F_x h}{F_y}$$

so

$$\frac{\varphi(x + h) - \varphi(x)}{h} = -\frac{F_x}{F_y}$$

and thus as $h \to 0$, $\varphi'(x)$ converges to

$$-\frac{F_x(x, \varphi(x))}{F_y(x, \varphi(x))}.$$

□

§22 December 1, 2020

Example

Consider the partial derivatives of the function

$$f(x,y) = \sqrt{1-x^2-y^2}.$$

If we fix y_0 and let $g(x) = f(x, y_0)$, then we can take the derivative of $g(x)$, where

$$g'(x_0) = \frac{\partial f}{\partial x}(x_0, y_0) = \frac{-2x}{2\sqrt{1-x^2-y^2}} = -\frac{x}{\sqrt{1-x^2-y^2}}.$$

Likewise, if we fix x_0 and let $h(y) = f(x_0, y)$, then

$$h'(y_0) = \frac{\partial f}{\partial y}(x_0, y_0).$$

Definition (Partial Derivative)

Consider $D \subseteq \mathbb{R}^n$ with $f : D \to \mathbb{R}$. Then, fixing all variables except the ith,

$$\frac{\partial f}{\partial x_i} = f'(x_1, \dots, x_{i-1}, s, x_{i+1}, \dots, x_n).$$

Definition (Higher Order Partial Derivatives)

We can also take higher order partial derivatives, such as

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Fact (Laplace Equation)

For a function $u : D \to \mathbb{R}$,

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \cdots + \frac{\partial^2 u}{\partial x_n^2} = Lu$$

and u is said to be a harmonic function if

$$Lu = 0.$$

Definition (Directional Derivatives)

Suppose we have a function $f : \mathbb{R}^2 \to \mathbb{R}$, a point (x_0, y_0), and a direction (u, v). Considering the function

$$h(s) = f(x_0 + su, y_0 + sv),$$

the **directional derivative** at (x_0, y_0) in the direction (u, v) is given by

$$h'(0) = D_{(u,v)} f(x_0, y_0).$$

We can evaluate

$$\frac{h(s) - h(0)}{s} = \frac{f(x_0 + su, y_0 + sv) - f(x_0, y_0)}{s}.$$

We can express this as

$$\frac{f(x_0 + su, y_0 + sv) - f(x_0 + su, y_0)}{s} + \frac{f(x_0 + su, y_0) - f(x_0, y_0)}{s}.$$

If we take the limit as $s \to 0$, the first term goes to $v\frac{\partial f}{\partial y}(x_0, y_0)$ (by the Fundamental Theorem of Calculus):

$$f(x_0 + su, y_0 + sv) - f(x_0 + su, y_0) = \int_0^s v\frac{\partial f}{\partial y}(x_0 + su, y_0 tv)dt$$

and the second term goes to $u\frac{\partial f}{\partial x}(x_0, y_0)$.

Fact (Directional Derivative)

For (u, v) a unit direction vector,

$$D_{(u,v)} f(x_0, y_0) = u\frac{\partial f}{\partial x}(x_0, y_0) + v\frac{\partial f}{\partial y}(x_0, y_0).$$

Example

Evaluate the directional derivative of

$$f(x, y) = x^2 + xy + 2$$

at $(1, 1)$ in the direction $(u, v) = (\frac{1}{2}, \frac{\sqrt{3}}{2})$.

We can evaluate

$$\frac{\partial f}{\partial x} = 2x + y, \frac{\partial f}{\partial y} = x,$$

so

$$D_{(u,v)}f = \frac{1}{2} \cdot 3 + \frac{\sqrt{3}}{2} \cdot 1 = \frac{3+\sqrt{3}}{2}.$$

Fact

We can generalize the Implicit Function Theorem proof more generally for $D \subseteq \mathbb{R}^n$ with

$$(x_1, \ldots, x_{n-1}, y) = (\underline{x}, y)$$

where $F : D \to \mathbb{R}$ with $D \subseteq \mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$ and $F(\underline{x_0}, y_0) = 0$. Then, in some open neighborhood O of $(\underline{x_0}, y_0)$ in $\mathbb{R}^{n-1}$,

$$\{F(\underline{x}, y) = 0\} = \{(\underline{x}, \varphi(\underline{x})\}.$$

Assuming that all partial derivatives are defined, then

$$\frac{\partial \varphi}{\partial x_i} = -\frac{F_{x_i}}{F_y}$$

§23 December 3, 2020

Theorem 48

Suppose we have a function $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ and assume the first and second derivatives of f exist, with the second derivatives being continuous. Then,

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}.$$

Proof. Consider the expression

$$\frac{1}{k}\left[\frac{f(h,k) - f(0,k)}{h} - \frac{f(h,0) - f(0,0)}{h}\right].$$

