Modern Analysis

Prof. Chang-Lara
Dept. of Mathematics
Columbia University
Notes by Yiqiao Yin in LaTeX

February 1, 2017

Abstract

This is the lecture notes from Prof. Chang-Lara Modern Analysis, an upper level course offered in the school year of 2016-2017 (two semesters, I & II). Similar to Calculus, Real Analysis studies the behavior of functions. In this course we mainly consider real valued functions defined over a subset of the real line. The emphasis will be to establish rigorous proofs for some of the theorems and properties you already know from your Calculus class: Sequences and series, continuity, differentiability, Riemann integral, sequences and series of functions, uniform continuity, Arzela-Ascoli Theorem and Wierstrass approximation Theorem.

This note is dedicated to Hector Chang-Lara.

Preface

When I was doing undergraduate studies in mathematics, I have not been able to fully appreciate the proofs of some of the basic knowledge encountered in the lectures. With this regret in mind, I have come to Columbia University to take the class I have missed out from my undergrad studies, Real Analysis.

I found Professor Hector Chang-Lara's lecture providing students a great variety of knowledge in the field of real analysis. Instead of following the textbooks like majority of undergraduate professors do, he implemented his own ideas and understanding about the theorems and proofs in the classroom. His lecture is, in its way, a poetry of logical ideas in the realm of real numbers. I found it difficult to give up this valuable information so I have decided to document each and every lecture he taught in this course.

Yiqiao Yin

2016 ∼ 2017 at Columbia University

Contents

1 §The Real and Complex Number Systems§

Go back to Table of Contents. Please click TOC

1.1 Introduction

Go back to Table of Contents. Please click TOC

We would see a lot of names from the 1600 to 1700 dealing with calculus problems and we would also see a lot of the names dealing with analysis in 1900. Newton and Leibniz contributions belong to the end of the XVII century. In the XVIII century we have the Bernoulli family, Euler, Lagrange, etc. The rigorous formulation of Calculus then started in the XIX century with Cauchy (also influenced by Lagrange), Abel, Riemann, and Weierstrass. We would try to answer the questions like the followings. When does the limit of continuous functions a continuous function? When does a differential equation have a unique solution? We would also try to discuss the story that Weistress came up with an example of a continuous function that is nowhere differential?

1.2 The Real Numbers: Sets, order and field axioms

Go back to Table of Contents. Please click TOC

Let us start by introducing some definitions.

Definition 1.1. A set is a collection of elements, $x \in S$: x belongs to S, and the negation is $x \notin S$.

Some famous examples are $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \emptyset$.

Definition 1.2. We use $A \subseteq B$ to denote A is a subset in B. We can also write $B \supseteq A$, that is to say, B is a superset of A (the two statements are equivalent). The opposite of this statement is that there is some element in A which is not in B.

Some examples: $A \subseteq B$ and $B \subseteq A \Rightarrow A = B$. More examples: $\emptyset \subseteq A$, and in fact we have $\emptyset \subseteq \mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$.

We are mostly dealing with real numbers. Are rational numbers enough? The answer is "no", first answered by Pythagoras. We can see from the following that rational numbers are not sufficient to describe magnitudes such as $\sqrt{2}$. Let a, b, c be the three sides of a triangle with the opposite angle of side c to be 90 degrees, then we have $a^2 + b^2 = c^2$. For example, we can choose $a = b = 1$, and compute $c = \sqrt{2}$. But $c = \sqrt{2} \notin \mathbb{R}$. This example shows us that there is a case we can a number that is not rational.

1.3 The Real Numbers: Ordered Fields and Completeness

Go back to Table of Contents. Please click TOC

Definition 1.3. $(S, <)$ is an ordered set if

(1) Trichotomy: $a < b$, $b > a$, or $a = b$

(2) Transitivity: $\forall a, b, c, \in S$, we have $a < b$ and $b < c \Rightarrow a < c$. Examples: $(\mathbb{N}, <)$, $(\mathbb{Z}, <)$, and $(\mathbb{Q}, <)$.

A field is a set where two (binary) operations are defined, namely addition and multiplication. We have the following axioms to define a field.

Definition 1.4. A field, $(S, +, \times)$, is a set where two binary operations are defined. A field, S, satisfies the following axioms [10].

Axioms of summation: $(S, +)$

(1) Closed: $\forall a, b \in S$, we have $a + b \in S$

(2) Commutative: $\forall a, b \in S$, $a + b = b + a$

(3) Associative: $\exists a, b, c \in S$, $(a + b) + c = a + (b + c)$

(4) Neutral: $\exists 0 \in S, s.t.\ \forall a \in S$, we have $a + 0 = a$

(5) Opposite: $\forall a \in S$, $\exists -a \in S$, s.t. $a + (-a) = 0$

Axioms of product: $(S, \times)$

(1) Closed: $\forall a, b \in S$, we have $a \times b \in S$

(2) Commutative: $\forall a, b \in S$, $a \times b = b \times a$ or $ab = ba$

(3) Associative: $\exists a, b, c \in S$, $(a \times b) \times c = a \times (b \times c)$ or $(ab)c = a(bc)$

(4) Unit: $\exists 1 \neq 0 \in S$, s.t. $\forall a \in S$, we have $a \times 1 = a$

(5) Inverse: $\forall a \in S|\{0\}$, $\exists \frac{1}{a} \in S$, s.t. $a \times (\frac{1}{a}) = 1$

We have Distributivity stating that $\forall a, b, c \in S$, we have $a \times (b + c) = a \times b + a \times c$. Some famous examples: $(\mathbb{Q}, +, \times)$, $(\{0, 1\}, +, \times)$.

Lemma 1.5. $0 \times a = 0$

Proof:

First, we have $0 \times a = 0 \times (0 + a)$ by neutral quality from summation. Then we have $0 \times a = 0 \times (0 + a) = 0 \times 0 + 0 \times a$ by distributivity.

Given $0 \times (0 + a) = 0 \times 0 + 0 \times a$, we use opposite to this equation and obtain $0 \times a + (-0 \times a) = 0 \times 0 + 0 \times a + (-0 \times a)$. We get $0 = 0 \times 0$.

Q.E.D.

Some examples of field are $\mathbb{Q}$ and $\mathbb{Z}_2 = \{0, 1\}$. For the set of $\mathbb{Z}_2 = \{0, 1\}$, we have addition and multiplication defined to be the following:

$$0 + 0 = 0, 0 + 1 = 1, 1 + 0 = 1, 1 + 1 = 0$$

$$0 \times 0 = 0, 0 \times 1 = 0, 1 \times 0 = 0, 1 \times 1 = 1$$

We have the field of $(\{0, 1\}, +, \times)$:

+	0	1
0	0	1
1	1	0

×	0	1
0	0	0
1	0	1

We looked at all the possibilities that we have to define the operations in $\mathbb{Z}_2$ and we conclude the operations listed above. One interesting deduction is that $0 \times 0 = 0$.

Then we can move on to discuss some of the properties we see.

First, we have cancellation, i.e. $x + y = x + z \Rightarrow y = z$. We also know for $x \neq 0$, we have $xy = xy \Leftrightarrow y = z$. We can prove this easily by subtracting x on both sides.

Secondly, we have uniqueness of inverses. That is, we have the following identity $x + y = 0 \Leftrightarrow y = -x$. For $x \neq 0$, we have $xy = 1 \Leftrightarrow y = \frac{1}{x}$.

Thirdly, We have inverse of inverse. That is, for $x \neq 0$, we have $\frac{1}{\frac{1}{x}} = x$. We also have opposite of opposite, $-(-x) = x$. Moreover, we have multiplication by zero, $\emptyset \times a = 0$. We can also set product to zero. That is, $\forall x, y \in \mathbb{R}$, $xy = \emptyset \Rightarrow x = 0 \vee y = 0$. We can also have different signs. We can have $(-x)y = -xy = x(-y)$, and $(-x)(-y) = xy$. After the understanding of definitions and basic properties, we move on to discuss some more advanced properties [10].

Proposition 1.6. *The axioms for addition imply*

(a) If $x + y = x + z$, then $y = z$.

(b) If $x + y = x$, then $y = 0$.

(c) If $x + y = 0$, then $y = -x$.

(d) If $-(-x) = x$.

Proposition 1.7. *The axioms for multiplication imply*

(a) If $x \neq 0$ and $xy = xz$, then $y = x$.

(b) If $x \neq 0$ and $xy = x$, then $y = 1$.

(c) If $x \neq 0$ and $xy = 1$, then $y = \frac{1}{x}$.

(d) If $x \neq 0$, then $1/(\frac{1}{x}) = x$.

Proposition 1.8. *The field axioms imply*

(a) $0x = 0$.

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

(c) If $(-x)y = -(xy) = x(-y)$.

(d) If $(-x)(-y) = xy$.

Remark 1.9. We can define $n \in F$, while F is a field, and $n = 1+1+...+1$. Is it possible for $1 + 1 = 0$? The answer is yes and we can observe such result in $\mathbb{Z}_2 = \{0, 1\}$.

□

We can define F to be a field of characteristic zero if $1+1+... = 1 \neq 0$, $\forall n \in \mathbb{N}$. The consequence is the following. If F has a characteristics zero, then we have $\mathbb{N} \subseteq F, \mathbb{Z} \subseteq F, \mathbb{Q} \in F$. This is where Newton's Binomial Identify comes in. We have $(a + b)^n$ and we want to see how to describe it. The idea is to write the term into a product of a series of terms, i.e. $(a+b)^n = (a+b)(a+b)...(a+b)$. For each of the n factor, for m terms, by either picking a or b. The next question we are interested is the amount of terms we have. The answer is 2^N, since this is a combination of a

and b. That is, each term of the function for m is in the form $a^r b^{n-r}$, $(r = 0, ..., n)$. For given r, how many times do we get $a^r b^{n-r}$.

With this idea in mind, another question that is equivalent to the combination question is the following: how many subsets are there in the set of $\{1, 2, ..., n\}$? The answer is $\binom{n}{r}$, i.e. among n we choose r. For example, $\binom{n}{1} = n$, $\binom{n}{n} = 1$, and $\binom{n}{0} = 1$.

In general, we know for n choose 2, we compute $\binom{n}{2} = \frac{n(n-1)}{2} = \frac{n!}{2!(n-2)!}$ and so on. Up to now, we can compute $(a+b)^n = \sum_{r=0}^{n} \binom{n}{2} a^r b^{n-r}$. For example, $(a+b)^2 = a^2 + 2ab + b^2$, and $(a+b)^3 = a^3 + 3a^2 b + 3ab^2 + b^3$. Thus, we have an illustrated understanding of the famous binomial theorem. And we need to prove the theorem to have a legit understanding of it.

Theorem 1.10. *Binomial Theorem.* $\forall a \in \mathbb{R}, b \in \mathbb{R}$, *and* $n \neq 0$, *we have*

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r}$$

Proof: We prove this theorem by mathematical induction. For the base cases, we have

$$(a+b)^0 = 1$$
$$(a+b)^1 = a + b$$
$$(a+b)^2 = a^2 + 2ab + b^2$$
$$...$$

Continuing the steps above, we can have base cases check for any arbitrary number, e.g. 0, 1, 2, and so on.

Now we assume that

$$(a+b)^k = \sum_{r=0}^{k} \binom{k}{r} a^r b^{k-r}$$

and we need to show, as an inductive conclusion, that this argument holds for $k+1$, that is, $(a+b)^{k+1} = \sum_{r=0}^{k+1} \binom{k+1}{r} a^r b^{k+1-r}$. We compute the following

$(a+b)^{k+1}$

$= (a+b)^k \times (a+b)$

$= \left(\sum_{r=0}^{k} \binom{k}{r} a^r b^{k-r} \right) \times (a+b)$ by induction hypothesis,

$= \left(\sum_{r=0}^{k} \binom{k}{r} a^r b^{k-r} \right) \times a + \left(\sum_{r=0}^{k} \binom{k}{r} a^r b^{k-r} \right) \times b$ by distributivity,

$= \left(\sum_{r=0}^{k} \binom{k}{r} a^{r+1} b^{k-r} \right) + \left(\sum_{r=0}^{k} \binom{k}{r} a^r b^{k-r+1} \right)$ by definition of power,

$= \left(\sum_{r=1}^{k+1} \binom{k}{r-1} a^r b^{k-r+1} \right) + \left(\sum_{r=0}^{k} \binom{k}{r} a^r b^{k-r+1} \right)$, since the first term can be written as $\sum_{r=0}^{k} \binom{k}{r} a^{r+1} b^{k-r} = \sum_{s=1}^{k+1} \binom{k}{s-1} a^s b^{k-s+1}$, and then replace s by r (dummy variable), and $\sum_{r=0}^{k} \binom{k}{r} a^{r+1} b^{k-r} = \sum_{r=1}^{k+1} \binom{k}{r-1} a^r b^{k-r+1}$

$= a^{k+1} + \sum_{1}^{k} \left(\binom{k}{r-1} \binom{k}{r} \right) a^r b^{k-r+1} + b^{k+1}$, since the ranges of the two terms have common range of $[1, k]$, so we sum this range together, and we are

left with the case of 0 and the case of $k+1$,
$= a^{k+1} + \sum_1^k \binom{k+1}{r} a^r b^{k-r+1} + b^{k+1}$, since $\binom{k}{r-1}\binom{k}{r} = \binom{k+1}{r}$
$= \sum_{r=0}^{k+1} \binom{k+1}{r} a^r b^{k+1-r}$, by taking the sum adding the case $k+1$ back to the range of the summation. Thus we have

$$(a+b)^n = \sum_{r=0}^{k+1} \binom{k+1}{r} a^r b^{k+1-r}$$

Q.E.D.

Definition 1.11. An ordered field is a field F which is also an ordered set, s.t.

(i) $x + y < x + z$ if $x, y, z \in F$ and $y < z$.

(ii) $xy > 0$ if $x \in F$, $y \in F$, $x > 0$, and $y > 0$.

If $x > 0$, we call x positive; if $x < 0$, x is negative.

Example 1.12. For example, $\mathbb{Q}$ is an ordered field.

□

Proposition 1.13. *The following statements are true in every ordered field [10]*

(a) If $x > 0$, then $-x < 0$, and vice versa.

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

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

(d) If $x \neq 0$, then $x^2 > 0$. In particular, $1 > 0$.

(e) If $0 < x < y$, then $0 < \frac{1}{y} < \frac{1}{x}$.

We shall discuss a little about the inequalities here.

(1) We have $x^2 \geq 0$ and the equivalence, $=$, holds if and only if $x = 0$.

(2) Assume $\forall x \geq 0$, $\exists \sqrt{x} \in F$, s.t. $x + \frac{1}{x} \geq 2$, $\forall x > 0$. Then we have $x - 2 + \frac{1}{x} \geq 0$. This is true from the first inequalities. We have a complete square $(\sqrt{x} + \frac{1}{\sqrt{x}})^2 \geq 0$.

(3) Cauchy-Schwarz Inequality, see the following theorem.

Theorem 1.14. *Let $(a_1, a_2, ..., a_n)$ and $(b_1, b_2, ..., b_n)$ be two sequences of real numbers, then*

$$\left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right) \geq \left(\sum_{i=1}^{n} a_i b_i\right)^2$$

Proof: The Cauchy-Schwarz inequality is an elementary inequality at the same time a powerful inequality. We consider the following complete square, which we know from the first inequality that it is larger or equal or zero. Then we open up the brackets and collect the identical terms.

$$\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i b_j - a_j b_i)^2$$

$$= \sum_{i=1}^{n} a_i^2 \sum_{j=1}^{n} b_j^2 + \sum_{i=1}^{n} b_i^2 \sum_{j=1}^{n} a_j^2 - 2\sum_{i=1}^{n} a_i b_i \sum_{j=1}^{n} b_j a_j$$

$$= 2\Big(\sum_{i=1}^{n} a_i^2\Big)\Big(\sum_{i=1}^{n} b_i^2\Big) - 2\Big(\sum_{i=1}^{n} a_i b_i\Big)^2$$

The left-hand side of the equation is a sum of the squares of real numbers, that is, greater or equal to zero, thus we have

$$2\Big(\sum_{i=1}^{n} a_i^2\Big)\Big(\sum_{i=1}^{n} b_i^2\Big) - 2\Big(\sum_{i=1}^{n} a_i b_i\Big)^2 \geq 0$$

$$\Rightarrow$$

$$\Big(\sum_{i=1}^{n} a_i^2\Big)\Big(\sum_{i=1}^{n} b_i^2\Big) \geq \Big(\sum_{i=1}^{n} a_i b_i\Big)^2$$

Q.E.D.

An interesting question is, when does = hold? This leads to some other famous inequalities.

Theorem 1.15. *AM-GM Theorem (Arithmetic Mean-Geometric Mean). The theorem states that for any set of nonnegative real numbers, the arithmetic mean of the set is greater than or equal to the geometric mean of the set. That is, for a set of nonnegative real numbers $a_1, a_2, ..., a_n$, the following holds*

$$\frac{a_1 + a_2 + ... + a_n}{n} \geq (a_1 a_2 ... a_n)^{\frac{1}{n}}$$

We can also write in the shorthand notation with summations and products:

$$\sum_{i=1}^{n} \frac{a_i}{n} \geq \prod_{i=1}^{n} a_i \frac{1}{n}$$

Proof: Note that the function $x \to ln(x)$ is strictly concave. Then, by Jensen's Inequality (proof see Jensen's Inequality), we have

$$ln \sum_i \lambda_i a_i \geq \sum_i \lambda_i ln(a_i) = ln \prod_i a_i^{\lambda_i}$$

with equality if and only if all the a_i are equal. Since $x \to ln(x)$ is a strictly increasing function, it then follows that

$$\sum_i \lambda_i a_i \geq \prod_i a_i^{\lambda_i}$$

with equality if and only if all the a_i are equal, as desired.

Q.E.D.

Theorem 1.16. *GM-HM Theorem (Geometric Mean-Harmonic Mean). For a set of nonnegative real numbers $a_1, a_2, ..., a_n$, the following holds*

$$(a_1 a_2 ... a_n)^{\frac{1}{n}} \geq \frac{n}{\frac{1}{a_1} + ... + \frac{1}{a_n}}$$

Proof: Proof is similar.

Theorem 1.17. *Jensen's Inequality. A real function $\phi(x)$ is convex on the interval $[a, b]$ if $\forall \theta \in (0, 1)$ we have*

$$\phi(\theta a + (1-\theta)b) \leq \theta\phi(a) + (1-\theta)\phi(b)$$

Notice that when $n = 2$, this is the definition of convexity.

Definition 1.18. Let ϕ is convex. $\forall a, b$, and $\forall \theta \in (0,1)$, we have $\phi(a) \leq \theta\phi(a) + (1-\theta)\phi(b)$

For this definition, we have the following observation. An ordered field has the characteristic of zero, and we have $1 + 1 + .. + 1 \neq 0$. The idea is the following:

(1) For $1 > 0$, we have $(1^2 = 1 \times 1 = 1)$ and $1 \neq 0 \Rightarrow 1 > 0$.

(2) $1 + 1 > 1 > 0$

(3) $1 + 1 + ... + 1 > 1 + 1 + ... + 1 > ... > 1 > 0$ by transitivity (the first term has n one's and the second term has $n - 1$ one's and so on)

Now we can go back ordered sets to introduce more definition and theorems.

Definition 1.19. Let $E \subseteq S$ to be bounded above (below) if $\exists M \in S$ s.t. $M \geq x$, $\forall x \in E$, ($M \leq x$, $\forall x \in E$).

Theorem 1.20. *Archimedean Property. The set or real numbers (an ordered field with the least upper bound property has the Archimedean Property.*

Lemma 1.21. *The set* $\mathbb{N}$ *of positive integers* $\mathbb{N} = \{0, 1, 2, ...\}$ *is not bounded from above.*

Proof: We prove this by contradiction, assume $\mathbb{N}$ is bounded from above. Since $\mathbb{N} \subset \mathbb{R}$ and $\mathbb{R}$ has the least upper bound property, then $\mathbb{N}$ has a least upper bound $\alpha \in \mathbb{R}$. Thus $n \leq \alpha$ for all $n \in \mathbb{N}$ and is the smallest such real number.

Consequently $\alpha - 1$ is not an upper bound for $\mathbb{N}$ (if it were, since $\alpha - 1 < \alpha$, then α would not be the least upper bound). Therefore there is some integer k with $\alpha - 1 < k$. But then $\alpha < k + 1$, which contradicts that α is an upper bound for $\mathbb{N}$.

Q.E.D.

Corollary 1.22. *Let x be any real number. Then there exists a natural number n such that $n > x$.*

Proof: Let $x \in \mathbb{R}$, and let $S = \{a \in \mathbb{N} : a \leq x\}$. The set S is either empty or nonempty.

If S is empty, let $n = 1$ and note that $x < n$ (since otherwise $1 \in S$).

If S is not empty, consider the following. Since S has an upper bound, S must have a least upper bound, call it b. Now consider $b - 1$. Since b is the least upper bound, $b - 1$ cannot be an upper bound of S. Thus, there exists some $y \in S$ such that $y > b - 1$. Let $n = y + 1$, then $n > b$. But y is a natural number, so n must also be a natural number. Since $n > b$, we know $n \notin S$; but is $n \notin S$, we can say $n > x$. Therefore, we have a natural number greater than x.

Q.E.D.

Corollary 1.23. *If x and y are real numbers with $x > 0$, there exists a natural n such that $nx > y$.*

Proof: Since x and y are real numbers, and $x \neq 0$, $\frac{y}{x}$ is a real number. By the Archimedean property, we can choose an $n \in \mathbb{N}$ such that $n > \frac{y}{x}$. Then $nx > y$.

Q.E.D.

Theorem 1.24. *The Density of Real Numbers. Let $x, y \in \mathbb{R}$ be any two real numbers where $x < y$. Then there exists a rational number $r \in \mathbb{Q}$ such that $x < r < y$.*

Proof: Suppose that $x > 0$. Since $x < y$ we have that $y >$0 and then we have $y - x > 0$. By the Archimedean properties, since $y - x > 0$, then there exists a natural number $n \in \mathbb{N}$ such that $\frac{1}{n} < y - x$. If we multiply n on both sides, we get $1 < ny - nx$ and rewrite as $nx + 1 < ny$. Now we know that since $n > 0$ and since $x > 0$, and by the Archimedean properties that since $nx > 0$ then there exists a natural number, call it $A \in \mathbb{N}$ such that $A - 1 \leq nx \leq A$ or equivalently $A \leq nx + 1 \leq A + 1$. Therefore $nx \leq A \leq nx + 1 \leq ny$ and so $nx < A < ny$ and so the rational number $r = \frac{A}{n}$ works for $x < r < y$.

Q.E.D.

Definition 1.25. For $E \subseteq S$ bounded above (below), sup$E$ has a smallest upper bound. Let sup$E = \alpha$ if

(1) $\alpha \geq x$, $\forall x \in E$

(2) assume $\exists \alpha' > x$, $\forall x \in E$, $\Rightarrow \alpha' \geq \alpha$

Remark 1.26. Let $E \subseteq S$ bounded above (below) $\not\Rightarrow$ sup$E$ exists. For example we have $S = \mathbb{Q}$ and $E = \{r \in \mathbb{Q} : r^2 < 2\}$. Is this bounded? Yes, we can simply consider $10 \geq r$, $\forall r \in E$. Is sup$E$ exists in $\mathbb{Q}$? No. If $\alpha =$sup$E \in \mathbb{Q}$, then we have $\alpha^2 = 2$. Thus we have $(\frac{m}{n})^2 = 2 \Rightarrow m^2 = 2n^2$ $\Rightarrow m, n$ are even which contradicts assumption.

Definition 1.27. Let $E \subseteq S$ bounded below in F, $E = \beta$ if

(1) $\beta \leq x$, $\forall x \in F$

(2) $\beta' \leq x$, $\forall x \in E \Rightarrow \beta' \leq \beta$

Theorem 1.28. *The Completeness Axiom. Every non-empty subset E of $\mathbb{R}$ that is bounded from above has a least upper bound sup$E \in \mathbb{R}$. Every non-empty subset E of $\mathbb{R}$ that is bounded from below has a greatest lower bound inf$E \in \mathbb{R}$.*

Proof: Let us prove the case that is bounded from below. Let T be the set $\{-s : s \in E\}$. Since E is bounded from below there is an $m \in \mathbb{R}$ such that $m \leq s$ for all $s \in E$. This implies that $-s \leq -m$ for all $s \in S$ and so $t \leq -m$ for all $t \in T$. Then T is bounded from above, hence by axiom, sup$T$ exists. Let $u = supT$. We show $-u =$inf$E$. That is, $\forall s \in E$, $-u \leq s$ and $\forall s \in E$, we have $t \leq s \Rightarrow u \leq q$. Set $q = -t$, we have $-s \leq -t$, $\forall s \in E \Rightarrow u \leq -t$. Hence, $\forall s \in E$, $t \leq s \Rightarrow t \leq -u$.

Q.E.D.

Next, we discuss Dedekind cut. The idea is simple. A Dedekind cut in an ordered field is a partition of it, (A, B), such that A is nonempty and closed downwards, B is nonempty and closed upwards. This satisfies all the axioms we discussed above, and hence is unique up to isomorphisms.

Let us recall the density of $\mathbb{Q}$. $\forall a, b \in F$, $\exists r \in \mathbb{Q}$ s.t. $a < r < b$. Consider the following question. How to define $\sqrt{2}$? Let $E = \{r \in \mathbb{Q} : r^2 < 2\}$, then

(1) $E \neq \emptyset$ while $0 \in E$

(2) Bounded above (we can consider $10 > r \ \forall r \in E$)

Therefore, we have $\sqrt{2} = supE = \alpha$. We can check the result. In this case, we have only three possibilities, $\alpha^2 > 2$, $\alpha^2 < 2$, or $\alpha^2 = 2$. We can see that it is neither of the first two cases. We construct $\sqrt{2} = sup\{r \in \mathbb{R} : r^2 < 2\}$, then it is well defined when $sup\{r \in \mathbb{R} : r^2 < 2\} = 2$, excluding the case $\alpha^2 > 2$ and $\alpha^2 < 2$. The consequence is $\sqrt{2} \notin \mathbb{R}$.

Let us take a closer look at this problem. If $\alpha^2 > 2$, then the goal is to find $r \in \mathbb{R}$. Since $r < \alpha$, α is an upper bound for E. If $\alpha^2 < 2$, then the goal is to find $r \in E$, s.t. $\alpha < 2$. We can take $r = \alpha + \frac{1}{n}$ and choose $n \to \infty$ and $n \in \mathbb{Q}$.

Proposition 1.29. *We have following properties for powers*

(i) $x^r x^s = x^{r+s}$

(ii) $(x^r)^s = x^{rs}$

(iii) $x^r y^r = (xy)^r$

Definition 1.30. Dedekind Cut. Let $E \subset \mathbb{Q}$, s.t.

(1) $\alpha \in E$ and $\alpha' < \alpha \Rightarrow \alpha' \in E$

(2) $\alpha \in E$, $\exists \beta \in E$, s.t. $\beta > \alpha$.

We can construct $\mathbb{R}$ and look at the set of all cuts.

(1) Ordered. $E_1 \subseteq E_2$ if $E_1 \subsetneq E_2$

(2) Operations. $E_1 < E_2$, if $E_1 \subset E_2$, with addition $(+)$ and multiplication $(\times)$

(3) Ordered field. We can use cancellation.

(4) Completeness. We can check bounded or not. We take a cut and check whether it is upper bound. Then look at the largest union. That is, there exists a complete ordered field that satisfies three properties (see theorem).

Theorem 1.31. *There exists a complete ordered field that*

$$\mathbb{R} = \{E \subseteq \mathbb{Q} : (1) E \neq \mathbb{Q}, (2) P \in E \ and \ q < p \Rightarrow q \in E, (3) p \in E \to \exists q \in E, s.t. q > p\}$$

With the concepts discussed in mind, we can ask questions like how big are these sets? Are all of the infinite sets countable? We can move on to the next section.

2 §Basic Topology§

Go back to Table of Contents. Please click TOC

2.1 Metric Spaces and Point Set Topology

Go back to Table of Contents. Please click TOC

Consider two sets A and B, whose elements may be any objects whatsoever. We can establish a map, a function f, from A to B. The set A is called the domain of f, and the elements $f(x)$ are called the values of f, or the range of f.

If for each $y \in B$, $f^{-1}(y)$ consists of at most one element of A, then f is said to be one-to-one mapping of A into B. If for all $y \in B$, there exists $x \in A$, s.t. $y = f(x)$. Then we say f is onto.

Definition 2.1. If there exists a one-to-one mapping of A onto B, we say that A and B can be put in one-to-one correspondence, or that A and B have the same cardinal number (or same cardinality), or, briefly, that A and B are equivalent, and we write $A \sim B$. Then it is clear that
It is reflexive: $A \sim A$
It is symmetric: $A \sim B$, then $B \sim A$
It is transitive: $A \sim B$ and $B \sim C$, then $A \sim C$.

Definition 2.2. For any positive integer n, let J_n be the set whose elements are the integers 1, 2, ..., n; let J be the set consisting of all positive integers. For any set A, we say:

(a) A is finite if $A \sim J_n$ for some n (the empty set is also considered to be finite)

(b) A is infinite if A is not finite

(c) A is countable if $A \sim J$

(d) A is uncountable if A is neither finite nor countable

(e) A is at most countable if A is finite or coutnable

Theorem 2.3. *Every infinite subset of a countable set A countable.*

Proof: Suppose $E \subset A$, and E is infinite. Arrange the elements x of A in a sequence $\{x_n\}$ of distinct elements. Construct a sequence $\{n_k\}$ as follows:

Let n_1 be the smallest positive integer such that $x_{n_1} \in E$. Having chosen $n_1, ..., n_{k-1} (k = 2, 3, 4, ...)$, let n_k be the smallest integer greater than n_{k-1} such that $x_{n_k} \in E$.

Putting $f(k) = x_{n_k} (k = 1, 2, 3, ...)$, we obtain a one-to-one correspondence between E and J. The theorem shows that, roughly speaking, countable sets represent the "smallest" infinity. No countable set can be a subset of a countable set [10].

Q.E.D.

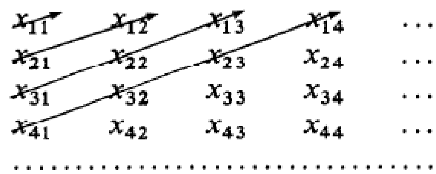
Theorem 2.4. *Let $\{E_n\}, n = 1, 2, 3, ...,$ be a sequence of countable sets, and put*

$$S = \bigcup_{n=1}^{\infty} E_n.$$

Then S is countable.

Proof: Let every set E_n be arranged in a sequence $\{x_{nk},\, k = 1, 2, 3, ...,$ and consider the infinite array

Figure 1: Graphic illustration of array



in which the elements of E_n from the nth row. The array contains all elements of S. As indicated by the arrows, these elements can be arranged in a sequence

$$x_{11}, x_{21}, x_{12}, x_{31}, x_{22}, x_{13}, ...$$

If any two of the sets E_n have elements in common, these will appear more than once in the series above. Hence there is a subset T of the set of all positive integers such that $S \sim T$, which shows that S is at most countable. Since $E_1 \subset S$, and E_1 is infinite, S is infinite, and thus countable [10].

Corollary 2.5. *Suppose A is at most countable, and for every $\alpha \in A$, B_α is at most countable. Put*

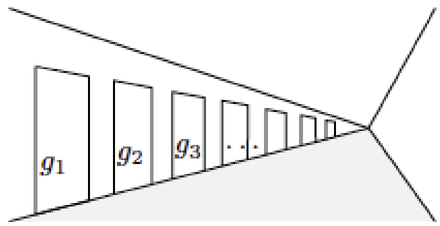
$$T = \bigcup_{\alpha \in A} B_\alpha.$$

Then T is at most countable. For T is equivalent to a subset of S such that $S = \bigcup_{n=1}^{\infty} E_n$.

There is famous story here called "Hilbert's hotel". Suppose a hotel has countably many rooms, numbered 1, 2, 3, ... with guest g_i occupying room i; so the hotel is fully booked. Now a new guest x arrives asking for a room, whereupon the hotel manager tells him: sorry, all rooms are taken. No problem, says the new arrival, just move guest g_1 to room 2, g_2 to room 3, g_3 to room 4, and so on, and I will then take room 1. To the manager's surprise this works; he can still put up all guests plus the new arrival x!

Now it is clear that he can also put up another guest y, and another one z, and so on. In particular, we note that, in contrast to finite sets, it may well happen that a proper subset of an infinite set M has the same size as M. In fact, this is a characterization of infinity: a set is infinite if and only if it has the same size as some proper subset [1].

Figure 2: Graphic illustration of Hilbert's Hotel



Theorem 2.6. *Let A be a countable set, and let B_n be the set of all n-tuples $(a_1, ..., a_n)$, where $a_k \in A$ $(k = 1, ..., n)$, and the elements $a_1, ..., a_n$ need not be distinct. Then B_n is countable.*

Proof: That B_1 is countable is evidence, since $B_1 = A$. Suppose B_{n-1} is countable $(n = 2, 3, 4, ...)$. The elements of B_n are of the form

$$(b, a), \ (b \in B_{n-1}, a \in A).$$

For every fixed b, the set of pairs (b, a) is equivalent to A, and hence countable. Thus B_n is the union of a countable set of countable sets. Then B_n is countable, which follows by induction [10].

Q.E.D.

Corollary 2.7. *The set of all rational numbers is countable.*

Proof: We apply previous theorem, with $n = 2$, noting that every rational r is of the form b/a, where a and b are integers. The set of pairs (a, b), and therefore the set of fractions b/a, is countable [10].

Q.E.D.

Theorem 2.8. *Let A be the set of all sequences whose elements are the digits 0 and 1. This set A is uncountable. The elements of A are sequences like 1,0,0,1,0,1,1,...*

Proof: Let E be a countable subset of A, and let E consist of the sequences $s_1, s_2, s_3, ...$ We construct a sequence s as follows. If the nth digit in s_n is 1, we let the nth digit of s be 0, and vice versa. Then the sequence s differs from every member of E in at least one place; hence $s \notin E$. But clearly $s \in A$, so that E is a proper subset of A.

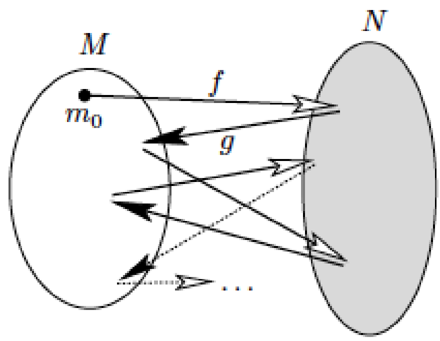
We have shown that every countable subset of A is a proper subset of A. It follows that A is uncountable (for otherwise A would be a proper subset of A, which is absurd) [10].

This leads us to the famous Cantor-Berstein theorem.

Theorem 2.9. *If each of two sets M and N can be mapped injectively into the other, then there is a bijection from M to N, hat is, $|M| = |N|$.*

Proof: We assume that M and N are disjoint - if not, then we just replace N be a new copy.

Figure 3: Graphic illustration of sets M and N in this theorem



The functions f and g map back and forth between the elements of M and those of N. One way to bring this potentially confusing situation into perfect clarity and order is to align $M \cup N$ into chains of elements: take an arbitrary element $m_0 \in M$, say, and from this generate a chain of elements by applying f, then g, the nf again, then g, and so on. The chain may close up (in Case 1.) if we reach m_0 again in this process, or it may continue with distinct elements indefinitely. (Note that the first "duplicate" in the chain cannot be an element different from m_0, by injectivity.)

If the chain continues in definitely, then we try to follow it backwards: from m_0 to $g^{-1}(m_0)$ if m_0 is in the image of g, then to $f^{-1}(g^{-1}(m_0))$ if $g^{-1}(m_0)$ is in the image of f, and so on. Three more cases may arise from here: the process of following the chain backwards may go on indefinitely (Case 2.), it may stop in an element of M that does not lie in the image of g (Case 3.), or it may stop in an element of N that does not lie in the image of f (Case 4.).

Thus $M \cup N$ splits perfectly into four types of chains, whose elements we may label in such a way that a bijection is simply given by putting $F : m_i \to n_i$. See the figures below for four cases.

Figure 4: Case 1. Finite cycles on $2k + 2$ distinct elements ($k \geq 0$).

$$\begin{tikzcd} m_0 \arrow[r, "f"] & n_0 \arrow[r, "g"] & m_1 \arrow[r, "f"] & \cdots & m_k \arrow[r, "f"] & n_k \arrow[lllll, bend left, "g"] \end{tikzcd}$$

Figure 5: Case 2. Two-way infinite chains of distinct elements.

$$\cdots \longrightarrow m_0 \xrightarrow{f} n_0 \xrightarrow{g} m_1 \xrightarrow{f} n_1 \xrightarrow{g} m_2 \xrightarrow{f} \cdots$$

Figure 6: Case 3. The one-way infinite chains of distinct elements that start at the elements $m_0 \in M \setminus g(N)$.

$$m_0 \xrightarrow{f} n_0 \xrightarrow{g} m_1 \xrightarrow{f} n_1 \xrightarrow{g} m_2 \xrightarrow{f} \cdots$$

Figure 7: Case 4. The one-way infinite chains of distinct elements that start at the elements $n_0 \in N \setminus f(M)$.

$$n_0 \xrightarrow{g} m_0 \xrightarrow{f} n_1 \xrightarrow{g} m_1 \xrightarrow{f} \cdots$$

Q.E.D.

With this understanding, we can move on to discuss metric spaces.

Definition 2.10. A set X, whose elements we shall call *points*, is said to be metric space if with any two points p and q of X there is associated a real number $d(p, q)$, called the distance from p to q, such that

(a) $d(p, q) > 0$ if $p \neq q$; $d(p, p) = 0$

(b) $d(p, q) = d(q, p)$;

(c) $d(p, q) \leq d(p, r) + d(r, q)$, $\forall r \in X$.

Any function with these three properties is called a distance function, or a metric. We can think of the following examples. For real set, $\mathbb{R}$, the distance between two points $d(x, y) = |x - y|$, and the triangle inequality tells us $|x + y| \leq |x| + |y|$. Moreover, for $\mathbb{R}^n$, we have $d(x, y) = ||x - y||$, and $||x|| = \sqrt{x \circ x}$ while $x > y = \sum_{i=1}^{n} x_i y_i$.

Here it worth to note the details of triangle inequality.

Theorem 2.11.

$$d(x, y) \leq d(x, z) + d(z, y)$$

Proof: Let $d(x, y) = ||x - y||$, $d(x, z) = ||x - z||$, and $d(z, y) = ||z - y||$. Taking the power of two on both sides and open up the brackets, we have $||a + b||^2 = ||a||^2 + ||b||^2 + 2ab$. Then $||a||^2 + ||b||^2 + 2ab \leq ||a||^2 + ||b||^2 + 2||a||||b||$.

$$\Rightarrow ab \leq ||a||||b||$$

which is the Cauchy - Schwarz equation.

Definition 2.12. Let X be a metric space. All points and sets mentioned below are understood to be elements and subsets of X.

(a) A neighborhood of p is a set $N_r(p)$ consisting of all q such that $d(p, q) < r$, for some $r > 0$. The number r is called the radius of $N_p(p)$.

(b) A point p is a limit point of the set E if every neighborhood of p contains a point $q \neq p$ such that $q \in E$.

(c) If $p \in E$ and p is not a limit point of E, then p is called an isolated point of E.

(d) E is closed if every limit point of E is a point of E.

(e) A point p is an interior point of E if there is a neighborhood N of p such that $N \subset E$. Or we can notate the definition with shorthand notations. The set of interior points is as the following: $E^o = \{x \in X : \ d(x, E^c) > 0\} = \bigcup_{Br(X) \subseteq E} Br(x)$

(f) E is open if every point of E is an interior point of E.

(g) The complement of E (denoted by E^c) is the set of all points $p \in X$ such that $p \notin E$.

(h) E is perfect is E is closed if every point of E is a limit point of E.

(i) E is bounded if there is a real number M and a point $q \in X$ such that $d(p, q) < M$ for all $p \in E$.

(j) E is dense in X if every point of X is a limit point of E, or a point of E (or both).

From the definition of metric space, limit point, interior point, perfect, and closures, we can show the following proofs.

Theorem 2.13. *Every neighborhood is an open set.*

Proof: Consider a neighborhood $E = N_r(p)$, and let q be any point of E. Then there is a positive real number h s.t.

$$d(p, q) = r - h.$$

For all points s s.t. $d(q, s) < h$, we have then $d(p, s) \leq d(p, q) < h$, we have then

$$d(p, s) \leq d(p, q) + d(q, s) < r - h + h = r,$$

so $s \in E$. Thus q is an interior point of E.

Q.E.D.

We can take a look at the following examples to understand closed, open, perfect, and bounded properties. We can simply consider the following subsets of R^2.

(a) The set of all complex z s.t. $|z| < 1$.

(b) The set of all complex z s.t. $|z| \leq 1$.

(c) A nonempty finite set.

(d) The set of all integers.

(e) The set consisting of the numbers $\frac{1}{n}$ $(n = 1, 2, 3, ...)$. Let us note that this set E has a limit point (namely, $z = 0$) but that no point of E is a limit point of E; we wish to emphasize the difference between having a limit point and containing one.

(f) The set of all complex numbers (that is, R^2).

(g) The segment (a, b).

Then we have the following checklist:

-	Closed	Open	Perfect	Bounded
(a)	No	Yes	No	Yes
(b)	Yes	No	Yes	Yes
(c)	Yes	No	No	Yes
(d)	No	Yes	No	No
(e)	No	No	No	No
(f)	Yes	Yes	Yes	No
(g)	No	Y/N	No	Yes

Here we have an interesting theorem about finite and infinite sets.

Theorem 2.14. *Let $\{E_\alpha\}$ be a finite or infinite collection of sets E_α. Then*

$$\left(\bigcup_\alpha E_\alpha\right)^c = \bigcap_\alpha (E_\alpha^c).$$

Proof: Let A and B be the left and right members of the equation. If $x \in A$, then $x \notin \bigcup_\alpha E_\alpha$ for any α, hence $x \in E_\alpha^c$ for every α, so that $x \in \bigcap E_\alpha^c$. Thus $A \subset B$.

Conversely, if $x \in B$, then $x \in E_\alpha^c$ for every α, hence $x \in E_\alpha$ for any α. This implies that $x \notin \bigcup_\alpha E_\alpha$. Thus $B \subset A$.

This completes the proof.

Q.E.D.

Theorem 2.15. *A set E is open if and only if its complement is closed.*

Proof: First, suppose E^c is closed. Choose $x \in E$. Then $x \in E^c$, and x is not a limit point of E^c. Hence there exists a neighborhood N of x s.t. $E^c \bigcap N$ is empty, that is, $N \subset E$. Thus x is an interior point of E, and E is open.

Next, suppose E is open. Let x be a limit point of E^c. Then every neighborhood of x contains a point of E^c, so that x not an interior point of E. Since E is open, this means that $x \in E^c$. It follows that E^c is closed.

Q.E.D.

Theorem 2.16. *This theorem states the following four properties.*

(a) For any collection $\{G_a\}$ of open sets, $\bigcup_\alpha G_\alpha$ is open.

(b) For any collection $\{F_\alpha\}$ of closed sets, $\bigcap_\alpha F_\alpha$ is closed.

(c) For any finite collection $G_1, ..., G_n$ of open sets, $\bigcap_{i=1}^n G_i$ is open.

(d) For any finite collection $F_1, ..., F_n$ of closed sets, $\bigcup_{t=1}^n F_i$ is closed.

Then we can introduce the following definition.

Definition 2.17. If X is a metric space, if $E \subset X$, and if E' denotes the set of all limit points of E in X, then the closure of E is the set $\bar{E} = E \cup E'$.

Theorem 2.18. *If X is a metric space and $E \subset X$, then*

(a) $\bar{X}$ is closed,

(b) $E = \bar{E}$ if and only if E is closed,

(c) $\bar{E} \subset F$ for every closed set $F \subset X$ such that $E \subset F$.

Remark 2.19. Suppose $E \subset Y \subset X$, where X is a metric space. To say that E is an open subset of X means that to each point $p \in E$ there is associated a positive number r such that the conditions $d(p,q) < r$, $q \in X$ imply that $q \in E$. But we have already observed that Y is also a metric space (see [10] 3E Page 31), so that our definitions may equally well be made within Y. To be explicit, we say that E is open relative to Y if to each $p \in E$ there is associated an $r > 0$ such that $q \in E$ whenever $d(p,q) < r$ and $q \in Y$.

Definition 2.20. If X is a metric space, if $E \subseteq X$, and if E' denotes the set of all limit points of E in X, then the closure of E is the set $\bar{E} = E \bigcup E'$.

Theorem 2.21. *If X is a metric space and $E \subset X$, then*

1. *$\bar{E}$ is closed,*
2. *$E = \bar{E}$ if and only if E is closed,*
3. *$\bar{E} \subset F$ for every closed set $F \subset X$ s.t. $E \subset F$.*

With this understanding, let us discuss some propositions and theorems.

Proposition 2.22. *The following propositions are about open sets.*

(1) *X and $\emptyset$ are open*

(2) *$\{G_\alpha\}$ be a collection of open sets $\Rightarrow \bigcup_\alpha G_\alpha$.*

(3) *$G_1, ..., G_n$ is open $\Rightarrow \bigcap_{i=1}^n G_i$ is open.*

Proposition 2.23. *Furthermore, we have the following propositions.*

(1) *E^o is the largest open set and $E^o \subset E$, i.e. if $G \subset E$ and G is open, then $\subset E^o$.*

(2) *$E^o = \bigcup_{G \text{ open and } G \subset E} G$.*

(3) *$(E^o)^o = E^o$.*

(4) *$E^o = E \Leftrightarrow E$ is open.*

Proof:

(1) Let G is open and $G \leq E$. $\forall x \in G$, $\exists r > 0$, s.t. $Br(x) \subset G$. Then $\because G \subset E \Rightarrow (by\ transi.)\ Br(x) \subseteq E \Rightarrow x \in E^o$.
$\therefore G \subseteq E^o$.

(2) E^o is an open set. Every $x \in E^o$, is an interior of E^o. $x \in E^o \Rightarrow \exists x > 0$, s.t. $Br(x) \subseteq E$. $\forall y \in Br(x)$, there is $r(y) > 0$ s.t. $Br(y) \subseteq Br(x) \subseteq E$. This implies that $y \in E^o$, which then implies $Br(x) \subseteq E^o \Rightarrow E^o \Rightarrow x$ is an interior point of E^o.

Then for $E^o = \bigcup_{G\ open\ G \subseteq E} G$, E^o is open and $E^o \subseteq E \Rightarrow E^o \Rightarrow E^o \subseteq \bigcup_{G\ open\ G \subseteq E}$

(3) $(E^o)^o = E^o$. We show the proof by "$L \subseteq R$", and "$R \subseteq L$".

For "$L \subseteq R$", let $x \in (E^o)^o$, and we can consider $x \in F^o$, *while* $F = E^o$. That is, $\exists x \in F^o\ s.t.\ (d, F^c) > 0$, which implies $x \in E^o$ since $F^c \subseteq E^o$.

For "$R \subseteq L$", the proof is backward.

(4) $E^o = E \Leftrightarrow E$ is open. We show the proof by "$L \Rightarrow R$", and "$R \Leftarrow L$". For "$L \Rightarrow R$", from definition, if $E^o = E$, then $\exists x \in E\ s.t.\ x \in E$ and vice versa. That is, $x \in E\ s.t.\ d(x, y) > 0, \forall y \in E$, which implies that E is open.

For "$L \Leftarrow R$", the proof is backward.

Q.E.D.

Theorem 2.24. *E is closed if and only if E^c is open.*

Proof: We prove "$\Rightarrow$" and we prove "$\Leftarrow$".

"$\Rightarrow$": Suppose that E is closed. Let x be in E^c, which means $x \notin E$. That means x is not a limit point of E. Then there exists a neighborhood N s.t. $N \subset E^c$, which means E^c is open.

"$\Leftarrow$": Assume E^c is open. Let x be a limit point of E. Then every neighborhood of x contains E. This implies that x is not an interior point of E^c. Then E^c is open $\Rightarrow x \in E$.

Q.E.D.

Theorem 2.25. *De Morgan's First Law:*

(1) $(A \bigcup B)^c = (A)^c \bigcap (B)^c$.

(2) $(A \bigcap B)^c = (A)^c \bigcup (B)^c$.

Proof:

(1) Consider $x \in (A \bigcup B)^c$. We can start with $x \in (A \bigcup B)^c$, then $x \notin A \bigcup B$, by definition of compliment.
$\Rightarrow x \in (A \bigcup B)^c$, by definition of $\notin$.
$\Rightarrow (x \in A \bigcup x \in B)^c$, by definition of $\bigcup$.
$\Rightarrow (x \in)^c \bigcap (x \in B)^c$
$\Rightarrow (x \notin A) \bigcap (x \notin B)$, by definition of $\notin$
$\Rightarrow (x \in A^c) \bigcap (x \in B^c)$, by definition of compliment.
$\Rightarrow x \in A^c \bigcap B^c$, by definition of $\bigcap$.

(2) Consider $x \in (A \bigcap B)^c$
$\Rightarrow x \in (A \bigcap B)^c$ then $x \notin A \bigcap B$, by definition of compliment
$\Rightarrow (x \in A \bigcap B)^c$, by definition of $\notin$
$\Rightarrow (x \in A \bigcap x \in B)^c$, by definition of $\bigcap$
$\Rightarrow (x \in A)^c \bigcap (x \in B)^c$
$\Rightarrow (x \notin A) \bigcap (x \notin B)$, by definition of $\notin$.
$\Rightarrow (x \in A^c) \bigcap (x \in B^c)$, by definition of compliment.
$\Rightarrow x \in A^c \bigcap B^c$, by definition of $\bigcap$.