First letting $h \to 0$, the expression will converge to

$$\frac{1}{k}\left[\frac{\partial f}{\partial x}(0,k) - \frac{\partial f}{\partial x}(0,0)\right].$$

Second, letting $k \to 0$, this expression converges to

$$\frac{\partial^2 f}{\partial y \partial x}(0,0).$$

Considering our original expression, we could rewrite as

$$\frac{1}{h}\left[\frac{f(h,k) - f(h,0)}{k} - \frac{f(0,k) - f(0,0)}{k}\right].$$

Letting $k \to 0$, this converges to

$$\frac{1}{h}\left[\frac{\partial f}{\partial y}(h,0) - \frac{\partial f}{\partial y}(0,0)\right]$$

and then letting $h \to 0$, our expression converges to

$$\frac{\partial^2 f}{\partial x \partial y}(0,0).$$

We can show that taking the limits in either order yields the same value. Considering again

$$\frac{1}{k}\left[\frac{f(h,k)-f(0,k)}{h} - \frac{f(h,0)-f(0,0)}{h}\right],$$

by the Mean Value theorem,

$$f(h,k) - f(0,k) = h\frac{\partial f}{\partial x}(z_1,k)$$

and

$$f(h,0) - f(0,0) = h\frac{\partial f}{\partial x}(z_2,0)$$

with $0 < z_1, z_2 < h$. Instead, considering as a function of h,

$$f(h,k) - f(0,k) - (f(h,0) - f(0,0)) = h\frac{\partial f}{\partial x}(z,k) - h\frac{\partial f}{\partial x}(z,0)$$

for $0 < z < h$. Thus,

$$\frac{1}{k}\left[\frac{f(h,k)-f(0,k)}{h} - \frac{f(h,0)-f(0,0)}{h}\right] = \frac{1}{kh}\left[f(h,k) - f(0,k) - (f(h,0) - f(0,0))\right]$$

which equals

$$\frac{1}{kh}\left[h\frac{\partial f}{\partial x}(z,k) - h\frac{\partial f}{\partial x}(z,0)\right] = \frac{1}{k}\left[\frac{\partial f}{\partial x}(z,k) - \frac{\partial f}{\partial x}(z,0)\right].$$

By the Mean Value Theorem, this equals

$$\frac{1}{k}\left[k\frac{\partial^2 f}{\partial y \partial x}(z,a)\right]$$

for some $0 < a < k$. Call this expression

$$\Delta = \frac{\partial^2 f}{\partial y \partial x}(z,a)$$

for some $0 < z < h$ and $0 < a < k$. We can similarly derive from the other, equal expression that

$$\Delta = \frac{\partial^2 f}{\partial x \partial y}(u,v)$$

for some $0 < u < h$ and $0 < v < k$. Then, as h and k approach 0, u, v, z, a are all evaluated at $(0,0)$. $\square$

Fact

We think about

$$\frac{\partial^2 f}{\partial x \partial y} : D \to \mathbb{R}$$

as being a continuous function of *two* variables.

Example

Consider

$$f(x,y) = \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

We show that

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = -1 \text{ and } \frac{\partial^2 f}{\partial x \partial y}(0,0) = 1.$$

First,

$$\frac{\partial f}{\partial x}(0,y) = \lim_{x\to 0} \frac{f(x,y) - f(0,y)}{x} = \lim_{x\to 0} \frac{y(x^2-y^2)}{x^2+y^2} = -y.$$

Likewise, we can compute

$$\frac{\partial f}{\partial y}(x,0) = x.$$

Then,

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{y\to 0} \frac{\frac{\partial f}{\partial x}(0,y) - \frac{\partial f}{\partial x}(0,0)}{y} = \lim_{y\to 0} \frac{-y-0}{y} = -1.$$

We can also consider

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{x\to 0} \frac{\frac{\partial f}{\partial y}(x,0) - \frac{\partial f}{\partial y}(0,0)}{x} = \lim_{x\to 0} \frac{x}{x} = 1.$$

In this case, the mixed second derivatives are not equal (the order of differentation matters). In particular, this arises due to the fact that

$$\frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}$$

are not continuous.

Fact (Differentiability in $\mathbb{R}$)
Let $o(x)$ be a function of x with the property that

$$\lim_{x \to 0} \frac{o(x)}{x} = 0.$$

A function $f : \mathbb{R} \to \mathbb{R}$ is *differentiable* at 0 with derivative c if and only if

$$f(x) = f(0) + cx + o(x).$$

This allows us to consider a generalization:

Definition (Differentiability in $\mathbb{R}^n$)
Letting $o(x)$ be a function of x with the property that for vector x,

$$\lim_{x \to 0} \frac{o(x)}{|x|} = 0.$$

A function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at 0 if

$$f(x) = c_1 x_1 + \cdots + c_n x_n + c + o(x).$$

In particular,

$$\frac{\partial f}{\partial x_i}(0, 0, \ldots, 0) = c_i.$$