Q.E.D.

2.2 Compact Sets

Go back to Table of Contents. Please click TOC

This subsection we introduce compact set. Before we start, let us review several properties from finite set.

Proposition 2.26. *Let E be a finite set: we have*

(1) $f : E \to \mathbb{R} \Rightarrow f$ *is bounded.*

(2) $f : E \to \mathbb{R}$: *max and min are attained in* E.

(3) $x_1, x_2 \in E \Rightarrow x_{n_1}, x_{n_2}, ...$ *is a constant series, which leads to the pigeon hole principle.*

(4) *If* $E \subseteq \bigcup_\alpha A \Rightarrow \alpha_1, ..., \alpha_n$ *s.t.* $E \subseteq \bigcap_{i=1}^n A_\alpha$.

An exercise can be the following: Assume (X, d) to be a metric space. Prove that E finite and $E \subseteq X \Rightarrow E$ is closed.

Here let us pause for a minute to discuss the famous Pigeon-hole Principle. The principle states the following: "If n objects are placed in r boxes, where $r < n$, then at least one of the boxes contains more than one object." ([1])

This is intuitive, but let us describe the proof in mathematics language. Let N and R be two finite sets with $|N| = n > r = |R|$, and let $f : N \to R$ be a mapping. Then there exists some $a \in R$ with $|f^{-1}(a)| \geq 2$. We may even state a stronger inequality: there exists some $a \in R$ with $|f^{-1}(a)| \geq \lceil \frac{n}{r} \rceil$. Otherwise we would have got $|f^{-1}(a)| < \frac{n}{r}$ for all a, and then $n = \sum_{a \in R} |f^{-1}(a)| < r \frac{n}{r} = n$, which cannot be. Hence the proof is complete. The concept of Pigeon-hole Principle is applied here in similar ways.

Figure 8: Graphic view for a pigeon-hole from a bird's point of view [1].

Definition 2.27. By an open cover of a set E in a metric space X we mean a collection $\{G_z\}$ of open subsets of X such that $E \subset \bigcup_\alpha G_\alpha$.

Definition 2.28. A subset K of a metric space X is said to be compact if every open cover of K contains a finite subcover.

More explicitly, the requirement is that if $\{G_\alpha\}$ is an open cover of K, then there are finitely many indices $\alpha_1, ..., \alpha_n$ s.t.

$$K \subset G_{\alpha_1} \cup ... \cup G_{\alpha_n}$$

The notion is very important in analysis, especially in connection with continuity. It is clear that every finite set is compact.

From the definition, we have the following theorems.

Theorem 2.29. *Any finite set is compact.*

Proof: Let $\{G_\alpha\}$ be a set covering $K = \{\alpha_1, ...\alpha_n\}$. Then $\forall x_i \in K$, $\exists X_i \in G_{\alpha_i} \Rightarrow G_{\alpha_1}, ...G_{\alpha_n}$ covers K.

Q.E.D.

Theorem 2.30. *The set of real numbers is not compact.*

Proof: Consider the set of real numbers, $\mathbb{R}$. Take the covering, say, $(-\infty, \sqrt{2} - \frac{1}{n})$ and $(\sqrt{2} + \frac{1}{2}, +\infty)$. $\nexists$ a subcover that will contain all of the elements of $\mathbb{R}$, since you can only go close to $\sqrt{2}$ but not exactly same to it.

Q.E.D.

An exercise following these two theorems can be trying to prove that E is unbounded implies that E is not compact. The idea is straightforward. Take E to be unbounded. By definition, there is not an element allows one to take the points inside E to be on the boundary (or the set of limit points, if you will). Then there cannot be a collection of open cover, hence completes the proof.

Another interesting exercise can be the following: let $I = (0, 1)$, and show that I is not compact. Consider $G_n = (\frac{1}{n+1}, \frac{1}{n-1})$ with $n \geq 2$ and $n \in \mathbb{N}$. Is there any open covering? $\forall x \in (0, 1)$, $n \geq 2$, we have $x \in (\frac{1}{n+1}, \frac{1}{n-1})$. Let $E = \{K : k \geq 2 \text{ } and \text{ } k \in \mathbb{N}\} \text{ } while \text{ } x > \frac{1}{n+1}\}$. Now we discuss n in the two cases:

(1) $E \neq \emptyset \Rightarrow$ we can use Archimedean property.

(2) E is bounded below, say choose 2, then $inf(E)$ is well defined.

$$\Rightarrow \frac{1}{inf(E) + 1} \leq x$$

and $inf(E) \geq 2$ while $inf(E) \in \mathbb{N}$. Then there are two more scenarios: (1) if we take "$<$", then it implies that $n = inf(E)$, which implies that $x < \frac{1}{n-1}$, which causes a contradiction; for (2) if we take "$=$", then it implies $x = \frac{1}{m}$ for some $m \in \mathbb{N}$, which is trivial.

Proposition 2.31. *Let (X, d) be a metric space:*

(1) $K \subseteq X$ is compact $\Rightarrow K$ is closed.

(2) F is closed $\subseteq K$ compact $\Rightarrow F$ is compact. (Reader can also go to Corllary 1.66 for more details.)

(3) E infinite $\subseteq K$ compact $\Rightarrow K \bigcap E' \neq \emptyset$.

Proof: We prove the propositions accordingly:

(1) Let K be compact and let $x \in K^c$. Choose for any point $p \in K$, s.t. $r = \frac{1}{3}d(x, p)$. That is, $\{B_{x,p}(p)\}$ covers $k \Rightarrow \exists p_1, p_2, ...$ s.t. $\{B(p_i)\}$ also covers k. Then we simply take $r_x = min(r_{x,p_1}, ..., r_{x,p_n}) > 0$, which implies $B(x) \bigcap k = \emptyset \Rightarrow x \in (X^c)^c$. Hence, $B \setminus x \in K^c$ is arbitrary $\Rightarrow K^c$ is open $\Leftrightarrow K$ is closed.

(2) Let $\{G_\alpha\}$ be an open covering of $F \subseteq K$. This implies that $F^c \bigcap \{G_\alpha\}$ open covering K. This is because K is compact and there exists a finite subcovering from $F^c \bigcap \{G_\alpha\}$. Thus, for F^c, we have $G_{\alpha_1}, ..., G_{\alpha_n}$ covers K, which implies that they also cover F while $F \subseteq K$, therefore also covers F. (Please see Remark after proposition for detailed comments.)

(3) We prove by contradiction. Assume that $\nexists p \in k$ s.t. $p \in E'$. That is, $\exists r_p > 0$, s.t. $B_p(p) \cap E = \emptyset$ by definition. Then we have $\{B_p(x)\}, p \subset K$ covers K, which implies that $\exists p_1, ...p_n$, s.t. $B(p_1), ..., B(p_n)$ covers $E \subseteq K$, which causes a contradiction since E is finite.

Q.E.D.

Remark 2.32. From part (3) of the propositions above, note that F is closed and K is compact. That is, "$F \bigcap K$ is closed and K is compact" implies "K is closed; and the intersection of closed sets is closed". On the other hand, $F \bigcap K$ is compact, because $F \bigcap K$ is closed and is a subset of K while K is compact.

Theorem 2.33. *Suppose $K \subset Y \subset X$. Then K is compact relative to X if and only if K is compact relative to Y.*

Proof: Suppose K is compact relative to X, and let $\{V_\alpha\}$ be a collection of sets, open relative to Y, s.t. $K \subset \bigcup_\alpha V_\alpha$. Then there are sets G_α, open relative to X, s.t. $V_\alpha = Y \cap G_\alpha$, for all α; and since K is compact relative to X, we have

$$K \subset G_{\alpha_1} \cap ... \cap G_{\alpha_n}$$

for some choice of finitely many indices $\alpha_1, ..., \alpha_n$. Since $K \subset Y$,

$$K \subset V_{\alpha_1} \cap ... \cap V_{\alpha_n}$$

which proves that K is compact relative to Y.

Q.E.D.

Theorem 2.34. *Compact subsets of metric spaces are closed.*

Proof: Let K be a compact subset of a metric space X. Suppose $p \in X$. If $q \in K$, let V_q and W_q be neighborhoods of p and q, respectively, of radius less than $\frac{1}{2}d(p, q)$. Since K is compact, there are finitely many points $q_1, ..., q_n$ in K s.t.

$$K \subset W_{q_1} \cap ... \cap W_{q_n} = W.$$

If $V = V_{q_1} \cap ... \cap V_{q_n}$, then V is a neighborhood of p which does not intersect W. Hence, $V \subset K^c$, so that p is an interior point of K^c, which completes the proof.

Q.E.D.

Theorem 2.35. *Closed subsets of compact sets are compact.*

Proof: Suppose $F \subset K \subset X$, F is closed (relative to X), and K is compact. Let $\{V_\alpha\}$ be an open cover of F. If F^c is ajoined to $\{V_\alpha\}$, we obtain an open cover Ω of K. Since K is compact, there is a finite subcollection Φ of Ω which covers K¡ and hence F. If F^c is a member of Φ, we may remove it from Φ and still retain an open cover of F. We have thus shown that a finite subcollection of $\{V_\alpha\}$ covers F.

Q.E.D.

Corollary 2.36. *If F is closed and K is compact, then $F \cap K$ is compact.*

Proof: We know from previous theorem that $F \cap K$ is closed; since $F \cap K \subset K$, then $F \cap K$ is compact.

Q.E.D.

Theorem 2.37. *If $\{K_\alpha\}$ is a collection of compact subsets of a metric space X s.t. the intersection of every finite subcollection of $\{K_\alpha\}$ is nonempty, then $\bigcap K_\alpha$ is nonempty.*

Proof: Fix a member K_1 of $\{K_\alpha\}$ and put $G_\alpha = K_\alpha^c$. Assume that no point of K_1 belongs to every K_α. Then the sets G_α form an open cover of K_1; and since K_1 is compact, there are finitely many indices $\alpha_1, ..., \alpha_n$ such that $K_1 \subset G_{\alpha_1} \cap ... \cap G_{\alpha_n}$. But this means that

$$K_1 \cap K_{\alpha_1} \cap ... \cap K_{\alpha_n}$$

is empty, in contradiction to out hypothesis.

Q.E.D.

Theorem 2.38. *If E is an infinite subset of a compact set K, then E has a limit point in K.*

Proof: if no point of K were a limit point of E, then each $q \in K$ would have a neighborhood V_q which contains at most one point of E (namely, q and $q \in E$). IT is clear that no finite subcollection of $\{V_1\}$ can cover E; and the same is true of K, since $E \subset K$. This contradicts the compactness of K.

Q.E.D.

Theorem 2.39. *If $\{I_n\}$ is a sequence of intervals in R^1, s.t. $K_{n+1} \subset K_n$ while $n = 1, 2, 3, ...$, then $\cap_1^\infty K_n$ is not empty.*

Proof: If $I_n = [a_n, b_n]$, let E be the set of all a_n. Then E is nonempty and bounded above (by b_1). Let x be the sup of E. If m and n are positive integers, then

$$a_n \leq a_{m+n} \leq b_{m+n} \leq b_m,$$

so that $x \leq b_m$ for each m. Since $a_m \leq x$, we see that $x \in I_m$ for $m = 1, 2, 3, ...$

Q.E.D.

Theorem 2.40. *Every k-cell is compact.*

Proof: Let I be a k-cell, consisting of all points $x = (x_1, ..., x_k)$ such that $a_j \leq x_j \leq b_j (1 \leq j \leq k)$. Put

$$\delta = \left\{ \sum_1^k (b_j - a_j)^2 \right\}^{1/2}.$$

Then $|x - y| \leq \delta$, if $x \in I$, y$\in I$.

Suppose, to get a contradiction, that there exists an open cover $\{G_\alpha\}$ of I which contains no finite subcover of I. Put $c_j = (a_j + b_j)/2$. The intervals $[a_j, c_j]$ and $[c_j, b_j]$ then determine w^k k-cells Q_1 whose union is I. At least one of these sets Q_1, call it I_1, cannot be covered by any finite subcollection of $\{G_\alpha\}$ (otherwise I could be so covered). We next subdivide I_1 and continue the process. We obtain a sequence $\{I_n\}$ with the following properties:

(a) $I \supset I_1 \supset I_2 \supset I_3 \supset ...$;

(b) I_n is not covered by any finite subcollection of $\{G_\alpha\}$;

(c) $if x \subset I_n$ and $y \subset I_n$, then $|x-y| \leq 2^{-n}\delta$.

By (a) and previous theorem, there is a point x^* which lies in every I_n. For some α, $x^* \in G_\alpha$. Since G_α is open, there exists $f > 0$ such that $|y - x^*| < r$ implies that $y \in G_\alpha$. If n is so large that $2^{-n}\delta < r$ (there is such an n, for otherwise $2^n \leq \delta r$ for all positive integers n, which is absurd since R is archimedean), then (c) implies that $I_n \subset G_\alpha$, which contradicts (b).

This completes the proof.

Q.E.D.

Theorem 2.41. *If a set E in R^k has one of the following three properties, then it has the other two:*

(a) E is closed and bounded.

(b) E is compact.

(c) Every infinite subset of E has a limit point in E.

Remark 2.42. This is covered in Heine-Borel Theorem.

Theorem 2.43. *Weierstrass. Every bounded infinite subset of R^k has a limit point in R^k.*

Proof: Consider the set E to be bounded. Let E to be a subset of a k-cell $I \subset R^k$. By previous theorem, I is compact, and so E has a limit point in I, which completes the proof.

Q.E.D.

Theorem 2.44. *If $K \subset X$ is compact and $K' \subset K$ is closed, then K' is compact.*

Remark 2.45. Two sources I am using for this famous theorem are one from University of Washington (`https://www.math.washington.edu/~morrow/334_14/heine-borel.pdf`) and another one from Princeton Math Department (`https://web.math.princeton.edu/~mdamron/teaching/F12/notes/MAT_215_Lecture_10.pdf`).

To understand Heine-Borel Theorem, we need some background on compactness. Throughout the proof of the theorem, we have two other lemmas. Hence, let us do them accordingly.

Proof: Let $\mathcal{C}$ be an open cover of K'. Define

$$\mathcal{D} = \mathcal{C} \cup \{K'^c\}$$

and note that $\mathcal{D}$ is actually an open cover of K. Therefore, as K is compact; we can extract from $\mathcal{D}$ a finite subcover $\{D_1, ..., D_n\}$. If $D_i \in \mathcal{C}$ for all i, then we are done; otherwise K'^c is in this set (say it is D_n) and we consider the collection $\{D_1, ..., D_{n-1}\}$. This is a finite sub collection of $\mathcal{C}$. We claim that it is an open cover of K' as well. Indeed, if $x \in K'$ then there exists $i = 1, ..., n$ such that $x \in D_i$. Since $x \notin K'^c$, D_i cannot equal K'^c, meaning that $i \neq n$, which completes the proof.

Q.E.D.

Theorem 2.46. *Heine-Borel. A set $K \subseteq \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Remark 2.47. To prove this theorem, we will need some preliminary results. Recall that Rudin defines an n-cell to be a subset of $\mathbb{R}^n$ of the form

$$[a_1, b_1] \times ... \times [a_n, b_n] \ for \ a_i \leq b_i, \ i = 1, ..., n.$$

Please note that the two lemmas below should be considered part of the proof of Heine-Borel Theorem. One should really start the proof by proving the lemmas first.

Proof: Suppose that K is closed and bounded $\mathbb{R}^n$. Then there exists an $n - cell$ C such that $K \subset K$. By lemma (see below), C is compact. But K is a closed subset of C so K is compact.

Suppose conversely that K is compact (which is from the second lemma below). Then we have already shown K is closed and bounded.

Q.E.D.

Lemma 2.48. *Suppose that C_1, C_2, ... are n-cells that are nested; that is, if $C_i = \prod_{k=1}^{n}[a_i^{(k)}, b_i^{(k)}]$, then*

$$[a_{i+1}^{i+1}, b_{(k)}^{(k)}] \subset [a_i^i, b_k^k], \forall i \ and \ \forall k.$$

Then $\cup_i C_i$ is nonempty.

Proof: We first consider the case $n = 1$. That is, take $C_i = [a_i, b_i]$ for $i \geq 1$ and $a^i \leq b^i$. Define $A = \{a_1, a_2, ...\}$ and $a = sup\ A$. We claim that

$$a \in \cup_i C_i.$$

To see this, note that $a_i \leq b_j$ for all i, j. Indeed,

$$a_i \leq a_j \leq b_j, \ if \ i \leq j$$

and

$$a_i \leq b_i \leq b_j, \ if \ i \geq j.$$

Therefore b_j is an upper bound for A. But A is the least upper bound of A so $a \leq b_j$ for all m. This gives

$$a_i \leq a \leq b_i \ for \ all \ m,$$

or $a \in \cup_i C_i$.

For the case $n \geq 2$ we just do the same argument on each of the coordinates to find $(a(1), ..., a(n))$ such that

$$a_i^{(k)} \leq a(k) \leq b_i^{(k)}, \ for \ all \ i, k,$$

or $(a(1), ..., a(n)) \in \cup_i C_i$.

Q.E.D.

Lemma 2.49. *Any n-cell is compact in $\mathbb{R}^n$.*

Proof: For simplicity, take $K = [0,1] \times ... \times [0,1] = [0,1]^n$. Since $\mathbb{R}^n$ is a metric space (with the usual metric), it suffices to prove that K is limit point compact; that is, that each infinite subset of K has a limit point in K. (This is from a homework problem this week which states that compactness and limit point compactness are equivalent in metric spaces.)

Suppose that $E \subset K$ is infinite. We will produce a limit point of E inside K. We begin by dividing K into 2^n sub-cells by cutting each interval $[0,1]$ into two equal pieces. For instance, in $\mathbb{R}^2$ we would consider the 4 sub-cells

$$[0, \frac{1}{2}] \times [0, \frac{1}{2}], [0, \frac{1}{2}] \times [\frac{1}{2}, 1], [\frac{1}{2}, 1] \times [0, \frac{1}{2}], [\frac{1}{2}, 1] \times [\frac{1}{2}, 1].$$

At least one of these $2n$ sub-cells must contain infinitely many points of E. Call this sub-cell K_1. Repeatedly dividing K_1 into $2n$ equal sub-cells to find sub-sub-cell K_2 which contains infinitely many points of E.

We continue this procedure ad infinitum, at stage $i \geq 1$ finding a sub-cell K_i of K of the form

$$K_i = [r_{1,i}2^{-i}, (r_{1,i}+1)2^{-i}] \times ... \times [r_{n,i}2^{-i}, (r_{n,i}+1)r^{-i}]$$

which contains infinitely many points of K. Note that the K_i's satisfy the conditions of the previous lemma: they are nested n-cells. Therefore there exists $z \in \cap_i K_i$. Because each K_i is a subset of K, we have $z \in K$.

We claim that z is a limit point of E. To show this, let $r > 0$. Note that for all points $x, y \in K_i$ we have

$$|x-y|^2 = (x_1 - y_1)^2 + ... + (x_n - y_n)^2 \leq n(2^{-i})^2 = \frac{n}{4^i}.$$

Therefore

$$diam(K_i) = sup\{|x-y| : \ x, y \in K_i\} \leq \frac{\sqrt{n}}{2^i} \leq \frac{\sqrt{n}}{i}.$$

(You can prove this inequality $i \leq 2^i$ for all i by induction.) So fix any $i > \frac{\sqrt{n}}{r}$; then for all $x \in K_i$, we have (since $z \in K_i$)

$$|x-z| \leq diam(K_i) \leq \frac{\sqrt{n}}{i} < r,$$

so that $K_i \subset B_r(z)$. However K_i contains infinitely many points of E, so we can find one not equal to z in $B_r(z)$. This means z is a limit point of E.

Q.E.D.

Remark 2.50. From the idea of the theorem, we find that

$$(closed\ +\ bounded) \Rightarrow (compact)\ in\ \mathbb{R}^n,$$

$$(closed\ +\ bounded) \Leftarrow (compact)\ in\ metric\ spaces,$$

and

$$(compact) \Leftrightarrow (limit\ point\ compact)\ in\ metric\ spaces.$$

3 §Numerical Sequences and Series§

Go back to Table of Contents. Please click TOC

3.1 Sequences

Go back to Table of Contents. Please click TOC

Let us start this section with some terminologies. For $n \in \mathbb{N}$, we have a sequence $S(n)$. We say that $S(n)$ holds eventually if S_n is true for all $n \in \mathbb{N}$ sufficiently large. The logic statement is the following. To be precise, $\exists N \in \mathbb{N}$ such that $S(n)$ is true if $n \geq \mathbb{N}$. Then we say that $S(n)$ holds frequently if $S(n)$ is true for infinitely many n's. Also we describe, to be precise, $\forall N \in \mathbb{N}\ \exists n \geq \mathbb{N}$ such that $S(n)$ is true.

With this in mind, we can discuss the major definitions in this section. Notate a map from natural number to a metric space, that is, $f : \mathbb{N} \to X$.

Definition 3.1. Subsequence of $\{f_n\}_n$.

$$n : \mathbb{N} \to \mathbb{N}\ (\textit{injective and increasing})$$

such that

$$g_k = f_n(k) = f_{nk},\ \textit{which means } g(k) = f(n(k))$$

Example 3.2. For example, we can consider the following random sequence $\{1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, ...\}$ and we can extract a subsequence $\{1, \frac{1}{3}, \frac{1}{5}, ...\}$ or a subsequence $\{-\frac{1}{2}, -\frac{1}{4}, ...\}$. The first subsequence is defined by $x_n < \frac{1}{3}$ eventually if $n \geq 4 \Rightarrow x_n < \frac{1}{3}$. The second subsequence is defined by $x_n < 0$ frequently if for any $N \in \mathbb{N}$ and for any $n \in \mathbb{N}$ or $n \in \mathbb{N} + 1$, we have $x_n < 0$.

□

From the logical reasoning, we can also negate the statement above. We say that

$$\begin{aligned} Negate: &\quad S(n) \textit{ holds eventually} \\ \Rightarrow &\quad \textit{It is not the case that } S(n) \textit{ holds eventually} \\ \Rightarrow &\quad !S(n) \textit{ holds eventually},\ ! \textit{ means complement} \end{aligned}$$

Remark 3.3. For an example, consider the sequence $\{x_{n_{n \geq 1}}\} \subseteq \mathbb{N}$. We say either "$\{x_n\}$ is eventually even" or "$\{x_n\}$ is frequently odd".

□

Remark 3.4. Please notice that textbook offers an alternative but more complicated notation-wise definition for subsequences, which is described as the following [10].

Definition 3.5. Given a sequence $\{p_n\}$, consider a sequence $\{n_k\}$ of a positive integers, such that $n_1 < n_2 < n_3 < ....$ Then the sequence $\{p_{ni}\}$ is called a subsequence of $\{p_n\}$. If $\{p_{ni}\}$ converges, its limit is called a subsequential limit of $\{p_n\}$.

It is clear that $\{p_n\}$ converges to p if and only if subsequence of $\{p_n\}$ converges to p.

□

We consider (X, d) to be a metric space throughout this section. We define the concept of convergence in the following.

Definition 3.6. For $\{P_n\}_{n\geq 1} \subseteq X$, we say $\{P_n\}_{n\geq 1}$ converges to $p \in X$ if

$$\forall \epsilon > 0, \exists N \in \mathbb{N}\ s.t.\ d(P_n, P) < \epsilon$$

which can be interpreted $\forall n \geq N$ we have $P_n \in B_\epsilon(p)$. This is to say that $\forall \epsilon > 0$, we have $P_n \in B_\epsilon(p)$ eventually.

Remark 3.7. This is the same notation as in textbook [10]. We can consider a ball with center p and radius ϵ. As ϵ, the radius, decreases (no matter how small), the area of the ball is getting smaller. This would excludes some of points which originally included in the ball. However, it would be the case that there always exists some points within the radius of $\epsilon > 0$ (no matter how small it is).

□

We can also take the negation of the statements in the definition.

$$\begin{aligned} Negate: \quad & \text{“}\forall \epsilon > 0\ s.t.\ p_n \in B_\epsilon(p)\ eventually'' \\ \Rightarrow \quad & \exists \epsilon > 0,\ P_n \notin B_\epsilon(p)\ frequently \\ \Leftrightarrow \quad & \exists \epsilon > 0\ and\ subsequence\ \{p_{nk}\}_{k\geq 1} : P_{nk} \notin B_\epsilon(p)\ \forall k \geq \mathbb{N} \end{aligned}$$

Remark 3.8. We notate the following $lim_{n\to\infty} p_n = p$ or $p_n \to p$ as $n \to p$.

□

Proposition 3.9. *We have three propositions coming with the definition of convergence. Let $\{p_n\}$ be a sequence in a metric space X.*

1. *Uniqueness. If $p \in X$, $p' \in X$, and if $\{p_n\}$ converges to p and to p', then $p' = p$.*
2. *Subsequence Converges Too. If $p_n \to p$ as $n \to \infty$, then for any subsequence: $p_{nk} \to p$ as $k \to \infty$.*
3. *Stronger Convergence. If $\{P_n\}_{n\geq 1}$ such that any subsequence, which is a function of a subsequence (i.e. sub-subsequence) converges to p, then $p_n \to p$ as $n \to \infty$.*

Proof: Here we prove all three propositions, as shown in the following.

1. Uniqueness. We prove by contradiction. Assume otherwise that we have $p \neq p'$. This means that $d(p, p') = d > 0$ for some distance d. We can choose a smaller ϵ than d, so simply let $\epsilon = \frac{d}{2}$. Then, going back to the definition, we have $p_n \to p$ as $n \to \infty$, which implies that $\exists N \in \mathbb{N}$ s.t. $d(p_n, p) < \epsilon$ for some $n \in N$. However, we also have $p_n \to p'$ as $n \to \infty$, which implies that $\exists N' \in \mathbb{N}$ s.t. $d(p_n, p') < \epsilon$ for some $n \geq N'$.

   What would happen is the following: consider some n to be the larger one of N and N', that is, $n \geq max(N, N')$. Then, by Triangle Inequality, we have

   $$d(p, p') \leq d(dp, p_n) + d(p + n, p')$$

while $d(p, p') = 2\epsilon$, $d(p, p_n) < \epsilon$, and $d(p_n, p') < \epsilon$, we have

$$2\epsilon < 2\epsilon$$

which leads to a contradiction.

(2) Subsequence Converges Too. Assume we have left side of the argument, which states $p_n \to p$ as $n \to \infty$, and let $\{p_{nk}\}_{k\geq 1}$ to be a subsequence of p_n.

By definition, we have $\forall \epsilon > 0$, $\exists n \in \mathbb{N}$ s.t. $d(p_n, p) < \epsilon$, if we have $n \geq N \in \mathbb{N}$.

For the subsequence, notated as p_{nk} in the assumption, if $n_k \geq N \in \mathbb{N}$, then $d(p_{nk}, p) > \epsilon$. (Notice that n_k is injective and increasing, which is a property coming from its own definition, i.e. defined already.)

$$\Rightarrow \exists L \in \mathbb{N},\ s.t.\ \exists k \geq L,\ n_k \geq N \in \mathbb{N}.$$

Note that in the lecture the notation is a little confusing. It should be an element n_k which is greater or equal to another big element N such that N is in the set $\mathbb{N}$. Please correct me if you think this is wrong.

(3) Stronger Convergence. We prove this proposition by contradiction. Assume it is not the case that $p_n \to p$ as $n \to \infty$. Then there exists $\epsilon > 0$ such that for $k \in \mathbb{K}$ there is $n_k \geq k$ such that $d(p_{n_k}, p) > \epsilon$ (i.e. you want to find a subsequence of the subsequence). Then there exists a subsequence $\{p_{n_k}\}$ such that $d(p_{n_k}, p) > \epsilon$.

By induction, we can check the case $n_1 = 1$.

Now for $N_l = max(l, n1, ..., n_{l-1})$, there exists $m \geq N_l$ such that $d(p_n, p) \geq \epsilon$. Let $n_l = n$, we have a subsequence does not have a further subsequence that converges to p, which leads to a contradiction.

Q.E.D.

This completes the proof of the three properties in the proposition above.

We can now discuss the subsequence in $\mathbb{R}^k$ with $k \geq 1$. Consider $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1} \subseteq \mathbb{R}$. We have the following proposition.

Proposition 3.10. *Consider $X_n \to x$ and $Y_n \to y$ as $n \to \infty$, then we have*

(1) Monotonicity. $X_n \leq Y_n \Rightarrow x \leq y$.

(2) Lindear Combination. $\forall \alpha, \beta \in \mathbb{R} : \alpha X_n + \beta Y_n \to \alpha x + \beta y$ as $n \to \infty$.

(3) Product. $X_n Y_n \to xy$ as $n \to \infty$.

(4) Inverse. $\frac{1}{X_n} \to \frac{1}{x}$ as $n \to \infty$ (x and $x_n \neq 0$).

Proof:

(1) Monotonicity. $X_n \leq Y_n \Rightarrow x \leq y$.

(2) Linear Combination. $\forall \alpha, \beta \in \mathbb{R} : \alpha X_n + \beta Y_n \to \alpha X + \beta Y$ as $n \to \infty$.

(3) Product. Let $\epsilon > 0$ to be arbitrary and consider $X_n Y_n - xy = (X_n - x)y_n + x(Y_n - y)$. Take absolute values and we have the following inequalities.

$$|X_n Y_n - xy| \leq |X_n - x||Y_n| + |x||Y_n - y|$$

From $X_n \to x$ as $n \to \infty$, we have $\exists N \in \mathbb{N}$ such that $|X_n - x| < \epsilon$ if $n \geq N$. From $Y_n \to y$ as $n \to \infty$.
Let $m = max(N, N')$ and R be an upper bound of $\{|Y_n|\} \cup \{|x|\}$. This implies that $|X_n Y_n - xy| \leq \epsilon R + R\epsilon = 2\epsilon R$, which leads to a contradiction.

(4) Inverse. $\frac{1}{\mathrm{X}_n} \to \frac{1}{x}$ as $n \to \infty$ (x and $x_n \neq 0$).

Q.E.D.

Remark 3.11. It is also a reasonable argument to claim that "$Y_n \to y$ as $n \to \infty \Rightarrow \{Y_n\}$ is bounded. This is given as part (c) in the following theorem and the proof for the theorem is also provided as the following [10].

□

Proposition 3.12. *For $\mathbb{R}^k : k \geq 1$ arbitrary, we have $\mathbb{X}_n = (a_{1,n}, ..., a_{k,n})$, and then we have*

(1) $\mathbb{X}_n \to \mathbb{X} = (\alpha_1, ..., \alpha_n)$ *as* $n \to \infty$.
$\Leftrightarrow \alpha_{i,n} \to \alpha_n$ *as* $n \to \infty$, $\forall i = 1, ..., k$.

(2) $[\mathbb{X}_n \to \mathbb{X}$ *as* $n \to \infty$ *and* $\mathbb{Y}_n \to \mathbb{Y}$ *as* $n \to \infty]$
$\Rightarrow [\mathbb{X}_n \mathbb{Y}_n \to \mathbb{X}\mathbb{Y}]$.

(3) $[\mathbb{X}_n \to \mathbb{X}$ *as* $n \to \infty$ *and* $\mathbb{Y}_n \to \mathbb{Y}$ *as* $n \to \infty]$
$\Rightarrow [\alpha\mathbb{X}_n + \beta\mathbb{Y} \to \alpha\mathbb{X} + \beta\mathbb{Y}$ *as* $n \to \infty\ \forall \alpha, \beta \in \mathbb{R}]$

Theorem 3.13. *Let $\{p_n\}$ be a sequence in a metric space X.*

(a) $\{p_n\}$ converges to $p \in X$ if and only if every neighborhood of p contains p_n for all but finitely many n.

(b) If $p \in X$, $p' \in X$, and if $\{p_n\}$ converges to p and to p', then $p' = p$. (This is proved earlier.)

(c) If $\{p_n\}$ converges, then $\{p_n\}$ is bounded.

(d) If $E \subset X$ and if p is a limit point of E, then there is a sequence $\{p_n\}$ in E such that $p = lim_{n\to\infty} p_n$.

Proof:

(a) Suppose $p_n \to p$ and let V be a neighborhood of p. For some $\epsilon > 0$, the conditions $d(q, p) < \epsilon$, $q \in X$ imply $q \in V$. Corresponding to this ϵ, there exists N such that $n \geq N$ implies $d(p_n, p) < \epsilon$. Thus $n \geq N$ implies that $p_n \in V$.

(b) Proof is omitted since it is presented above.

(c) Suppose $p_n \to p$ as $n \to \infty$. That means there is an integer N such that $n > N$ implies $d(p_n, p) < 1$ (In class, we discussed that the "1" is not necessary, but referring to the textbook let us follow the authority here. After consulting with instructor, we have the following. Using $\epsilon = 1$ you get that p_n is within distance one from p eventually (let us say for all $n > N$). If $r = max(1, d(p1, p), ..., d(pN, p))$ then p_n is contained in $B_r(y)$ and it is hence bounded. You can use any epsilon you wish and the proof still works. Another question was: Is 1 really necessary inside the maximum? Can we just take $r = max(d(p1, p), ..., d(pN, p))$ instead? This would be fine if the N is optimal such that $d(p_N, p)$ is greater or equal than 1 and $d(p_n, p) < 1$ for $n > N$. We are not assuming that N is optimal in this sense and that's why 1 is included in the max.). Simple choose

$$r = max\{1, d(p_1, p), ..., d(p_N, p)\}.$$

Then $d(p_n, p) \leq r$ for $n = 1, 2, 3, ...$, which completes the proof.

(d) For each positive integer n, there is a point $p_n \in E$ such that $e(p_n, p) < \frac{1}{n}$. Given $\epsilon > 0$, choose N so that $N_\epsilon > 1$. If $n > N$, it follows that $d(p_n, p) < \epsilon$. Hence $p_n \to p$.

Q.E.D.

Theorem 3.14. *This theorem states some interesting facts about sequence, subsequence, and boundedness.*

(a) If $\{p_n\}$ is a sequence in a compact metric space X, then some subsequence of $\{p_n\}$ converges to a point of X.

(b) Every bounded sequence in $\mathbb{R}^k$ contains a convergent subsequence.

Proof:

(a) Let E be the range of $\{p_n\}$. If E is finite, there is a $p \in E$ and a sequence $\{n_i\}$ with $n_1 < n_2 < n_3 < ...$, such that

$$p_{n_1} = p_{n_2} = ... = p$$

The subsequence $\{p_{n_i}\}$ so obtained converges evidently to p.

If E is infinite, Theorem 2.37 [10] shows that E has a limit point $p \in X$. Choose n_1 so that $d(p, p_{n_1}) < 1$. Having chosen $n_1, , n_{i-1}$, we see from Theorem 2.20 [10] that there is an integer $n_1 > n_{i-1}$ such that $d)p, p_{n-1}) < 1/i$.

(b) This follows from (a), since Theorem 2.41 [10] implies that every bounded subset of R^k lies in a compact subset of R^k.

Q.E.D.

Corollary 3.15. *In metric space $X = \mathbb{R}^k$, if $\{P_n\}_{n \geq 1}$ is bounded, then $\{P_n\}_{n \geq 1} \subseteq \bar{B}_R(0)$ which is compact by Heine-Borel Theorem. Then P_n has convergent subsequence.*

A reciprocal of the statement above is the following. In a metric space $K \subset R^k$, any sequence in k has a convergence sub-sequence. This implies that K is compact.

Proof: Consider the statement "K is compact $\Rightarrow$ "K is closed and bounded. We are showing the proof of "K is compact $\Rightarrow$ K is closed.

We prove by contradiction. Assume $p \in K$ and $p \notin K$, i.e. limit point p is not in K.

$\forall n \in \mathbb{N}$, we have $B_{\frac{1}{n}}(p) \cap K \setminus \{p\} \neq 0$. Then we take $p_n \in B_{\frac{1}{n}}(p) \cap K$.

This implies that $p_n \to p$ as $n \to \infty$. Then by uniqueness limit point $p \in K$, which leads to a contradiction.

The other part of statement is the same procedure. "K is compact" $\Rightarrow$ "K is bounded".

Q.E.D.

Definition 3.16. $\{P_n\}_{n\geq 1}$ is cauchy if $\forall N \in \mathbb{N}: \ n, m \geq N$ such that $d(p_n, p_m) < \epsilon$.

Remark 3.17. An alternative definition can be said as the following

$$\forall \epsilon > 0, \ \exists N \in \mathbb{N} : p_n \in B_\epsilon(p_m) \ \textit{eventually} \ \forall m \geq \mathbb{N}$$

□

Remark 3.18. Referred textbook offers a similar yet wordy description. A sequence $\{p_n\}$ in a metric space X is said to be Cauchy sequence if for every $\epsilon > 0$ there is an integer N such that $d(p_n, p_m) < \epsilon$ if $n \geq N$ and $m \geq N$.

□

Just for notation purpose, we offer the definition of diameter of E by textbook [10].

Definition 3.19. Let E be a nonempty subset of a metric space X, and let S be the set of all real numbers of the form $d(p, q)$, with $p \in E$ and $q \in E$. The *sup* of S is called the *diameter* of E.

If $\{p_n\}$ is a sequence in X and if E_N consists of the points p_N, p_{N+1}, p_{N+2}, ..., it is clear from the two preceding definitions that $\{p_n\}$ is a Cauchy sequence if and only if

$$lim_{N\to\infty} diam \ E_N = 0.$$

Proposition 3.20. *A convergence sequence is a Cauchy Sequence.*

Proof: Let $p_n \to$p as $n \to \infty$. Then for $\epsilon > 0$ no matter how small, we have $\exists N \in \mathbb{N} : \ d(p_n, p) < \epsilon/2$. This implies that if $m, n \geq N$, then $d(p_n, p_m) \leq d(p_n, p) + d(p, p_m) < \epsilon$.

Q.E.D.

Remark 3.21. Notice that the opposite direction is may or may not be true. Consider $X = \mathbb{R} \setminus \{0\}$, $p_n = \frac{1}{n}$ and p_n is Cauchy and p_n does not converge.

□

Definition 3.22. A metric space in which every Cauchy sequence converges is said to be complete.

Remark 3.23. Theorem 3.11 says that all compact metric spaces and all Euclidean spaces are complete [10]. It also implies that every closed subset E of a complete metric space X is complete. That is, every Cauchy sequence in E is a Cauchy sequence in X, hence it converges to some $p \in X$, and actually $p \in E$ since E is closed.

An example can be the following. Consider a metric space which is not complete, this metric space is the space of all rational numbers, with $d(x, y) = |x - y|$.

□

Remark 3.24. Theorem 3.2 (c) and example (d) of Definition 3.1 show that convergent sequences are bounded, but that bounded sequences in R^k need not converge [10].

□

Proposition 3.25. *This proposition states the following:*

(1) If X is compact, then X is complete.

(2) $\mathbb{R}^k$ is complete.

Proof:

(1) Let $\{p_n\}_{n\geq 1} \subseteq X$. This implies that there is a subsequence $p_{n_k} \to p \in K$, since K is compact.

Assume $\{p_n\}$ is Cauchy and our goal is to get $p_n \to p$. Let $\epsilon > 0$ no matter how small.

Since p_n is Cauchy, $M \in \mathbb{N}$: $m, n \geq M$ implies that $d(p_n, p_m) < \epsilon/2$.

Since $p_{n_k} \to p$, we have $k \in \mathbb{N}$ such that $k \geq K$, which implies $d(p_{n_k}) < \epsilon/2$.

Then we have $n \geq M$, which implies that $d(p_n, p_m) \leq d(p_n, p_{n_k}) + d(p_{n_k}, p) < \epsilon$.

(2) Consider $k = 1$. Then we claim that $\{p_n\}$ is Cauchy implies that $\{p_n\}$ is bounded.

$\exists N \in \mathbb{N}$ such that $d(p_n, p_m) < 1$, $\forall n, m \geq N$.

This implies that $p_m \in B_1(p_N)$, $\forall m, n \in \mathbb{N}$.

Consider $R = \max(d(p_1, p_n), ..., d(p_{n_k}, p_n))$, then we have $p_m \in B_R(p_N)$, $\forall m \geq 1$.

Since $\{p_n\}$ is Cauchy, this sequence is bounded. Then by the theorem of subsequence we had earlier, $p_{n_k} \to p \in \mathbb{R}^k$.

Q.E.D.

Now we introduce the famous Bolzano-Weierstrass Theorem.

Theorem 3.26. (Bolzano-Weierstrass Theorem). *Let $(a_n)_{n=0}^{\infty}$ be a bounded sequence (i.e., there exists a real number $M > 0$ such that $|a_n| \leq M$ for all $n \in N$). Then there is at least one subsequence of $(a_n)_{n=0}^{\infty}$ which converges.*

Proof: Let L be the limit superior of the sequence $(a_n)_{n=0}^{\infty}$. Since we have $-M \leq a_n \leq M$ for all natural numbers n, it follows from the comparison principle (Lemma 6.4.13) that $-M \leq L \leq M$. In particular, L is a real number (not $+\infty$ or $-\infty$). By Proposition 6.4.12(e), L is thus a limit point of $(a_n)_{n=0}^{\infty}$. Thus by Proposition 6.6.6, there exists a subsequence of $(a_n)_{n=0}^{\infty}$ which converges (in fact, it converges to L).

Q.E.D.

Remark 3.27. To understand the proof of the theorem above, we need to notate the lemma and proposition in textbook [11].

Lemma 3.28. Comparison Principle (Lemma 6.4.13 from textbook [11]). *Suppose that $(a_n)_{n=m}^{\infty}$ and $(b_n)_{n=m}^{\infty}$ are two sequences of real numbers s.t. $a_n \le b_n$ for all $n \ge m$. Then we have the inequalities*

$$\sup(a_n)_{n=m}^{\infty} \le \sup(b_n)_{n=m}^{\infty}$$

$$\inf(a_n)_{n=m}^{\infty} \le \inf(b_n)_{n=m}^{\infty}$$

$$\lim \sup_{n\to\infty} a_n \le \lim \mathit{sup}_{n\to\infty} b_n$$

$$\lim \inf_{n\to\infty} a_n \le \mathit{lim} \inf_{n\to\infty} b_n$$

Proposition 3.29. Proposition 6.4.12 (e) from referred textbook [11]. *Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers, let L^+ be the limit superior of this sequence, and let L^- be the limit inferior of this sequence (thus both L^+ and L^- are extended real numbers).*

(e) If L^+ is finite, then it is a limit point of $(a_n)_{n=m}^{\infty}$. Similarly, if L^- is finite, then it is a limit point of $(a_n)_{n=m}^{\infty}$.

Proposition 3.30. Subsequences related to limit points (Proposition 6.6.6 from textbook [11]). *Let $(a_n)_{n=0}^{\infty}$ be a sequence of real numbers, and let L be a real number. Then the following two statements are logically equivalent.*

(a) L is a limt point of $(a_n)_{n=0}^{\infty}$.

(b) There exists a subsequence of $(a_n)_{n=0}^{\infty}$ which converges to L.

□

3.2 Series

Go back to Table of Contents. Please click TOC

The concept of convergence is really important in the study of real numbers. From textbook [10], recall the following definition from sequence.

Definition 3.31. A sequence $\{p_n\}$ in a metric space X is said to converge if there is a point $p \in X$ with the following property: for every $\epsilon > 0$ there is an integer N such that $n \ge N$ implies that $d(p_n, p) < \epsilon$.

Then we can start to introduce series.

Definition 3.32. Given a sequence $\{a_n\}$, we use the notation

$$\sum_{n=p}^{q} a_n, \ p \le q$$

to denote the sum $a_p + a_{p-1} + \cdots + a_q$. With $\{a_n\}$ we associate a sequence $\{S_n\}$, where $s_n = \sum_{k=1}^{n} a_k$. For $\{s_n\}$ we also use the symbolic expression $a_1 + a_2 + \cdots + a_n$ or more concisely, $\sum_{n=1}^{\infty} a_n$, which is called infinite series, or just a series. The numbers s_n are called the partial sums of the series. If $\{s_n\}$ converges to s, we say that the series converges, and write $\sum_{n=1}^{\infty} a_n = s$.

We start by discussing the types of series converge. Then we discuss how to compute the examples, which leads to Fundamental Theorem of Calculus. Lastly, we introduce convergence/divergence tests.

(1) Types of Convergence.

We have conditional and absolute convergence. Need to refer to text.[11] For absolute convergence, we have rearrangement, product, and convolution. Product states the following $\sum_{n=0}^{\infty} a_n$ and $\sum_{a=0}^{\infty} b_n$ convergences absolutely. Convolution states for $(a \star b)_n = \sum_{a_k}^{n} a_k b_{n-k}$, then $\sum_{n=0}^{\infty}(a+b)_n = (\sum_{n=0}^{\infty} a_n)(\sum_{n=0}^{\infty} b_n)$.
From text [11], we have the following definition and test.

Definition 3.33. (Absolute convergence). Let $\sum_{n=m}^{n} a_n$ be a formal series of real numbers. We say that this sereis is aboslutely convergent if and only if the series $\sum_{n=m}^{\infty} |a_n|$ is convergent.

Proposition 3.34. *(Absolute convergence test). Let* $|\sum_{n=m}^{\infty} a_n$ *be a formal series of real numbers. If this series is absolutely convergent, then it is also conditionally convergent. Furthermore, in this case we have the triangle inequality*

$$\left|\sum_{n=m}^{\infty} a_n\right| \leq \sum_{n=m}^{\infty} |a_n|$$

Example 3.35. For example, convolution has the following properties: linearity, commutative, associative, and unit element.

□

Linear: $\forall \alpha, \beta \in \mathbb{R}, [(\alpha a + \beta b) \star c]_n = \alpha(a \star c) + \beta(b \star c)$.
Commutative: $(a \star b)_n = (b \star a)_n$.
Associative: $(a \star b) \star c = a \star (b \star c)$.
Unit:

$$s_n = \begin{cases} 1, & if\ n = 0 \\ 0, & elsewhere \end{cases} \Rightarrow \delta \star a = a \star \delta = a$$

Fubini: $\{a_{mn}\}$ sequence of complex numbers says the following: "For all $n \in \mathbb{N}$, there is $A_n = \sum_{m=1}^{\infty} a_{mn}$ converges absolutely and $\sum_{n=1}^{\infty} A_n$ converges absolutely." $\Rightarrow$

$$\sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{mn} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{mn},$$

which is as illustrated as the matrix below,

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$$

Example 3.36. For example consider $p \in (-1, 1)$, then we have

$$\begin{bmatrix} 1 & p & p^2 & p^3 & ... \\ 0 & p & p^2 & p^3 & ... \\ 0 & 0 & p^2 & p^3 & ... \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{bmatrix}$$

□

Then we count the sum of each columns. Column 1 has 1. Column 2 has $2p$. Column 3 has $3p^2$, etc. We can also count the sum of rows. Row 1 has $\frac{1}{1-P}$. Row 2 has $\frac{p}{1-p}$. Row 3 has $\frac{p^2}{1-p}$.
Then we have

$$\sum_{n=1}^{n} np^{n-1} = \sum_{n=1}^{\infty} \frac{p^{n-1}}{1-p} = \frac{1}{(1-p)^2}.$$

(2) How to compute them? Recall the telescopic theorem, (which is also similar to Fundamental Theorem of Calculus, i.e. $\int_a^b f' = f|_a^b$). We have

$$(\triangle a)_n = a_{n+1} - a_n \Rightarrow \sum_{n=p}^{q} (\triangle a)_n = a_{n+1} - a_1$$

Example 3.37. For example, consider $a_n = p$. We have

$$\Rightarrow (\triangle a_n) = p^{n+1} - p^n = p^n(1-p)$$
$$\Rightarrow \sum_{n=0}^{N} p^n(1-p) = \frac{p^{n+1}-1}{p-1} \to \frac{1}{1-p} \ as \ N \to \infty \ if \ p \in (-1, 1)$$

□

Example 3.38. For another example, consider

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = lim_{N\to\infty} \sum_{n+1}^{N} (\frac{1}{n}) - \frac{1}{n+1}) = lim_{N\to\infty}(1 - \frac{1}{N+1}) = 1$$

This is summation by parts, (similar to integration by parts, $\int udv = uv - \int vdu$), and we have the following

$$\sum_{n=p}^{q} a_{n+1}(\triangle b)_n = (a_q + b_{q+1}) - \sum_{n=p}^{q} b_n(\triangle a)_n$$

The idea is to combine them by the product rule, and also follow Fundamental Theorem of Calculus. That is

$$\begin{aligned} \triangle(ab)_n &= (\triangle a_{n+1})(\triangle b)_n + (\triangle a)_n b_n \\ &= a_{n+1}b_{n+1} - a_n b_n \end{aligned}$$

As an exercise, solve

$$\sum_{n=1}^{\infty} np^n = ?$$

by using summation by parts, also by considering $\triangle n = 1$.

□

(3) Convergence/Divergence Test

  (a) Comparison Tests. In this section we have two parts.
  Part I. $|a_n| < b_n$ eventually and $\sum b_n$ convergence $\Rightarrow \sum a_n$ converges absolutely.
  Part II. $|a_n| = b_n > 0$ eventually and $\sum b_n$ divergence $\Rightarrow \sum |a_n|$ divergence.

  (b) Alternative Series Test. This test states the following:
  If $a_n \geq 0$, $a_n \to 0$ as $n \to \infty$,
  then we have $\sum(-1)^n a_n$ convergence.
  Remark 3.39. Note that $\sum b_n \to \infty \Rightarrow |a_n| \to \infty$ as well.

□

Consider the following example about Alternative Harmonic Series.

$$\sum_{n=1}^{n}(-1)^n \frac{1}{n} = -ln(2)$$

Figure 9: Graphic view for a standard Alternative Harmonic Series.

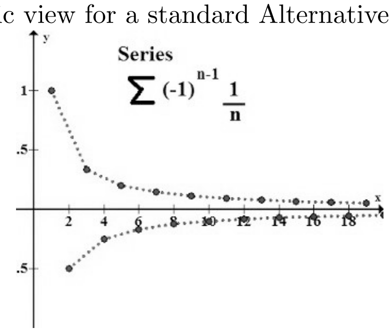
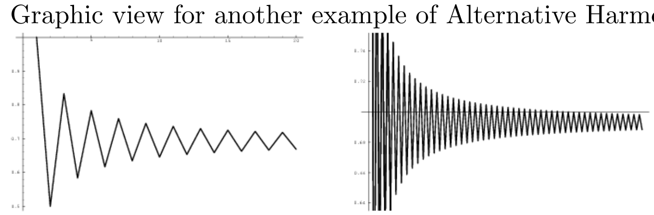


Figure 10: Graphic view for another example of Alternative Harmonic Series.



We have the following strategy, just as illustrated in the graph. That is, consider s_{2n+1} increasing and s_{2n} decreasing. Then we can prove that $s_{2n+1} < 2_{2n}$. Then $s \leq s_{2n+1} \leq s_{2n} \leq s_2$.
Remark 3.40. Notice that each term in the series is decreasing and bounded below.

□

For another example about harmonic series: $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$, that is, $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots$ The first term, $1 \geq 1$, the second term, $\frac{1}{2} \geq \frac{1}{2}$, the third term, $\frac{1}{3} \geq \frac{1}{4}$ and so on. We observe the first term is greater or equal to 1, the second term is greater or equal to $\frac{1}{2}$. The third to sixth term are summed up to be $\frac{1}{2}$ and so on. We would have a series greater or equal to $1 + \frac{1}{2} + \frac{1}{2} + \cdots = \infty$. For an example, we can consider p-series.

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = \begin{cases} convergence & , \quad if\ p > 1 \\ divergence & , \quad if\ p \leq 1 \end{cases}$$

If $p \leq 1$, $\frac{1}{n^p} \geq \frac{1}{1/n}$. That is,

$$1 + \frac{1}{2^p} + \frac{1}{3^p} + ...$$

and we apply the same method. Consider the first term to be less or equal 1. The second and third term to be less or equal $\frac{1}{2^p}$. Starting from the third term, there are four terms less or equal $\frac{1}{4^p}$. Then we have the first term equal to 1. The second and third term sum up to be less or equal to $\frac{2}{2^p}$. The third to sixth term to sum up be less or equal to $\frac{4}{4^p}$ and so on. This becomes

$$\sum_{k=0}^{\infty} \frac{2^k}{2^{kp}} = \sum_{k=0}^{\infty} 2^{k(1-p)} = \frac{1}{1 - p^{1-p}}$$

which is convergent.

The text [11] offers an alternative:

Proposition 3.41. *(Alternating series test). Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers which are non-negative and decreasing, thus $a_n \geq 0$ and $a_n \geq a_{n+1}$ for every $n \geq m$. Then the series $\sum_{n=m}^{\infty} (-1)^n a_n$ is convergent if and only if the sequence a_n converges to 0 as $n \to \infty$.*

(c) Comparison w/ Geometric Series. The motivation here is to understand that $\{a_n\}$ is a geometric series.
If (1) $\frac{a_{n+1}}{a_n} = r$, whch is common ratio, or (2) $(a_n)^{1/n}$, which is common ratio, then we have $a_n = r^n$.

(i) Ratio: $lim_{n\to\infty}\ sup\ \frac{|a_{n+1}|}{|a_n|} = r$
(i.1) $r \in [0, 1)$, $\Rightarrow \sum a_n$ convergence absolutely.
(i.2) $r > 1$, $\Rightarrow \sum a_n$ divergence. This infers that $lim_{n\to\infty}\ inf\ \frac{|a_{n+1}|}{|a_n|} > 1 \Rightarrow \sum a_n$ divergent.
(i.3) $r = 1$, there is no information. This is saying $lim_{n\to\infty}\ sup\ \frac{|a_{n+1}|}{a_n} = 1 \Rightarrow$ no information.
The idea is straightforward. For (i.1), we have $r < 1$, then $\frac{|a_{n+1}|}{|a_n|} < \frac{r+1}{2}$ eventually. This implies that $|a_n| \leq c(\frac{r+1}{2})^n$ eventually, while c is a constant. This is done by applying Comparison Test Part I. For (i,2), we assume $r > 1$, then we have $\frac{|a_{n+1}|}{a_n} > \frac{r+1}{2}$ frequently. This implies that compare $|a_n|$ with $c(\frac{r+1}{2})^n$ by using Comparison Test Part II.

(ii) Root: $lim_{n\to\infty}\ sup|a_n|^{1/n} = r$.
(ii.1) $r < 1 \Rightarrow \sum a_n$ which convergence absolutely.
(ii.2) Unif$(|a_n|)^{1/n} > 1 \Rightarrow \sum a_n$ divergence.

(ii.3) $lim_{n\to\infty}(|a_n|)^{1/n} = 1 \Rightarrow$ no information implied.
An exericse $(|a_n|)^{1/n} < \frac{r+1}{2}$ eventually $\Rightarrow |a_n| < c(\frac{r+1}{2})^n$ eventually.

Remark 3.42. Referred textbook has the following lemma for comparison principle [11].

Lemma 3.43. Comparison Principle. *Suppose that* $(a_n)_{n=m}^{\infty}$ *and* $(b_n)_{n=m}^{\infty}$ *are two sequences of real numbers such that* $a_n \leq b_n$ *for all* $n \geq m$. *Then we have the inequalities*

$$sup(a_n)_{n=m}^{\infty} \leq sup(b_n)_{n=m}^{\infty}$$

$$inf(a_n)_{n=m}^{\infty} \leq inf(b_n)_{n=m}^{\infty}$$

$$\lim_{n\to\infty} sup\; a_n \leq \lim_{n\to\infty} sup\; b_n$$

$$\lim_{n\to\infty} inf\; a_n \leq \lim_{n\to\infty} inf\; b_n$$

□

Now we can look at some examples. Consider $\{a_n\}_{n\geq 1} \subseteq \mathbb{R}$ (or $\mathbb{C}$). Consider $s_n = \sum_{k=1}^{n} a_k$ a partial sum.

For another example, consider geometric series: $a^{n+1} - b^{n+1} = (a - b)(a^n + a^{n-1}b + \cdots + ab^{n-1} + b)$. If $b = 1$, it implies $\frac{1-a^{n+1}}{1-a} = \sum_{k=1}^{n} a^k$. If $a \in (0,1)$, then $\sum_{k=1}^{\infty} a^k = \frac{1}{1-a}$. If $a \geq 1$, then $\sum_{k=1}^{\infty} a^k = \infty$.

Example 3.44. For example, consider harmonic series, $\sum_{i=1}^{\infty} \frac{1}{n} = \infty$.

□

Example 3.45. Also we can consider example, p-series, $\sum_{i=1}^{\infty} \frac{1}{n^p}$. We have

$$\sum_{i=1}^{\infty} \frac{1}{n^p} = \begin{cases} converge & , if\ p > 1 \\ diverge & , if\ p \leq 1 \end{cases}$$

□

A goal can be: absolute value implies commutative proposition, which leads to the following proposition.

Proposition 3.46. *Absolute convergence implies convergence.*

Proof:

$$|S_n - S_m| = |\sum_{k=n}^{m} a_k| \leq \sum_{k=n}^{m} |a_k| = |\bar{S}_m - \bar{S}_n|$$

$$\bar{S}_k = \sum_{l=1}^{k} |a_l|$$

For one that is small, it implies that the other is sufficiently large.

3.3 Rearrange

Go back to Table of Contents. Please click TOC

We start with a definition.

Definition 3.47. Consider $\sigma : \mathbb{N} \to \mathbb{N}$ a bijection, then $\sum a_{\sigma(n)}$ is a real of $\sum a_n$.

Theorem 3.48. *If $\sum a_n$ converges absolutely, then $\sum a_{\sigma(n)} = \sum a_n$ (also the commutative property).*

Lemma 3.49. $a_n \leq 0 \Rightarrow \sum a_{\sigma(n)} = \sum a_n$.

Proof: (of lemma) Let $\epsilon > 0$, and let $\sum a_n = s$, and we have $\exists N \in \mathbb{N}$ s.t. $|s - s_n| < \epsilon$, for $n \geq \mathbb{N}$. Take large neough such that $\{1, 2, ..., N\} \subseteq \{\sigma(1), \sigma(p)\}$. Let $m \geq p$, compute $|s - s'_n|$ while we define $s'_n = \sum_{n=1}^{n} a_{\sigma(k)}$. Then we have

$$|s - s'_n| = |a_1 + \cdots + a_N + \cdots - (a_{\sigma(1)} + a_{\sigma(2)} + \cdots + a_{\sigma(p)} + \cdots + a_{\sigma(n)})|$$

$$\leq \sum_{n=N+1}^{\infty} a_n| = |s - s_N| < \epsilon$$

Proof of Lemma. Q.E.D.

Proof: (of theorem)

$$\begin{array}{rcl} \sum |a_n| \text{ converges} & \Rightarrow & \sum |a_{\sigma(n)}| \text{ converges} \\ & \Leftrightarrow & \sum a_{\sigma(n)} \text{ converges absolutely} \\ & \Rightarrow & \sum a_{\sigma(n)} \text{ converges} \end{array}$$

Proof of Theorem. Q.E.D.

Now we discuss the Application Product.

Take $(a + b + c)(\alpha + \beta + \gamma)$, which will be 9 terms. We consider the matrix form

-	α	β	γ
a	...		
b	...		
c	...		

Then we take $\sum a_n$ and $\sum b_n$ and we would have the expanded form of the equation above.

Definition 3.50. For $\{a_n\}$, $\{b_n\} \subseteq \mathbb{R}$, we have $(a \star b)_n = \sum_{k=1}^{n} a_k b_{n-k}$.

For an exercise, consider $a_n = 1$ and $b_n = n$. Then compute $(a \star b)_n$.

$$\begin{array}{rcl} (a \star b)_n & = & a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_n b_0 \\ & = & \frac{n(n+1)}{2} \end{array}$$

I need to redo this part since I haven't figured out why.

Theorem 3.51. *If $\sum a_n$ and $\sum b_n$ converge absolutely, then $\sum (a \star b)_n = (\sum a_n)(\sum b_n)$.*

Next, is the Convergence/Divergence Test, which has four parts to be discussed.

(1) Cauchy. $\sum a_n$ is convergent $\Leftrightarrow$ $\{S_n\}$ is Cauchy. That is, $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$, s.t. $|S_n - S_m| < \epsilon$, $\forall n, m \geq N$. Notice that $|S_n - S_m|$ is the term $\sum_{k=n}^{m} a_k$.

(2) Divergence Test. "$lim_{n \to \infty} a_n \neq 0$ does not exist" $\Rightarrow$ "$\sum a_n$ diverges".

(3) Bounded. $\bar{S}_n = \sum_{k=1}^{n} |a_k|$ increasing. That is, $\sum a_n$ converge absolutley $\Leftrightarrow$ $\{\bar{S}_n\}$ bounded above. Moreover, $\sum |a_n| = sup \bar{S}_n$.

(4) Comparison (Monotonicity). We discuss two scenarios:
   (4.a) "$|a_n| \leq c_n$ eventually" $\Rightarrow$ "$\sum a_n$ converge if $\sum c_n$ converge".
   (4.b) "$|a_n| \geq c_n > 0$" $\Rightarrow$ "$\sum |a_n|$ diverge if $\sum c_n$ diverge".

We need to continue with Power Series, which finishes up Chapter 3.

4 §Continuity§

Go back to Table of Contents. Please click TOC

4.1 Continuity

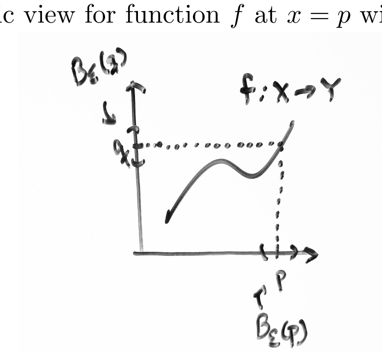
Go back to Table of Contents. Please click TOC

In this section, we discuss the following topics: limit, continuity, continuity and compactness, and uniform continuity.

(1) Limit

Consider the metric space (X, dx), (Y, dy). Consider function $f : X \to Y$. That is, consider $\lim_{x\to p} f(x) = q$ is $f(x)$ is close to q if x is sufficiently closely to p.

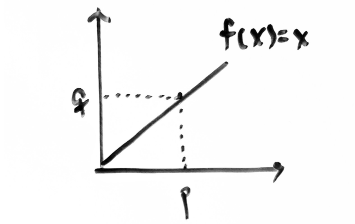
Figure 11: Graphic view for function f at $x = p$ with $\epsilon > 0$ and $B_\epsilon(q)$.



Definition 4.1. "$\lim_{x\to p} f(x) = q$ or $f(x) \to q$ *as* $x \to p$" if "$\forall \epsilon > 0$, $\exists \delta > 0$, such that $f(x) \in B_\epsilon(q)$ for all $x \in B_\delta(p) \setminus \{p\}$."

Example 4.2. For example, consider X and Y in $\mathbb{R}$, $lim_{x\to p} x = p$, $\forall \epsilon > 0$. That is, $\forall \epsilon > 0$, take $\delta = \epsilon$, then "$x \in B_\epsilon(p)$" if "$x \in B_\delta(p) \setminus \{p\}$ while interval is $(p - \epsilon, p + \epsilon) \setminus \{p\}$", which is illustrated in the following graph.

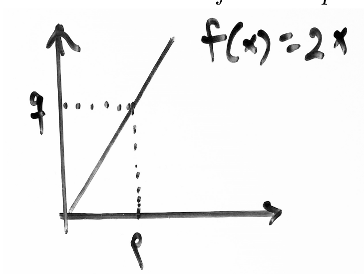
Figure 12: Graphic view for function f at $x = p$ with $\epsilon > 0$ and $B_\epsilon(q)$.



$\square$

Consider an alternative but similar example, $f(x) = 2x$. In this case, we can no longer take $\delta = \epsilon$, instead we take $\delta = \epsilon/2$. The rest is the same procedure.

Figure 13: Graphic view for function f at $x = p$ with $\epsilon > 0$ and $B_\epsilon(q)$.



We can also state the negation of the statement above. That is, "$f(x) \not\to q$ *as* $n \to \infty$", $\Leftrightarrow$ "$\exists \epsilon > 0 : \forall \delta > 0$, there exists some $x \in B_\delta(p) \setminus \{p\}$ such that $f(x) \notin B_\epsilon(q)$".

Remark 4.3. The procedure is always the following: choose a point, no matter how small, it will always fall outside of target.

$\square$

Proposition 4.4. *We have the following statements:*
"$f(x) \to q$ as $x \to p$" if and only if "$f(p_n) \to q$ as $n \to \infty$ for every sequence $p_n \to p$ as $n \to \infty$ and $p_n \neq p$."

Proof:
($\Rightarrow$) $\forall \epsilon > 0$, $\exists \delta > 0$, such that $f(x) \in B_\epsilon(q)$ if $x \in B_\epsilon(p) \setminus \{p\}$. If $p_n \to p$ as $n \to \infty$, then $p_n \in B_\delta(p)$ eventually.
$\Rightarrow f(p_n) \in B_\epsilon(q)$ eventually.
$\Rightarrow f(p_n) \to q$.
($\Leftarrow$) By condition, $f(x) \not\to q$ as $x \to p$. $\exists \epsilon >$), s.t. $\delta = \frac{1}{n}$, $\exists p_n \in B_{1/n}(p) \setminus \{p\}$ s.t. $f(p_n) \notin B_\epsilon(q)$.
$\Rightarrow p_n \neq q$ and $p_n \to p$ as $n \to \infty$. But since $f(p_n) \not\to q$ as $n \to \infty$, which is a contradiction.

Q.E.D.

Now we move on to discuss limsup and liminf.

Definition 4.5. Consider $\{a_n\}_{n\geq 1}\} \subseteq \mathbb{R}$. We have $a_n^* = sup_{m\geq n} a_m$.

Notice that a_n^* is decreasing.

Definition 4.6.

$$lim_{n\to\infty}\ sup\ a_n = inf_{n\geq 1} a_n^* \overset{Remark}{=} lim_{n\to\infty} a_n^*$$

$$= lim_{n\to\infty} sup_{m\geq n} a_m.$$

Remark 4.7. Notice that *since* a_N^* *decreasing.*

□

In a similar way, we also have

Definition 4.8.

$$lim_{n\to\infty}\ inf\ a_n = lim_{n\to\infty}\ inf_{m\geq n} a_m$$

$$= sup_{n\geq 1}\ inf_{m\geq n} a_m$$

An exercise can be: consider $a_n = (-1)^n$. Then $lim_{n\to\infty}\ sup(-1)^n =$ 1 and $lim_{n\to\infty}\ inf(-1)^n = -1$.

Proposition 4.9.

$$lim_{n\to\infty}\ inf\ a_n = -lim_{n\to\infty}\ sup(-a_n)$$

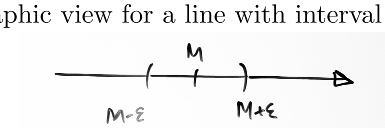
The moral of this proposition is that for *lim sup*, it implies dual proposition for liminf. Check textbook

Proposition 4.10. *This proposition states:*

(1) $lim_{n\to\infty}\ sup\ a_n \leq N \Leftrightarrow \forall\epsilon > 0,\ a_n < M + \epsilon$ *eventually.*

(2) $lim_{n\to\infty}\ sup\ a_n \geq M \Leftrightarrow \forall\epsilon > 0,\ a_n > M - \epsilon$ *frequently.*

Figure 14: Graphic view for a line with interval $(M-\epsilon, M+\epsilon)$.



From the graph illustration, we discuss two points. First, consider $M+\epsilon$, we have "lim$sup\ a_n \leq N$" $\Leftrightarrow$ "a_n only crosses this a finite number of times". Then consider $M-\epsilon$, we have "lim$sup\ a_n \geq M$" $\Leftrightarrow$ "a_n crosses an infinite number of times".

Proof: We discuss two cases accordingly:

First, $lim_{n\to\infty}\ sup\ a_n \leq M$

$$\begin{aligned} \underset{n\geq 1}{inf} a_N^* &\Leftrightarrow \forall\epsilon > 0, \exists N \in \mathbb{N} : a_n < M+\epsilon,\ for\ n \geq N, \\ &\quad from\ definition \\ &\Leftrightarrow a_m < M+\epsilon,\ form \geq n \geq N \end{aligned}$$

Second, $\lim_{n\to\infty} \sup\ a_n \geq M$

$$\begin{aligned} \underset{n\to\infty}{inf}\ a_N^* \geq M \quad &\Leftrightarrow \quad a_n^* M, for\ all\ n \\ &\Leftrightarrow \quad \forall\epsilon > 0, \forall N \in \mathbb{N}, a_m > M - \epsilon, for\ some\ m \geq N \end{aligned}$$

Q.E.D.

Then we have the following proposition.

Proposition 4.11. *This proposition states two cases:*
(1) liminf $a_n \leq M \Leftrightarrow \forall\epsilon > 0, a_n < M + \epsilon$ frequently
(2) liminf $a_n \geq M \Leftrightarrow \forall\epsilon > 0, a_n > M_\epsilon$ eventually.

In the discussion of limits, we are given (X, dx) and (Y, dy) as metric spaces with a function $f : X \to Y$. Then we have the following definition.

Definition 4.12. $f(x) \to q$ if $\forall\epsilon > 0$, $\exists\delta > 0 : x \in B_\delta(p) \setminus \{p\} \Rightarrow f(x) \in B_p(q)$.

Then we have the analogy with limit of sequences, which is the following. Consider $p_n \to p$ as $n \to \infty$, $\forall\epsilon\ \exists N \in \mathbb{N}$, we have $n \geq N \Rightarrow p_n \in B_\epsilon(p)$, i.e. converge eventually. Then for $f(x) \to q$ as $n \to \infty$, we have $\forall\epsilon > 0$, $\exists\delta > 0 : x \in B_\epsilon(p) \setminus \{p\} \Rightarrow f(x) \in B_\epsilon(q)$.

Remark 4.13. Note that $f(p)$ plays no role here. We can consider $E \subseteq X$ while (E, dx) is a metric space with function $f : E \to Y$.

□

Definition 4.14. Consider $p \in X$, $f(x) \to q \in Y$ as $x \to p$. Then we have the following argument: if $\forall\epsilon > 0$, $\exists\delta > 0$ s.t. $\forall x \in B_\delta(p) \cap E \setminus \{p\}$, then it implies that $f(x) \in B_\epsilon(q)$.

Remark 4.15. Note that (1) if $p \in X \setminus \bar{E}$, then the definition holds for any q; (2) if $p \in E' \setminus \{E\}$, then we can discuss possible limit, something we are interested, and (3) if $p \in E \setminus E'$, then definition holds for any p.

□

Remark 4.16. Note that we can also consider the metric space X to be restricted if we define $X \to (E, dx)$.

□

Remark 4.17. For the metric space E, it is inherited from X.

□

Then we have the following proposition.

Proposition 4.18. Limit of functions and limit of sequences.
"$f(x) \to q$ as $x \to q$"
$\Rightarrow$
"$\forall p_n \neq p \to p$ as $n \to \infty$, we have $f(p_n) \to q$ as $n \to \infty$".

Starting from here, we would discuss the consequences from series. Notice that the proposition for limit of sequences also implies the same propositions for limit of functions.

(1.1) Uniqueness. Consider $p \in E'$, then "$f(x) \to q$ as $x \to p$ and $f(x) \to q'$ as $x \to p$" $\Rightarrow$ "$q = q'$.
Proof: If "$p \in E'$, $\exists p_n = p$ as $n \to \infty$ s.t. $f(p_n) \to q$ as $n \to \infty$, and $f(p_n) \to q'$ as $n \to \infty$", then "$q = q'$ is implied".

Q.E.D.

(2.2) Limits and operation of $\mathbb{R}^n$. Given $f : E \subseteq X \to \mathbb{R}^k$, and $g : E \subseteq X \to \mathbb{R}^k$ such that $f(x) \to A$ as $x \to p$ and $g(x) \to B$ as $x \to p$, we have
(2.2.a) $(f+g)(x) \to A+B$ as $x \to p$, $\forall \alpha, \beta \in \mathbb{R}, (\alpha f + \beta g)(x) \to \alpha A + \beta B$ as $x \to p$.
(2.2.b) "$(f \circ g)(x) \to AB$ as $x \to p$" $\Rightarrow$ "If $[g(x)e^i = (0, ..., 0, 1, 0, ...)$ with 1 to be the ith element], then $[f_i(x) \to A_i$ as $x \to p]$".
(2.2.c) For k=1, $g(x) \neq 0$, and $B \neq 0$, then we have $(\frac{f}{g})(x) \to \frac{A}{B}$ as $n \to p$.
(2.2.d) Given $f(x) = (f_1(x), ..., f_k(x))$, we have $f_i(x) \to A_i$ as $x \to p$ for all $i = 1, ..., k$
$\Rightarrow f(x) \to A = (A_1, ..., A_k)$ as $k \to p$.
(2.2.e) For k=1, $max(f, g)(x) \to max(A, B)$ as $n \to p$. Also $min(f, g)(x) \to min(A, B)$ as $n \to p$.

(2) Continuity

Definition 4.19. Consider $f : X \to Y$, f is continuous at $p \in X$ if $f(x) \to f(p)$ as $x \to p$. In other words, f is continuous if it is continuous $\forall p \in X$.

Remark 4.20. If p is isolated, then f is automatically continuous at p. If δ is small, then $B_\delta(p) \setminus \{p\} = \emptyset$, nothing to check.

□

Remark 4.21. $E \subseteq X$, $f : X \to Y$ continuous at p is defined if $p \in E$.

□

For an example, consider the identity map: $f : X \to Y$ s.t. $I(x) = x$, we can check $\epsilon = \delta$. This will be simply be an equal size ball on p on both axis.

For another example, consider $E \subseteq X$, $d_E : X \to \mathbb{R}$, then $d_E(x) = d(x, E) = inf\{d(x, y) : y \in E\}$. Then we can check $d(x, E) \to d(p, E)$ as $x \to p$, and $d(p, E) - d(x, E) \leq d(p, y) - d(x, E), \forall y \in E$, of which $d(p, y)$ is the best lower bound.

Then for every $\epsilon > 0$, $\exists z \in E$ s.t. $d(x, Z) \leq d(x, E) \leq d(x, Z) + \epsilon$. For the last step to be true, simply take $y = z$, then we have

$$\begin{aligned} d(p, E) - d(x, E) &\leq d(p, Z) - [d(x, Z) - \epsilon] \\ &\leq d(p, Z) + \epsilon \end{aligned}$$

Moreover, if $d(x, Z)$, then $f(x) - f(p) < \epsilon + \delta$. The case that bounded from below is similar. That is, $|f(x) - f(p)| < \epsilon + \delta$, and simply choose $\delta = \epsilon$ then we have $|f(x) - f(p)| < 2\epsilon$ if $d(x, p) < \epsilon$.

Proposition 4.22. *Consider $f : E \subseteq X \to \mathbb{R}^k$, and $g : E \subseteq X \to \mathbb{R}^k$, then we have*

(2.1) "f and g are continuous" $\Rightarrow$ "$\alpha f + \beta g$ is continuous $\forall \alpha, \beta \in \mathbb{R}$"

(2.2) "f and g continuous $\Rightarrow$ $f \circ g$ continuous" $\Rightarrow$ "f is continuous $\Rightarrow$ f_i continuous".

(2.3) For $k = 1$, f and g are continuous and $g \neq 0 \Rightarrow \frac{f}{g}$ continuous.

(2.4) For $k = 1$, f and g continuous $\Rightarrow$ we have $max(f, g)$ and $min(f, g)$ continuous.
Note: we define $max(f, g)(x) = max(f(x), g(x))$.

For an exercise, consider for $k = 1$ the polynomial, $p(x) = a_n x^n + \cdots + a_1 x + a_0$. We have (1) $I : \mathbb{R} \to \mathbb{R}$, there is $I(x) = x$ continuous, (2) x^i is continuous by definition of product, and (3) p is continuous by super position (linear).

We can discuss the following scenarios.

Monomials. Consider $x_1^{i1}, x_2^{i2}, ..., x_k^{ik}$ are continuous. Then $x_1^{i1}, ..., x_k^{ik}$ are polynomials, which implies that monomials are continuous by

Remark 4.23. Note that consider the $x_1, ..., x_k$ are continuous. The first term x_1 will be x_{e1}, and so on.

□

Polynomials. Polynomials are linear combination of monomials, which is continuous.

Example 4.24. For example, we can consider xy or simply $x^2 + y^2$ from calculus.

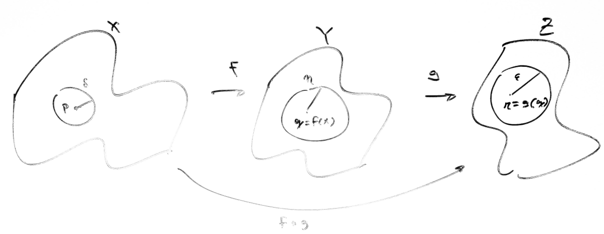
□

Rationals. Rationals are defined as the ratio of polynomials, which is restricted to $\{x \in \mathbb{R}^k,\ denomination \neq 0\}$.

Example 4.25. For example, $\frac{1}{x^2+1}$ or simply the form $\frac{ax+by}{cx+dy}$. An exericse can be $|x| = max(x, 0) + max(-x, 0)$.

Composition. Consider metric spaces X, Y, Z. Consider maps $f : X \to Y$ and $g : Y \to Z$. Then $f \circ g$ is well-defined.

Figure 15: Graphic view for a composition: for metric spaces X, Y, and Z, we have $p \in X$, $q = f(x) \in Y$, and $r = g(q) \in Z$.



□

We have the following property:

Proposition 4.26. *"f continuous at p and g continuous at q" $\Rightarrow$ "$g \circ f$ is continuous at p".*

Proof: Let $\epsilon > 0$ g continuous at $q \Rightarrow \exists \eta > 0 : g(y) \in B_\epsilon(r)\ if\ y \in B_\eta(q)$.
And also f continuous at p implies that $\exists \delta > 0,\ f(x) \in B_\eta(q)\ if\ x \in B_\delta(p)$.
This implies: "$x \in B_\delta(p) \Rightarrow f(x) \in B_\eta(q)$" $\Rightarrow$ "$g \circ f(x) \in B_\epsilon(r)$".

Q.E.D.

An exercise can be $|x^2 - 1|$ is continuous, or $[d_E(x)]^2$ is continuous.
Proof: Consider open scenario with closed scenario to be exercise.
Consider $G \subseteq Y$ that is open. The goal is to show $\forall p \in q^{-1}(G)$ is an interior point, which implies that $f(p) \in G$ that is open.
Then we have $\exists \epsilon > 0 : B_\epsilon(f(p)) \subseteq G$, since f is continuous, then $\exists \delta > 0$ s.t. $f(x) \in B_\epsilon(f(p))$ if $x \in B_\delta(p)$.
$\Rightarrow B_\delta(p) \subseteq f^{-1}(g)$ since "$B_\epsilon(f(p)) \subseteq G \to B_\delta(p) \subseteq f^{-1}(g)$".

Q.E.D.

(3) Continuity and Compactness (Weierestrass)

(4) Uniform Continuity

Recall that we have a function $f : X \to Y$ is continuous if and only:

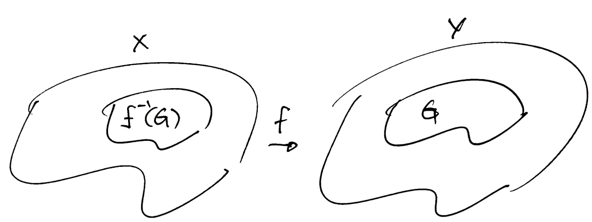
(1) $f(x) \to f(p)$ as $x \to p$ $\forall p \in X$.

(2) $f(p_n)_n \to f(p)$ as $\to \infty$ $\forall p_n \to p$.

(3) $\forall G \subseteq Y$ open if and only if $f^{-1}(G) \subseteq X$ is open

(4) $\forall F \subseteq Y$ closed if and only if $f^{-1}(F) \subseteq X$ is closed

Figure 16: Graphic view for a continuous function $f : X \to Y$.

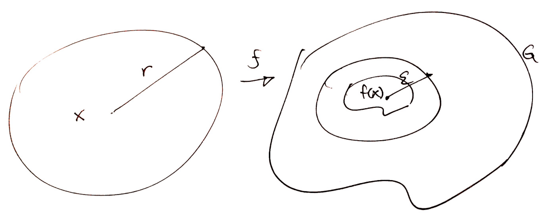


Proof: (Proof for (3)).

[$\Rightarrow$] Goal is $\forall x \in f^{-1}(G)$, $\exists r > 0$ s.t. $B_\epsilon(x) \subseteq f^{-1}$. The idea is that for $f(x) \subseteq G$ with G open, we have $\exists \epsilon > 0$ such that $B_\epsilon(f(x)) \subseteq G$. By continuity, $\exists r > 0$, $\exists \epsilon > 0$, $B_\epsilon(f(x)) \subseteq G$. Notice that $f^{-1}(B_\epsilon(f(x))) \supseteq$

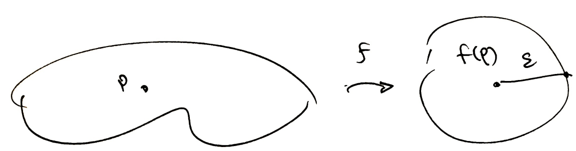
$B_r(x)$, but $f^{-1}(B_\epsilon(f(x))) \subseteq f^{-1}(G)$. This implies that $B_r(x) \subseteq f^{-1}(G)$, referring to the graph below.

Figure 17: Graphic view for proof (3) [$\Rightarrow$].



[$\Leftarrow$] Use $B_\epsilon(p)$ which is an open set. $\forall \epsilon > 0$, we have $f^{-1}(B_\epsilon(f(p)))$ is open and $p \in f^{-1}(B_\epsilon(f(p)))$, i.e. $\exists \delta > 0$ such that $B_\delta(p) \subseteq f^{-1}(B_\epsilon f(p)))$. This is equivalent as saying $\forall x \in B_\delta(p)$, we have $f(x) \in B_\epsilon(f(p))$, which implies that $f(x) \to f(p)$, as $x \to p$.

Figure 18: Graphic view for proof (3) [$\Leftarrow$].



Proof (4) will be an exercise.

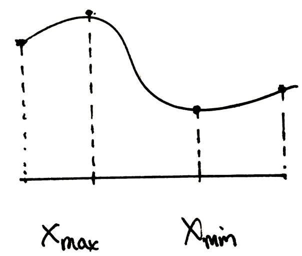
Q.E.D.

4.2 Weierstrass Theorem (Extreme Value Theorem)

Go back to Table of Contents. Please click TOC

This section we introduce the Extreme Value Theorem, which is also the Weierstrass Theorem. Consider the following graph, there is a function f and there exists x_{max} and x_{min} such that $f(x_{max})$ gives us maximum and $f(x_{min})$ gives us minimum.

Figure 19: Graphic view for Weierstrass Theorem.



Theorem 4.27. *(Extreme Value Theorem, a.k.a., Weierstrass Theorem). Consider $f : X \to Y$ continuous. $\forall k$ to be compact and $k \in K \subseteq X \Rightarrow f(k) \subseteq Y$ is compact.*

Let us start by considering some particular cases.

First, consider k is finite. This implies that $f(k)$ is finite if f is not continuous.

Second, consider $f : [a, b] \to \mathbb{R}$. Then it is the case that $f([a, b]) \subseteq \mathbb{R}$ to be compact, which is equivalent as saying it is closed and bounded. (It is closed because there is *sup* and *inf* of $f([a, b])$ are actually *max* and *min*, and it is bounded because we can define *sup* and *inf* of $f([a, b])$).

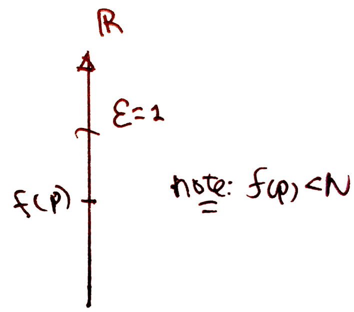
Also we have the following lemmas.

Lemma 4.28. *Assume $f : [a, b] \to \mathbb{R}$ and also f is continuous. f must be bounded. That is, $\exists M \in \mathbb{R}, f(x) \in (-M, M) \forall x \in [a, b]$.*

Proof: (Prove that it is bounded above.) By contradiction, $\forall n \in \mathbb{N}$, $\exists x_n \in [a, b]$ such that $f(x_n) > n$. Then we have $\{x_n\}$ is a sequence $\subseteq [a, b]$. This implies that there exists a convergent subsequence $\{x_{n_k}\}$, that $x_{n_k} \to p \in [a, b]$. Then we have f defined at $p \in [a, b]$, $\exists N \in \mathbb{N}$, such that $f(p) < N$.

From Archimedean Property, use continuity of f at p with $\epsilon = 1$ (note that ϵ can be anything), we have $f(x) < N + 1$, for $x \in (p - \delta, p + \delta)$. Now $x_{n_k} \in (p - \delta, p + \delta)$ eventually, which contradicts $f(x_{n_k})$, which can be referred to the following graph.

Figure 20: Graphic view for proof of Lemma 1.



An exercise can be the following. Show that f is bounded below. The proof is to reproduce the same argument as above. The fact that f is continuous would imply that $-f$ is also continuous.

Here is an alternative proof. Let M be the least upper bound of $f(x)$ for any x in $[a, b]$. If this is true, then we have the following, $\forall \epsilon > 0$, $\exists n \in \mathbb{N}$ such that for x_1 we have $f(x_1) - M = \epsilon$, for x_2 we have $f(x_2) - M = \epsilon/2$, and so on. In other words, as $n \to \infty$, the image for x_n, that is $f(x_n)$, will go very close to M. This is the definition of boundedness.

Consider this a series, $x_1, x_2, ..., x_n$, which we call $\{x_n\}$. Then we take any subsequence of $\{x_n\}$. We find the subsequence by the following steps (by Archimedean Property). First, we split the interval $[a, b]$ into two equally sized pieces. In other words, let us split $[a, b]$ into two and pick the piece with infinitely many elements and call it $[a_1, b_1]$, call it d_1. Then we consider the sub-interval $[a_1, b_1]$ and split in half again, which is $[a_2, b_2]$ and let us call it d_2. We choose the piece with infinitely many elements and call it $[a_3, b_3]$ and call it d_3. We keep repeating this action and we will get $d_1, d_2, ..., d_n$. This action will cut d_n smaller and smaller as n goes to infinity. We can denote a d that is the limit as n goes infinity, which is $\lim_{n\to\infty} d_n = d$. Hence, we have

$$f(d) = f(\lim_{n\to\infty} d_n) = \lim_{n\to\infty} f(d_n) = \underset{n\to\infty}{f}(x_n)$$

Because we chose x_n arbitrarily such that the images $f(x_n)$ were getting closer and closer to M, we see this limit must equal to M. From the equal above, we know this $f(d) = M$, which is what we want.

Here is another alternative, by contradiction. First of all, we want to show that it is bounded above and then we show that there is a *sup*. Assume by contradiction, that it is not bounded above, that is, $\forall n > N_1 \in \mathbb{N}$, $\exists x_n \in [a, b]$ such that $f(x_n) > N_1$. Then $\{x_n\}$ can be sequence. By Bolzano-Weierstrass, subsequence $\{X_{n_k}\} \to p_1$, for some $p - 1$. At p_1, $f(p_1)$ is well defined and continuous. From Archimedean Property, we have $x_{n_k} \in \{p_1 - \delta, p_1 + \delta\}$, $\forall \delta > 0$ eventually, which leads to contradiction.

Then we know it has to be bounded above. There exists $p_1 \in [a, b]$

$f(p_1) = \sup\limits_{x\in[a,b]} f(x)$, call it M, and we have $\forall n \in \mathbb{N}$ such that $M - \frac{1}{n} < M$, which implies that $\exists x_M \in [a,b]$ such that $M - \frac{1}{n} < f(x_M) \leq M$. We conclude, Squeeze Theorem, there exists $\exists x_M$ such that $f(x_M) = M$.

Second, we want to show that it is bounded below and there is an *inf*. Assume by contradiction that "it is not bounded below". That gives us $\forall n > N_2 \in \mathbb{N}$, $\exists x_n \in [a,b]$, $f(x_n) < N_2$. Then $\{x_n\}$ can be a sequence. By Bolzano-Weierstrass, subsequence $\{x_{n_k}\} \to p_2$ for some p_2, for $f(p_2)$ defined and continuous. By A.P., we have $x_{n_k} \in \{p_2 - \delta, p_2 + \delta\} \forall \delta > 0$ eventually, which leads to contradiction.

Then we show that there exists *inf*. There $\exists p_2 \in [a,b]$ such that $f(p_2) = \underset{x\in[a,b]}{inf} f(x)$, let us call it m. $\forall n \in \mathbb{N}$, we have $m < m + \frac{1}{n}$, which implies that $\exists x_m \in [a,b]$ such that $m \leq f(x_m) < m + \frac{1}{n}$. From squeeze theorem, we say that $\exists x_m : f(x_n) = m$.

Q.E.D.

Lemma 4.29. *There exists $q \in [a,b]$ such that $f(p) = \sup_{x\in[a,b]} f(x)$*

Proof: Observe that $\sup_{x\in[a,b]} f(x)$ is well-defined, and that $f(a) \in \{f(x) \in \mathbb{R} : x \in [a,b]\}$, so it is not an empty set.

From Lemma 1 above, we know that the set is bounded above. Let $\sup_{x\in[a,b]} f(x) = M$, we have $M \geq f(x)$ for all $x \in [a,b]$. We have $\forall \epsilon > 0$, $M - \epsilon$ is not an upper bound. Then $\exists x \in [a,b]$ such that $f(x) > M - \epsilon$.

Take $n \in \mathbb{N}$, $\epsilon = \frac{1}{n} > 0$, let $x_n \in [a,b]$ such that $M - \epsilon < f(x_n) \leq M$. Thus, $\{x_n\} \subseteq [a,b]$ has a convergent subsequence that $x_{n_k} \to p \in [a,b]$.

Now the goal is to show that $f(p) = M$. By continuity at p, we have $f(x_{n_k}) \to f(p)$ as $k \to \infty$. which is what we want. By contradiction of $f(x_n)$, we have $f(x_n) \to M$ as $n \to \infty$ (checked). Then $\{f(x_{n_k})$ is a subsequence of $\{f(x_n)\}$. This implies that $f(x_{n_k}) \to M$ as $k \to \infty$, which implies that $f(p) = M$. Done.

Q.E.D.

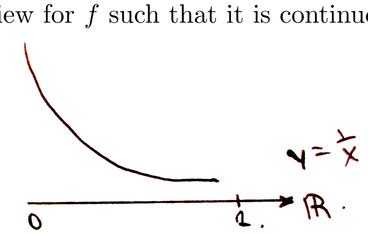
Example 4.30. For example, $\exists q \in [a,b]$ such that $f(q) = inf_{x\in[a,b]} f(x)$. It is necessary of the proof for boundedness and closeness.

First example, consider $f(0,1) \to \mathbb{R}$, which is continuous and unbounded in what situation?

□

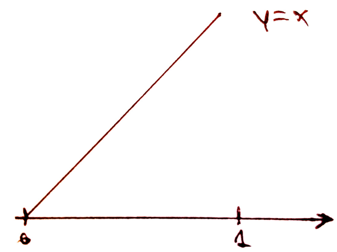
We can refer to the following graph, i.e. $y = \frac{1}{x}$. In this case, we have f well defined on $(0,1)$ and we have a continuous function. Note that f is not bounded.

Figure 21: Graphic view for f such that it is continuous and unbounded.



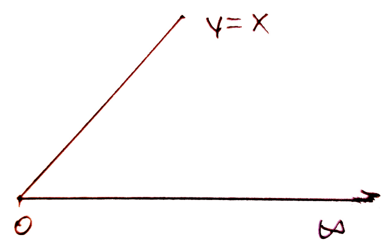
Second example, consider $f : (0,1) \to \mathbb{R}$ continuous and bounded but $f(p) \neq sup_{x\in(0,1)} f(x)$. We can simply take $f(x) = x$. It is continuous for certain and also bounded since $x = 1$ is not included.

Figure 22: Graphic view for f that is continuous and bounded.



Third example, consider $f : [0,\infty) \to \mathbb{R}$ continuous and unbounded. Take $f(x) = x$ again and it is straightforward.

Figure 23: Graphic view for f that is continuous and bounded.



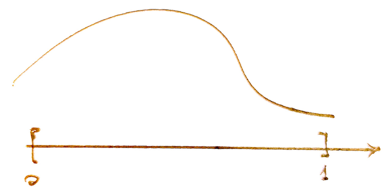
Fourth example asks the following: for $f : [0,\infty) \to [R)$ that is continuous and bounded, but $f(p) \neq sup_{x\in[0,\infty)} f(x), \forall p \in (0,\infty)$

Fifth examples states: $f[0,1] \to \mathbb{R}$ bounded but $f(p) \neq sup_{x\in[0,1]} f(x), \forall p \in [0,1]$. Consider the function:

$$f(x) = \begin{cases} x - x^2 & , \quad if\ x \neq \frac{1}{2} \\ 0 & , \quad if\ x = \frac{1}{2} \end{cases}$$

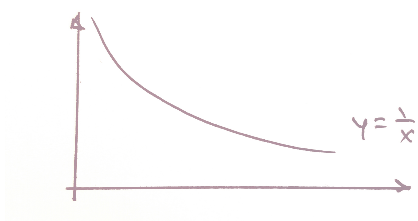
which is presented in the following graph.

Figure 24: Graphic view for $f[0,1] \rightarrow \mathbb{R}$ bounded but $f(p) \neq sup_{x\in[0,1]}f(x), \forall p \in [0,1]$.



Sixth example: consider $f : [0,1] \rightarrow \mathbb{R}$ to be unbounded. Take $f(x) = \frac{1}{x}$, as shown in the graph below.

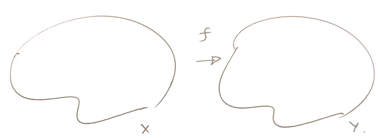
Figure 25: Graphic view for $f : [0,1] \rightarrow \mathbb{R}$ to be unbounded.



With the idea in mind, we can introduce the general theorem.

Theorem 4.31. *(General Theorem. Weierstrass Theorem). Consider $f : X \rightarrow Y$, $K \subseteq X$ and K is compact. Then we have $f(k)$ to be compact while $k \in K$.*

Figure 26: Graphic view for Weierstrass Theorem.



Referring to the graph we have:
(1) G open $\Rightarrow f^{-1}(G)$ to be open.
(2) F closed $\Rightarrow f^{-1}F(G)$ to be closed.
(3) K compact $\Rightarrow f(k)$ to be compact.

Proof: The goal is the following: for any open covering of $f(k)$ has a finite sub-covering, we have $\{G_\alpha\}$ is covering of $f(g)$ while f is continuous. This implies that $f^{-1}(G_\alpha) \subseteq X$ is open for all α and $\{f^{-1}(G_\alpha)\}$ is an open covering for that compact set K.

$\Rightarrow \exists \alpha_1, \ldots, \alpha_n$ such that $K \subseteq f^{-1}(G_\alpha) \cup \cdots \cup f^{-1}(G_{\alpha_n})$

$\Rightarrow f(k) \subseteq G_{\alpha_2} \cup \cdots \cup G_{\alpha_n}$. That means

$$\begin{aligned} f(k) \subseteq \bigcup_\alpha G_\alpha &\Leftrightarrow \forall q \in f(k), q \in G_\alpha for\ some\ alpha \\ &\Leftrightarrow \exists p \in K,\ q = f(p) \\ &\Leftrightarrow p \in f^{-1}(G_\alpha) \end{aligned}$$

An application can be the following: consider $f : X \to Y$ with X compact. Given that f is continuous and bijective, we have $r^{-1} : Y \to X$ is continuous.

4.3 Intermediate Value Theorem

Go back to Table of Contents. Please click TOC
This subsection we discuss Intermediate Value Theorem, a.k.a., Weierstrass Theorem.

Theorem 4.32. *Given a function $f : X \to Y$ that is continuous, and also suppose we have $K \subseteq X$ that K is compact. Then the theorem states that $f(k) \subseteq Y$ while $f(k)$ is compact.*

Corollary 4.33. *Given $k = [a, b]$, and also given $Y = \mathbb{R}$, then we have $f([a, b]) \subseteq \mathbb{R}$ compact if and only if $f([a, b])$ is bounded and closed.*

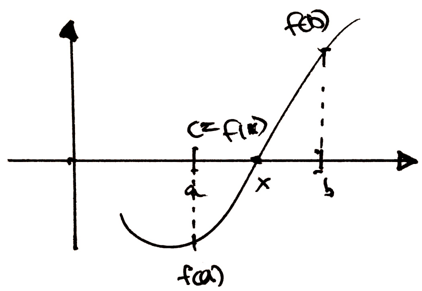
This also implies that

$$\underset{x \in [a,b]}{sup} f(x) = f(x^*)$$

for some $x^* \in [a, b]$.

Corollary 4.34. *Given $f : X \to Y$ compact, given X to be compact, and f a bijection, we have $f^{-1} = g : Y \to X$ to be continuous.*

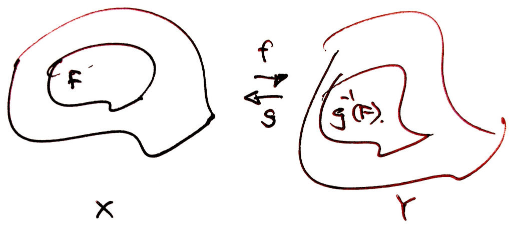
Figure 27: Graphic view for Intermediate Value Theorem.



Proof: Referring to the graph below, we want g to be continuous, which is equivalent as $g^{-1}(F)$ to be closed if F is closed while $F \subseteq X$.

Then take $F \subseteq X$ to be closed. We have F to be compact (because X is compact). This implies that $f(F)$ is compact because f is continuous. Thus, we have $f(F)$ to be closed and $f(F) = g^{-1}(F)$.

Figure 28: Graphic view for the proof.

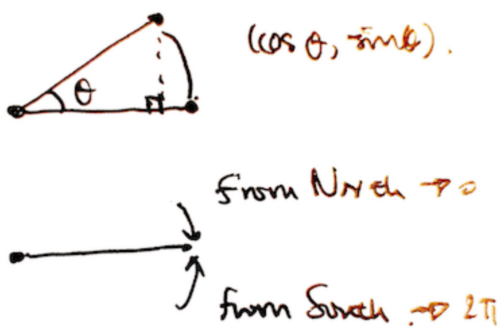


Q.E.D.

For an example, we can consider the following, $f : [0, 2\pi) \to \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. Then we have $f(\theta) = (cos\theta, sin\theta)$, which is shown in the graph. Consider f to be continuous, and bijection, which are both assumptions. We would have $f^{-1} = g$ to be not continuous.

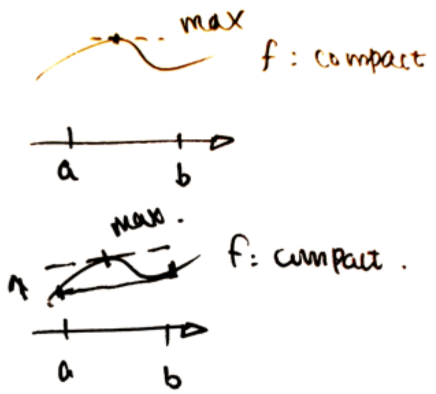
Why is this the case? Consider $\theta = 0$ and at this point we need to consider $\theta \to 0^+$ and also $\theta \to 0^-$. From the definition of sine and cosine function, we know that the limit from North would tend to 0 but the limit from the South would tend to 2π, hence causing a discontinuity.

Figure 29: Graphic view for example.



Later on, we can apply W.T. to Rolles Theorem, and mean value theorem. Consider f to be compact and define an interval $[a, b]$. We would have a max that can be found by connecting $f(a)$ and $f(b)$ and shift the line above.

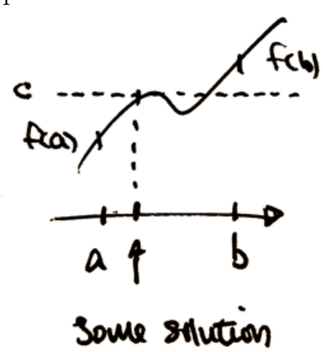
Figure 30: Graphic view for some application, i.e. Rolles Theorem.



Here we discuss Intermediate Value Theorem.

Theorem 4.35. *Suppose $f : [a, b] \to \mathbb{R}$ is a continuous function. Assume $f(a) \leq f(b)$ without loss of generality and $c \in [f(a), f(b)]$. Then there exists $x^* \in [a, b]$ such that $f(x^*) = c$, hich can be shown in the graph.*

Figure 31: Graphic view for Intermediate Value Theorem.



The idea is straightforward as the below. Let us say the goal is to compute $\sqrt{3}$. This is equivalent as finding x such that $x^2 - 3 = 0$. We choose $x^2 - 3|_{x=0} = -3 < 0$. Then choose $x^2 - 3|_{x=2} = 1 > 0$. Observe one positive and negative relation. We choose $x^2 - 3|_{x=1} = -2 < 0$, at $x = 1$ in between the values of x in previous two statements. We repeat this process and we would eventually get an answer satisfies our goal.

Now we can prove the theorem.

Proof: Let us present the proof from the class and we will present the proof from textbook. Consider $I_0 = [a, b] = [a_0, b_0]$ so we have $c_0 = \frac{b_0 + a_0}{2}$. Define $I_{n+1} = [a_{n+1}, b_{n+1}]$ and $c_{n+1} = \frac{b_{n+1} + a_{n+1}}{2}$ recursively. If $f(c_n) > c$, then $I_{n+1} = [a_n, c_n]$. If else, then we take $I_{n+1} = [c_n, b_n]$.

Here let us recall the sequence of nested intervals $I_0 \supseteq I_1 \supseteq I_2 \supseteq \dots$ such that $||I_n|| = \frac{b-a}{2^n}$ and also $f(a_n) \leq f(b_n)$. Let $x^* = \underset{n\geq 1}{sup} a_n$ and also assume that $\{a_n\}$ is increasing and bounded above. Then we have $x^* = sup\ a_n = \lim_{n\to\infty} a_n$ since f is continuous $f(x^*) = \lim_{n\to\infty} f(a_n)$.

By contradiction, $\exists \epsilon > 0$ such that $f(a_n) \nsubseteq (c-\epsilon, c+\epsilon)$ eventually (see remark). This is equivalent as $f(a_{n_k}) \nsubseteq (c-\epsilon, c+\epsilon)$ for some subsequence. Notice that $f(a_{n_k})$ has a convergent subsequence, because it is contained in a compact set $f([a,b])$.

Thus we infer that $\exists f(a_{n_{k_l}}) \to d$ as $l \to \infty$ and that $f(a_{n_{k_l}}) \notin (c-\epsilon, c+\epsilon)$. Here, $d \notin (c-\epsilon, c+\epsilon)$, otherwise we would contradict $f(a_{n_{k_l}}) \to d$.

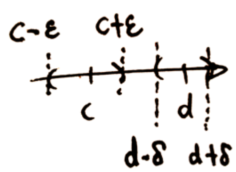
Remark 4.36. Notice that $f(a_n) \nsubseteq (c-\epsilon, c+\epsilon)$ eventually means that it is outside frequently.

□

Referring to the graph, we can argue that $\exists \delta > 0$ such that $(c-\epsilon, c+\epsilon) \cup (d-\delta, d+\delta) = \emptyset$ and that f is continuous at d. Moreover, we have $||I_n|| \to 0$, i.e. the norm of I_n goes to zero, e.g. $a_{n_{k_l}} \to x^*$.

Hence, $f(I_n) \subseteq (d-\delta, d+\delta)$ eventually, which contradicts the fact $f(a_n) \leq c \leq f(b_n) \Rightarrow f(a_n) \to c as\ n \to \infty \Rightarrow f(x^*) = c$.

Figure 32: Graphic view for Intermediate Value Theorem.



Q.E.D.

Here is an alternative proof from textbook.[11]

Theorem 4.37. *(Intermediate Value Theorem). Let $a < b$, and let $f : [a,b] \to \mathbb{R}$ be a continuous function on $[a,b]$. Let y be a real number between $f(a)$ and $f(b)$, i.e., either $f(a) \leq y \leq f(b)$ or $f(a) \geq y \geq f(b)$. Then there exists $c \in [a,b]$ such that $f(c) = y$.*

Proof: We have two cases: $f(a) \leq y \leq f(b)$ or $f(a) \geq y \geq f(b)$. We will assume the former (without loss of generality), that $f(a) \leq y \leq f(b)$; the latter is proven similarly.

If $y = f(a)$ or $y = f(b)$, then the claim is easy, as one can simply set $c = a$ or $c = b$, so we will assume that $f(a) < y < f(b)$. Let E denote the set

$$E := \{x \in [a,b] : f(x) < y\}.$$

Clearly E is a subset of $[a,b]$, and is hence boudned. Also, since $f(a) < y$, we see that a is an element of E, so E is non-empty. By the least upper bound principle, the supremum

$$c := sup(E)$$

is thus finite. Since E is bounded by b, we know that $c \leq b$; since E contains a, we know that $c \geq a$. Thus we have $c \in [a, b]$. To complete the proof we now show that $f(c) = y$. The idea is to work from the left of c to show that $f(c) = y$. The idea is to work from the left of c to show that $f(c) \leq y$, and to work from the right of c to show that $f(c) \geq y$.

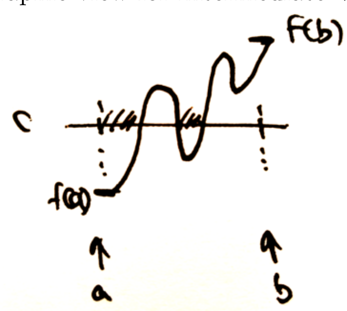
Let $n \geq 1$ be an integer. The number $c - \frac{1}{n}$ is less than $c = sup(E)$ and hence cannot be an upper bound for E. Thus there exists a point, call it x_n, which lies in E and which is greater than $c - \frac{1}{n}$. Also $x_n \leq c$ since c is an upper bound for E. Thus,

$$c - \frac{1}{n} \leq x_n \leq c.$$

By the squeeze test (Corollary 6.4.14, see remark), we thus have $\lim\limits_{n\to\infty} x_n = c$. Since f is continuous at c, this implies that $\lim\limits_{n\to\infty} f(x_n) = f(c)$. But since x_n lies in E for every n, we have $f(x_n) < y$ for every n. By the comparison principle (Lemma 6.4.13, see remark), we have $f(c) \leq y$. Since $f(b) > f(c)$, we conclude that $c \neq b$.

Since $c \neq b$ and $c \in [a, b]$, we must have $c < b$. In particular there is an $N > 0$ such that $c + \frac{1}{n} < b$ for all $n > N$ (since $c + \frac{1}{n}$ converges to c as $n \to \infty$). Since c is the supremum of E and $c + \frac{1}{n} > c$, we thus have $c + \frac{1}{n} \notin E$ for all $n > N$. Since $c + \frac{1}{n} \in [a, b]$, we thus have $f(c + \frac{1}{n}) \geq y$ for all $n \geq N$. But $c + \frac{1}{n}$ converges to c, and f is continuous at c, thus $f(c) \geq y$. But we already knew that $f(c) \leq y$, thus $f(c) = y$, as desired.

Figure 33: Graphic view for Intermediate Value Theorem.



Remark 4.38. Here we introduce Corollary 6.4.14 of textbook [11].

Corollary 4.39. *(Squeeze Test). Let $(a_n)_{n=m}^{\infty}$, $(b_n)_{n=m}^{\infty}$, $(c_n)_{n=m}^{\infty}$ be sequences of real numbers such that*

$$a_n \leq b_n \leq c_n$$

for all $n \geq m$. Suppose also that $(a_n)_{n=m}^{\infty}$ and $(c_n)_{n=m}^{\infty}$ both converge to the same limit L. Then $(b_n)_{n=m}^{\infty}$ is also convergent to L.

□

Remark 4.40. Here we introduce Lemma 6.4.13 of textbook [11].

Lemma 4.41. *(Comparison Principle). Suppose that* $(a_n)_{n=m}^{\infty}$ *and* $(b_n)_{n=m}^{\infty}$ *are two sequences of real numbers such that* $a_n \leq b_n$ *for all* $n \geq m$. *Then we have the inequalities*

$$\sup(a_n)_{n=m}^{\infty} \leq \sup(b_n)_{n=m}^{\infty}$$

$$\inf(a_n)_{n=m}^{\infty} \leq \inf(b_n)_{n=m}^{\infty}$$

$$\lim \sup_{n\to\infty} a_n \leq lim \sup_{n\to\infty} b_n$$

$$\lim \inf_{n\to\infty} a_n \leq lim \inf_{n\to\infty} b_n$$

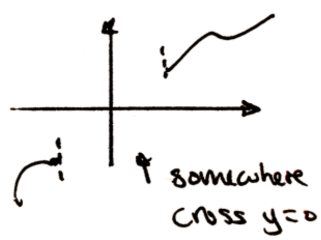
□

Q.E.D.

Now we can discuss some applications. There exists a solution to a certain equations. Let us say for $c > 0$, $\sqrt{c}$ exists by solving $x^2 - c = 0$. Then $c^{1/n}$ exists that is the roots of certain polynomials.

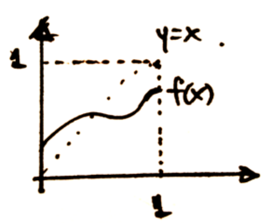
This is also the homework problem. Consider $P(x)$ a polynomial of odd degree. We can infer that there eixsts a rela root. For example, consider $P(x) = x^3 + x + L$. As shown below, there is a solution somewhere crosing $y = 0$.

Figure 34: Graphic view for Intermediate Value Theorem.



We can also fix a point. Consider $f : [0,1] \to [0,1]$ continuous. There eixsts $x^* \in [0,1]$ such that $f(x^*) = x^* \Leftrightarrow f(x^*) - x^* = 0$. As shown below, consider $g(x) = f(x) - x$. We can compare two functions. For $x \to 0$, we have f to be above. For $x \to 1$, we have f to be below. Then it must be the case that it crosses over in the interval $[0,1]$.

Figure 35: Graphic view for Intermediate Value Theorem.



Another application is the famous Meteorology theorem, which we can refer to textbook [5].

Theorem 4.42. *(Meteorology Theorem). Consider T to be temperature and P to be pressure.*

$$\exists a \in \mathbb{S}^2 \Rightarrow T(a) = T(-a) \textit{ and } P(a) = P(-a)$$

Remark 4.43. (Meteorology Theorem). The theorem can be interpreted as: somewhere on the Earth, there is a pair of antipodal points having simultaneously the same temperature and pressure.

□

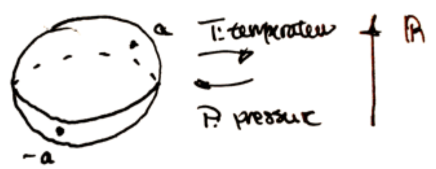
Proof: Let T, P respectively denote the temperature and pressure functions, assumed to be continuous, on the Earth's surface. Then we have a continuous map

$$f : \mathbb{S}^2 \to \mathbb{R}^2 : x \to (T(x), P(S))$$

and the theorem applies.

Q.E.D.

Figure 36: Graphic view for Metheology Theorem.



More ideas about the proof is the following. Let $f(x) = T(x) - T(-x)$. If $T(x_0) = T(-x_0)$, then we are done. Else if $T(x_o) > T(-x_0)$ (or $<$), x goes from x_0 to $-x_0$, that is, f goes from greater than 0 to less than 0.

Figure 37: Graphic view for proof of Metheology Theorem.

Here we attempt to introduce Toeplitz' conjecture, an unsolved problem formally published by the referred text [13]. In page 143 of referred text [13], it is introduced as the Squaring the Lake question. The questions is the following: prove that every simple closed curve in the plane contains four points forming the vertices of a square.

Figure 38: Graphic view for Intermediate Value Theorem.

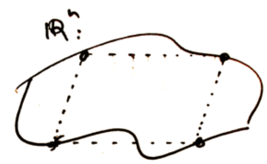
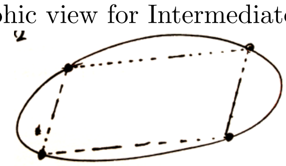


Figure 39: Graphic view for Intermediate Value Theorem.



Proof: I attempted the proof myself. Here is my thought. Suppose we have a simple closed curve in the plane.

Recall the definition of a simple closed curve:

Definition 4.44. A Jordan curve or a simple closed curve in the plane $\mathbb{R}^2$ is the image C of an injective continuous map of a circle into the plane, $\varphi : \mathbb{S}^1 \to \mathbb{R}^2$. A Jordan arc in the plane is the image of an injective continuous map of a closed interval into the plane.

Definition 4.45. Let f be a mapping of a metric space X into a metric space Y. We say that f is uniformly continuous on X if for every $\epsilon > 0$ there exists $\delta > 0$ such that

$$d_Y(f(p), f(q)) < \epsilon$$

for all p and q in X for which $d_X(p, q) < \delta$.

Proof starts here. Consider $\varphi : \mathbb{S}^1 \to \mathbb{R}^2$. We have φ an injection and continuous. The goal is to show φ is convex.

We have φ continuous, so we have $\varphi(tx+(1-t)y) = \varphi((1-t)y) \to \varphi(y)$ as $t \to \infty$, which implies that $\varphi(tx + (1-t)y)$ is bounded. By H.B. Theorem, we also know that it is compact.

Last, from Toeplitz [12] and Emch [6], we know it is always possible find four points that form a square.

Corollary 4.46. *On every simple closed continuous curve in the plane there are four points that form a square.*

Remark 4.47. From Toeplitz [12] and Emch [6], we have already known that for convex curves: on every simple closed continuous curve in the plane there are four points that form a square.

$\square$

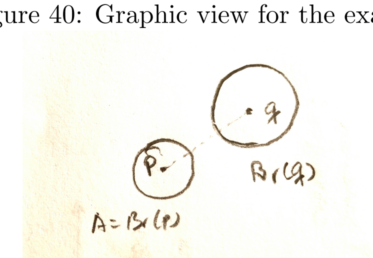
Q.E.D.

Next, consider general case as the following.

Definition 4.48. Suppose $E \subseteq X$ is disconnected. If $\exists A, B \subseteq X$ open and disjoint such that $E \subseteq A \cup B$ open covering and $E \cap A$ and $E \cap B \neq \emptyset$.

Example 4.49. For example, suppose $E = \{p, q\}$ is disconnected. We have $A = B_r(p)$ and $B = B_r(q)$. Simply take $r = \frac{1}{3}d(p, q)$, and we are done.

Figure 40: Graphic view for the example.



$\square$

Definition 4.50. Suppose $E \subseteq X$ is connected. If $\forall A, B \subseteq X$ open and disjoint, then we have $E \subseteq A \cup B$. This implies that $E \cap A = \emptyset$ or $E \cap B = \emptyset$.

Consider the following example. Suppose $X = \mathbb{R}$, $I = [a, b]$ is connected. By contradiction A and B are open and disjoint such that $I = A \cup B$, $I \cap A \neq \emptyset$ and $I \cap B \neq \emptyset$. Some $p \in I$ for some $q \in I$. Assume without loss of generality, $p < q$, $r \in I$, because I is an interval $r \in [p, q] \leq I$.

Question: $r \in A$ or B?

If $r \in A$, and $r \in I$, then $\exists \epsilon \in I$, then $\exists \epsilon > 0$, such that $(r - \epsilon, r + \epsilon) \subseteq A \cap I$. Then r is not an optimal upper bound.

If $r \in B$, and $r \in I$, then $(r - \epsilon, r + \epsilon) \subseteq B \cap I$. Then r is not an optimal upper bound for $A \cap I$.

Theorem 4.51. *Suppose $f : X \to Y$ is continuous and $E \subseteq X$ connected. Then $f(E)$ is connected.*

Corollary 4.52. *Suppose $E = [a, b]$, $f : E \to \mathbb{R}$. Then $f([a, b])$ is connected. Hence, the interval is shown and we can show I.V.T.*

From Intermediate Value Theorem, we consider $f : [a, b] \to \mathbb{R}$ to be continuous. Consider $c \in [f(a), f(b)]$. Then we have the following, $\exists x^* \in [a, b]$ such that $f(x^*) = c$.

There are three types of proofs that we discussed and we present accordingly.

Proof: Proof 1, Algorithm. Consider the graph below. We have some constants $c = 0$. Assume without loss of generality, $f(a) \leq c \leq f(b)$. We start the proof with considering interval $I_0 = [a_0, b_0] = [a, b]$. In each step, i.e. each n, we have $I_n = [a_n, b_n]$ such that $f(a_n) \leq c \leq f(b_n)$.

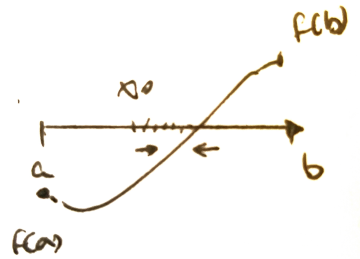
For the next stop, evaluate $f(\frac{b_n+a_n}{2})$. If $f(\frac{b_a+a_n}{2}) \geq c$, then $I_{n+1} = [a_n, \frac{b_n+a_n}{2}]$. If else, then we have $I_{n+1} = [\frac{b_n+a_n}{2}, b_n]$.

We claim that 1) a_n is increasing and bounded above (by a) which implies that a_n is increasing (consider a_∞); 2) b_n is decreasing and bounded below (by b) which implies that b_n is decreasing (consider b_∞).

Finally, we have $f(a_n) \leq c \leq f(b_n)$.

$$\left.\begin{array}{c} f(a_n) \to f(x^*) \\ f(b_n) \to f(x^*) \end{array}\right\} \Rightarrow f(x^*) = c$$

Figure 41: Graphic view for the Proof 1.



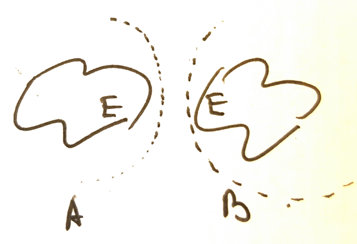
Proof: Proof 2, textbook [11]. Assume we have $f(a) \leq c \leq f(b)$. Consider $E = \{X \in (a,b) : f(x) \leq c\}$. Then we have 1) $E \neq \emptyset$ ($a \in E$), and 2) E bounded above (by b).

Then we have $X^* = sup\ E$, and claim that $f(x^*) = c$. We now can check whether the sign is greater or less than. We would have either "$f(x^*) > c$" or "$f(x^) < c$", yet both cases lead to contradiction. More specifically, $f(x^*) > c$ leads to contradiction of least upper bound; and $f(x^*) < c$ leads to contradiction of upper bound.

Proof: Proof 3, textbook [10].

Definition 4.53. $E \subseteq X$ is connected if A, B are open in X such that $E \subseteq A \cup B$ and $A \cap B = \emptyset$. That is, $E \subseteq A$ or $E \subseteq B$.

Figure 42: Graphic view for the definition.



Proposition 4.54. *Consider* $[a, b]$ *or any interval of* $\mathbb{R}$ *and is also connected. The hint here is to prove by contradiction. Take interval* $I \subseteq \mathbb{R}$*,*

and we would have form of $[a, b] \ldots$*, which can be expanded to* $(-\infty, b)$ *or* (a, ∞)*.*

I am not sure about this one and I will update this one later after I ask professor.

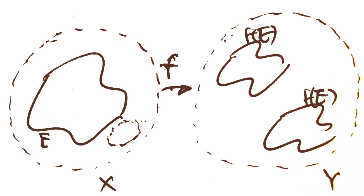
Now we consider the following property.

Proposition 4.55. *(Intermediate Value Prop). Take* $\alpha, \beta \in I \Rightarrow [\alpha, \beta] \subseteq I$*. Suppose a map* $f : X \to Y$*, and* E *is connected. Then we have* $f(E)$ *is connected.*

Corollary 4.56. *Consider* $[a, b] \to \mathbb{R}$*, and* $[a, b]$ *is connected. This implies that* $f([a, b])$*, any connected set of* $\mathbb{R}$*. For any connected set like this, and has I.V.P. satisfied, we would have* $\forall c \in (f(a), f(b))$*,* $\exists x^* \in [a, b]$*, such that* $f(x^*) = c$*.*

Now we show the proof (a proof within Proof 3.) *Proof:* Let A, B to be open sets in Y and given that they are disjoint. Let us say $f(E) \subseteq A \cup B$. Consider f to be continuous, and then we have $f^{-1}(A)$ and $f^{-1}(B)$ are open in X. That is, we would have $x \in E$ which implies that it should be the case that $f(x) \in f(E) \subseteq A \cup B$. This would then imply $f(x) \in A$ or $f(x) \in B$, which is equivalent as saying $x \in f^{-1}(A)$ or $x \in f^{-1}(B)$. However, we also know that $f^{-1}(A) \cap f^{-1}(B) = \emptyset$, which is equivalent as saying $f(x) \in A \cap B = \emptyset$. Then we have a contradiction. *Q.E.D.*

Figure 43: Graphic view for Proof 3.



This completes Proof 3.

Q.E.D.

Now we can put Intermediate Value Property and Extreme Value Theorem together.

Consider $f : [a, b] \to \mathbb{R}$ and f to be continuous. E.V.P. states that $\exists M, m$ such that

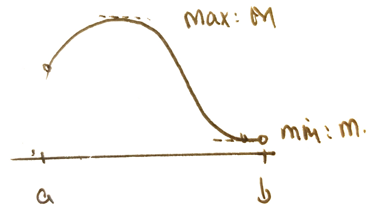
$$\forall x^* \in [a, b], \ M = \underset{x \in [a,b]}{sup} f(x) = f(x^*)$$

and

$$\forall x_* \in [a, b], \ m = \underset{x \in [a,b]}{inf} f(x) = f(x_*)$$

Then E.V.T. and I.V.T. states tells us $f([a, b]) = [m, M]$. Referring to the graph below, we can see that this means we would have an interval such that in the interval the domain gets mapped to an image that is bounded just exactly by maximum, M, and minimum, m.

Figure 44: Graphic view for putting E.V.T. and I.V.T. together.



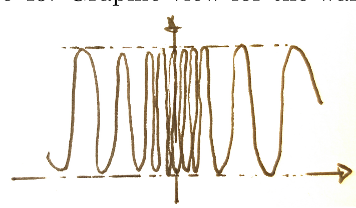
As a warning, we should be cautious that it is not the case I.V.P. implies continuity.

Consider the following example, which is shown in the graph. We would have

$$f(x) = \begin{cases} sin(\frac{1}{x} & , \quad if\ x \neq 0 \\ 0 & , \quad if x = 0 \end{cases}$$

which is presented in the graph. We would have a map satisfying I.V.P., but this is not continuous since it is undefined at $x = 0$.

Figure 45: Graphic view for the warning.



Next section, we will be discussing uniform continuity.

4.4 Uniform Continuity

Go back to Table of Contents. Please click TOC

To start this subsection, we need to understand the motivation behind the concept of Uniform Continuity. Consider an example, a map f, which is shown below. This map, f, takes elements from X to Y. Consider two points p and q in X and there will be images for them from the map, say $f(p)$ and $f(q)$. We call this map *isometry* if we have the distance between point p and q preserved after the map. Let us see the definition.

Definition 4.57. (Isometry). Suppose a map $f : X \to Y$. The map f is isometric if

$$\forall p, q \in X,\ d_Y(f(p), f(q)) = d_X(p, q)$$

Figure 46: Graphic view for Isometry.

Then a natural question to ask is: are isometric functions continuous?

Consider an example, $p \in X$, $\forall \epsilon > 0$ $\exists \delta > 0$ we have $d(f(x), f(p)) < \epsilon$ if $d(x, p) < \delta$. Simply take $\delta = \epsilon$, that is, δ only depends on ϵ, we can show that the function f is continuous.

4.5 Lipschitz Function

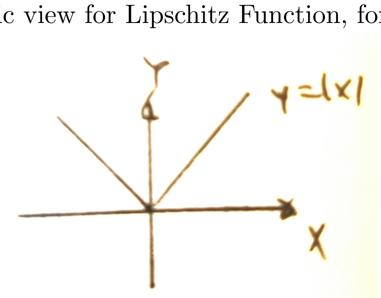
Go back to Table of Contents. Please click TOC

For another example, let us say we have $d(f(p), f(q)) = d(p, q)$. What if we switch "$=$" to become "$\leq$"? In this case, we would have $d(f(p), f(q)) \leq c$ for some constant c. This is the famous Lipschitz Function.

Consider a Lipschitz Function, notice that every isometry is a Lipschitz Function. Take a simple example, $f(x) = |x|$, as shown below. One can use triangle inequality to show the characteristic. Suppose $f \in C$ with f bounded. Then f is a Lipschitz Function.

An intuitive question would be: does Lipschitz imply continuity? $\forall b \in X$ take $\epsilon > 0$, $\exists \delta > 0$ such that $d(f(p), f(q)) < \epsilon$ if $d(x, p) < \delta$. Simply take $\delta = \frac{\epsilon}{c}$ and again δ is independent of p (the location of p does not matter) while $c > 0$.

Figure 47: Graphic view for Lipschitz Function, for example, $y = |x|$.



Remark 4.58. We can also refer to page 151, section 12.3, of textbook [7] for an alternative definition. A function f is *Lipschitz continuous* with *Lipschitz constant* L_f on I, if there is a (necessarily nonnegative) constant L_f such that $|f(x_1) - f(x_2)| \leq L_f |x_1 - x_2|, \forall x_1, x_2 \in I$.

$\square$

4.6 Hölder Function

Go back to Table of Contents. Please click TOC

We can also introduce another type of continuity, Hölder Function. For $\alpha \in (0,1)$, let us say we have $d(f(p), f(q)) \leq (d(p,q))^{\alpha}$. Consider the example, $f(x) = |x|^{\alpha}$, given that $f : \mathbb{R} \to \mathbb{R}$, which would satisfy to be a Hölder Function.

Again, we ask the question would an α-Hölder Function be continuous? $\forall p \in X$, $\forall \epsilon > 0$, $\exists \delta > 0$ such that $d(f(p), f(q)) < \epsilon$ if $d(p,x) < \delta$. Simply choose $(\frac{\epsilon}{c})^{1/\alpha}$.

Remark 4.59. We can refer to page 52 of textbook [8] for an alternative definition. Let x_0 be a point in $\mathbb{R}^n$ and f a function defined on a bounded set D containing x_0. If $0 < \alpha < 1$, we say that f is Hölder continuous with exponent α at x_0 if the quantity $[f]_{\alpha;x_0} = \sup\limits_{D} \frac{|f(x)-f(x_0)|}{|x-x_0|^{\alpha}}$.

□

Remark 4.60. Note that so far we have the following relationships: Continuously differentiable $\subseteq$ Lipschitz continuous $\subseteq$ α-Hölder continuous $\subseteq$ Uniformly continuous $\subseteq$ Continuous. Important!

□

Now we introduce *Oscillation.*

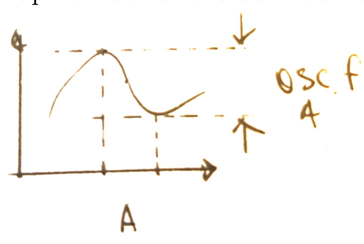
Definition 4.61. (Oscillation). An oscillation is defined as the following, suppose $A \subset X$,

$$\underset{A}{osc}\, f = \sup_{x,y \in A} d(f(x), f(y))$$

which is shown below in the graph.

Referring to the graph, consider f a function. In some interval in the domain of f, we have the image of the map move up and down between two values. A collection of distances that can bound these values would be an upper bound of the distances of any two images mapped from domain. An oscillation is defined to be the least upper bound of that collection of distances.

Figure 48: Graphic view for the definition of Oscillation.



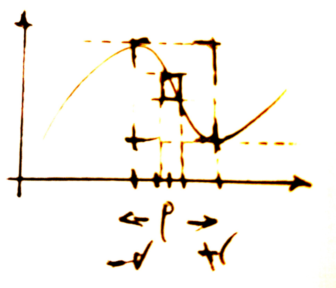
Also consider the following argument.

Theorem 4.62. *A map $f : X \to Y$ to be continuous at $p \in X$ if and only if*

$$\underset{B_r(p)}{osc}\, f \to 0 \text{ as } r \to 0^+.$$

Referring to the graph, we simply have a map, f, taking values from X to Y. For any point p, we can draw a ball with radius r. Given some r, we have a range of maps, and can be represented by some kind of notation of f, which is $\underset{B_r(p)}{osc} f$. Then we take r, no matter how small, we would have the oscillation goes to 0. If all of these are satisfied, we say f is continuous.

Figure 49: Graphic view for the theorem.



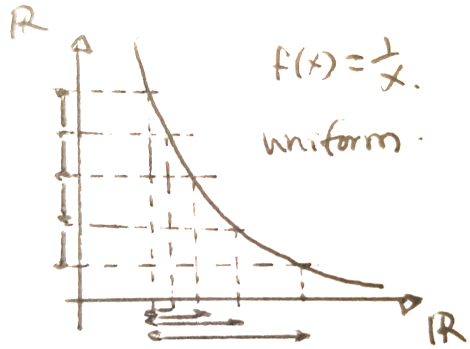
Definition 4.63. (Modulus of Continuity). At p, suppose we have $\omega_p : (0, \infty) \to (0, \infty)$ such that $\omega_p(r) \geq \underset{B_r(p)}{osc} f$ and $\omega_p(r) \to 0$ as $r \to 0$. Then we say f is continuous at p and we also have ω_p measures the convergence, that is, $f(x) \to f(p)$ as $x \to p$.

Example 4.64. For example, f is Lipschitz, $d(f(p), f(q)) \leq cd(p, q)$, then $\omega_p(r) = cr$. For another example, f is α-Hölder function and $d(f(p), f(q)) \leq cd(p, q)$, then $\omega_p(r) = cr^{\alpha}$.

$\square$

A natural question is: for f, a continuous map, will there exists a common modulus of continuity for all $p \in X$? The answer is *no*. Consider the function $y = \frac{1}{x}$, which is shown below. As $x \to 0$, the modulus of continuity becomes worse, that is, you cannot choose a uniform δ.

Figure 50: Graphic view for the common modulus.



Definition 4.65. (Uniform Continuity). Suppose there is a map, $f : X \to Y$. We say it is uniformly continuous if

(1) $\forall p, \exists \omega(r)$ a common modulus of continuity, or

(2) $\forall \epsilon > 0$, $\exists \delta > 0$ (independent of p), there is $d(f(p), f(q)) < \epsilon$ if $d(p,q) < \delta$.

Common examples can be Lipschitz and Hölder.

5 §Differentiation§

Go back to Table of Contents. Please click TOC

5.1 Derivative

Go back to Table of Contents. Please click TOC

This section we discuss the derivative of real functions.

Definition 5.1. Let f be defined and real-valued on $[a,b]$. For any $x \in [a,b]$ from the quotient

$$\phi(t) = \frac{f(t) - f(x)}{t - x}(a < t < b, t \neq x),$$

and define

$$f'(x) = \lim_{t \to x} \phi(t),$$

Then we can discuss function f and function f' whose domain is the set of points x at which the limit exists; f' is called the derivative of f.

Example 5.2. For example, consider $f(x) = 1$, and $f'(x_0) = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} = 0$ since the numerator goes to zero.

□

Example 5.3. For another example, consider $f(x) = mx + b$, we have $f'(x_0) = \lim_{h \to 0} \frac{(m(x_0+h)+b)-(mx_0+b)}{h} = \lim_{h \to 0} \frac{mh}{h} = m$.

$\square$

Example 5.4. For example, consider $f(x) = \frac{1}{x}$, assuming $x_0 \neq 0$, we have $f'(x) = \lim_{h\to 0} \frac{\frac{1}{x_0+h} - \frac{1}{x_0}}{h} = \lim_{h\to 0} \frac{(x_0-x_0-h)/((x_0+h)x_0)}{h} = \lim_{n\to 0} \frac{-h}{x_0(x_0+h)h} = \frac{-1}{x_0}$.

$\square$

Theorem 5.5. *Let f be defined on $[a,b]$. If f is differentiable at a point $x \in [a,b]$, then f is continuous at x.*

Proof: The goal is to show that $f(x+h) \to f(x)$ as $h \to 0$. Then we need $f(x+h) - f(x) = \frac{f(x+h)-f(x)}{h} h \to f'(x) \times 0 \to 0$ as $h \to 0$.

Q.E.D.

Theorem 5.6. *Suppose f and g are defined on $[a,b]$ and are differentiable at a point $x \in [a,b]$. Then $f+g$, fg, and f/g are differentiable at x,*

(a) Linearity. $(f+g)'(x) = f'(x) + g'(x)$;

(b) Product Rule. $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$;

(c) Quotient Rule. $\left(\frac{f}{g}\right)'(x) = \frac{g(x)f'(x)-g'(x)f(x)}{g^2(x)}$*, assume* $g(x) \neq 0$.

Proof: We present the proof accordingly:

(a) Let $h = fg$. Then $h(t) - h(x) = f(t)[g(t) - g(x)] + g(x)[f(t) - f(x)]$. If we divide both sides by $t - x$, we would have had $[f(t) - f(x)] \to 0$. That is, we would have $f'(x)g(x) + f(x)g'(x)$ left, which completes the proof.

(b) Recall that $f(x+h)g(x+h) - f(x)g(x) = 0 \Rightarrow (f(x+h) - f(x))g(x+h) + f(x)(g(x+h) - g(x)) \sim (df)g + f(dg)$. Where did it come from? Then we divide the equation both sides by h and let $h \to 0$. Thus

$$\begin{aligned} &\frac{f(x+h)g(x+h) + f(x)g(x)}{h} \\ = &\frac{f(x+h) - f(x)}{h} g(x+h) + f(x)\frac{g(x+h) - g(x)}{h} \\ &\Rightarrow f'(x)g(x) + f(x)g'(x) \end{aligned}$$

as $h \to 0$ since g is continuous at x.

(c) Let $h(x) = \frac{f}{g}(x) \Leftrightarrow f(x) = g(x)h(x)$. Formally we have $f'(x) = g'(x)h(x) + g(x)h'(x)$. Solving $h'(x)$, we have

$$\begin{aligned} h'(x) &= \frac{f'(x) - g'(x)\frac{f(x)}{g(x)}}{g(x)} \\ &= \frac{f'(x)g(x) - g(x)f(x)}{g(x)^2} \end{aligned}$$

Notice here that $h'(x)$ is not well defined since we may have the case that the denominator is zero. In that case, we need to evaluate

$$\begin{aligned} &\lim_{h\to 0} \frac{1}{h}\left[\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}\right] \\ = &\lim_{h\to 0} \frac{1}{h}\left[f(x+h)g(x) - g(x+h)f(x)\right] \frac{1}{g(x+h)g(x)} \\ = &\lim_{h\to 0} \left[\frac{1}{h} f(x+h)g(x) - \frac{1}{h} g(x+h)f(x)\right] \frac{1}{g(x+h)g(x)} \end{aligned}$$

The numerator would tend to zero as $h \to 0$, which completes the proof.

Q.E.D.

Theorem 5.7. *Suppose f is continuous on $[a,b]$, $f'(x)$ exists at some point $x \in [a,b]$, g is defined on an interval I which contains the range of f, and g is differentiable at the point $f(x)$. If*

$$h(t) = g(f(t)), a \leq t \leq b,$$

then h is differentiable at x, and

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

Example 5.8. Consider $(fg)'' = (f'g + fg')' = f''g + 2f'g' + fg''$, which is similar as $(a+b)^2 = a^2 + 2ab + b^2$.

□

Example 5.9. Prove that

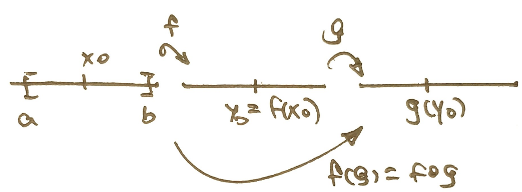
$$(fg)^{(k)} = \sum j = 0^k \binom{k}{j} f^{(i)} g^{k-j}$$

which is the same idea as the Newton's identity (by induction).

□

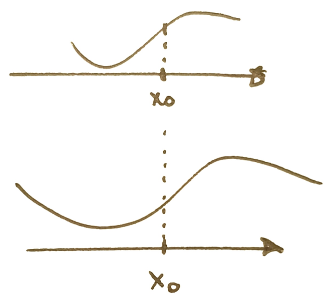
Consider the chain rule, we have h, f, and g. We want the following, $f'(x) = g'(x)f'(x)$.

Figure 51: Graphic view for the Chain Rule.



Take $f(x) = \frac{x}{2}$ as an example, $h(x) = g \circ f(x) = g(\frac{x}{2})$. Referring to the following graph, we know that the geometrical sketch will be stretched from the multiplying of half, i.e. $\frac{1}{2}$.

Figure 52: Graphic view for the example $f(x) = \frac{x}{2}$. The map is stretched wider.



Then we present the following attempt (although wrong, yet worth noticing), remember we defined above that $h(x) = g \circ f(x) = g(\frac{x}{2})$, say $x/2$ as an example,

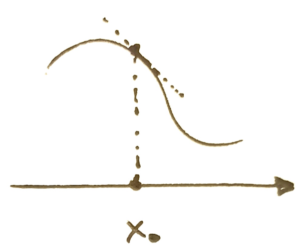
$$\begin{aligned}
\frac{h(x+h)-h(x)}{h} &= \frac{g(f(x+h))-g(f(x))}{h} \\
&= \frac{g(f(x+h))-g(f(x))}{f(x+h)-f(x)} - \frac{f(x+h)-f(x)}{h} \\
&= g'(y)f'(x)
\end{aligned}$$

The problem is that we might be dividing zero. To correct this, we introduce a definition. We need to define the derivative without using dangerous denominations.

Definition 5.10. Newton's Linear Approximation. Suppose there is a function f. The tangent line at any point, say x_0, is the approximation of f. That is,

$$y = f'(x_0)(x - x_0) + f(x_0)$$

Figure 53: Graphic view for the *Newton's Linear Approximation.*



That is, we consider slope m and at point (a, b), we have $y - b = m(x - a) \Rightarrow y = m(x - a) + b$.

Proposition 5.11. *Consider* $u(x - x_0) \to 0$ *as* $x \to x_0$.

Proof: Start with u, we have $u(x-x_0) = \frac{f(x)-f(x_0)}{x-x_0} = f(x_0)$ as $x \to x_0$. Then we have $u(x-x_0) \to 0$. The advantage of this approach is that there is no denominator.

Recall Chain Rule. Suppose we have f differentiable at x_0, g differentiable at $y_0 = f(x_0)$, and $h = f \circ g$. Then we have h differentiable at x_0 and $h'(x_0) = g'(x_0)f'(x_0)$.

Proof: $f(x) = f(x) + f'(x_0)(x-x_0) + (x-x_0)u(x-x_0)$ and $g(y) = g(y_0) + g'(y_0)(y-y_0) + (y-y_0)v(y-y_0)$. Notice that $u(x-x_0) \to 0$ as $x \to x_0$ and $v(y-y_0) \to 0$ as $y \to y_0$. Take $y = f(x)$, we have $g(f(x)) = g(y_0) + g'(y_0)(f(x) - f(x_0)) + (f(x) - f(x_0))v(y-y_0)$, which implies that $h(x) = h(x_0) + g'(y_0)(f'(x_0)(x-x_0) + (x-x_0)u(x-x_0)) + (f(x) - f(x_0))v(y-y_0)$. Then as $x \to x_0$, $g'(x_0)u(x-x_0) \to 0$. Then $y = f(x_0) \to f(x_0) = y_0 \Rightarrow v(y-y_0) \to 0$ as $x \to x_0$.

Therefore, the second and third term cancel and we conclude the argument.

Example 5.12. An exercise can be the following. Consider

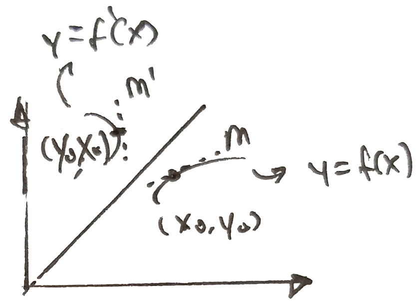
$$f(x) = \begin{cases} x^2 sin(\frac{1}{x}) & , \quad x \neq 0 \\ 0 & , \quad x = 0 \end{cases}$$

□

Definition 5.13. (Inverse Function). Suppose $f : X \to Y$ is well-defined and a bijection. Then there exists $f^{-1} : Y \to X$ such that $f \circ f^{-1} = \mathbb{I}_Y$ and $f^{-1} \circ f = \mathbb{I}_Y$. If f and f^{-1} are differentiable, then $f \circ f^{-1}(y) = y$.

Applying $\frac{d}{dy}|_{y=y_0} : f^{(1)}(x_0)f^{-1(1)}(y_0) = 1$, then we have $f^{-1}(y_0) = \frac{1}{f'(x_0)}$. Geometrically, we have some function f. Let us say f is well-defined and $y = f(x)$. Then the inverse of f, which is f^{-1} is well-defined and mirror reflected by the line $y = x$. Thus, some arbitrary point, say (x_0, y_0) will be (y_0, x_0). The slope at the point (x_0, y_0) will be m and will be reflected to $m' = \frac{1}{m}$.

Figure 54: Graphic view *Inverse Function*.



This concept leads to the following theorem.

Theorem 5.14. *(Inverse Function Theorem). Suppose $f : (x_0 - \delta, x_0 + \delta) \to Y \subseteq \mathbb{R}$ and a bijection. Also suppose the map f is differentiable at x_0 and $f^{(1)}(x_0) \neq 0$. Then f^{-1} is continuous at $y_0 = f(x_0)$. Then f^{-1} is differentiable at y_0.*

Proof: The goal is the following. We want to show $\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} \to \frac{1}{f(x_0)}$ as $y \to y_0$. It suffices to show that if $y_n \neq y_0 \to y_0$ as $n \to \infty$ then $\frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} \to \frac{1}{f^{-1}(x_0)}$ as $n \to \infty$. Consider f a bijection. Then we have $x_n = f^{-1}(y_n)$ or $y_n = f(x_n)$. Then $\frac{f^{-1}(Y_n) - f^{-1}(y_0)}{y_n - y_0} = \frac{x_n - x_0}{f(x_n) - f(x_0)} = \left(\frac{f(x_n) - f(x_0)}{x_n - x_0)}\right)^{-1}$. It is important to notice that f^{-1} is continuous at y_0 and this gives us $f^{-1}(y_n) \to f^{-1}(y_0)$ which is also $x_n \to x_0$.

Q.E.D.

Example 5.15. Suppose $f : (x_0 - \delta, x_0 + \delta) \to Y \subseteq \mathbb{R}$ and a bijection. Also suppose f is continuous. Then we have f^{-1} a continuous function. For this problem, simply take $f : [x_0 - \delta/2, x_0 + \delta/2] \to Z \subseteq Y \subseteq \mathbb{R}$ continuous. Notice that the interval is compact. This gives us that f^{-1} is continuous. Then we take $\delta/3$, $\delta/4$, ... δ/N such that $N \to \infty$, which will show the complete proof.

□

Example 5.16. Suppose $f(x) = x^n$ and $x > 0$ and $f^{-1} = \frac{1}{n}$. Then we have $f^{-1}(x) = \frac{1}{n(x^{\frac{1}{n}})^{n-1}} = \frac{1}{n}x^{\frac{1}{n}-1}$.

□

Example 5.17. We can also compute x^α such that $\alpha \in \mathbb{Q}$. Then take $\alpha = \frac{m}{n}, n \neq 0$. Also the same for $\alpha \in \mathbb{R}$. We use chain rule and power rule.

□

5.2 Mean Value Theorem

Go back to Table of Contents. Please click TOC

Definition 5.18. Let f be a real function defind on a metric space X. We say that f has a local maximum at a point $p \in X$ if there exists $\delta > 0$ such that $f(q) \leq f(p)$ for all $q \in X$ with $d(p, q) < \delta$.

Remark 5.19. Local minima are defined likewise.

□

Theorem 5.20. *Let f be a defined on $[a, b]$; if f has a local maximum at a point $x \in (a, b)$, and if $f'(x)$ exists, then $f'(x) = 0$.*

Theorem 5.21. *If f and g are continuous real functions on $[a, b]$ which are differentiable in (a, b), then there is a point $x \in (a, b)$ at which*

$$[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x).$$

Note that differentiability is not required at the endpoints.

Theorem 5.22. *If f is a real continuous function on $[a,b]$ which is differentiable in (a,b), then there is a point $x \in (a,b)$ at which*

$$f(b) - f(a) = (b-a)f'(x).$$

Proposition 5.23. *Suppose f is differentiable in (a,b),*

(a) If $f'(x) \geq 0$ for all $x \in (a,b)$, then f is monotonically increasing.

(b) If $f'(x) = 0$ for all $x \in (a,b)$, then f is constant.

(c) If $f'(x) \leq 0$ for all $x \in (a,b)$, then f is monotonically decreasing.

Proof: We prove Monotonicity accordingly. For (a), by contradiction, we assume $f(x) > f(y)$ for some $a \leq x < y \leq b$. By M.V.T., $\exists c \in (x,y)$ such that $f'(c) = \frac{f(y)-f(x)}{y-x} < 0$, which is a contradiction. For (b), by contradiction, we assume that $f(x) < f(y)$ for some $a \leq x < y \leq b$. By M.V.T., $\exists c \in (x,y)$ such that $f'(c) = \frac{f(y)-f(x)}{y-x} > 0$, which leads to contradiction. The last proof is the same.

Q.E.D.

Theorem 5.24. *Suppose f is a real differentiable function on $[a,b]$ and suppose $f'(a) < \lambda < f'(b)$. Then there is a point $x \in (a,b)$ such that $f'(x) = \lambda$.*

Note a similar result hold for $f'(a) > f'(b)$.

Proof:

Q.E.D.

A particular case is to let $f(a) = f(b)$, then we would get local maximum and local minimum, which leads to the following definition.

Definition 5.25. $f : [a,b] \to \mathbb{R}$ has a local maximum (or minimum) at $c \in [a,b]$ if $\exists \delta > 0$ such that $f(x) \leq f(c)$, $\forall x \in (c-\delta, c+\delta) \cap [a,b]$.

Proposition 5.26. *$f : [a,b] \to \mathbb{R}$ continuous and differentiable in (a,b) has a local maximum (or minimum) at $c \in (a,b)$. This implies that $f'(c) = 0$.*

Proof: Consider f has local max at c, then $\exists c > 0$ such that $f(x) \leq f(c)$, $\forall x \in (c-\delta, c+\delta)$ by using $c \in (a,b)$ to get $x \in (c-\delta, c+\delta)$.

If $x \to c^+$ (i.e. $x > c$), then $\frac{f(x)-f(c)}{x-c} \leq 0$. This implies that $f'(c) \leq 0$. If $x \to c^-$ (i.e. $x < c$), then $\frac{f(x)-f(c)}{x-c} \geq 0$. This implies that $f'(c) \geq 0$.

Hence, we have $f'(c) = 0$.

Q.E.D.

Now a question needs to be asked. Can we have some $c = a$ (or $c = b$) for any $c \in [a,b]$? The answer to this question leads to Rolle's Theorem.

Theorem 5.27. *(Rolle's Theorem). Suppose $f : [a,b] \to \mathbb{R}$ continuous and f differentiable in (a,b) and $f(a) = f(b)$. This implies that there exists $c \in (a,b)$ such that $f'(c) = 0$.*

Proof: Discuss all cases accordingly.

If $f(x) = f(a) = f(b)$ for all $x \in (a,b)$, then we have $\forall c \in (a,b)$, $f(c) = 0$, which is done. Else if $f(x)$ has a *max* or *min* in (a,b), since E.V.T. at such value c (which we obtained from previous proof), we have $f'(c) = 0$.

Q.E.D.

Theorem 5.28. *(Mean Value Theorem). Suppose* $f : [a, b] \to \mathbb{R}$ *and* f *continuous. Also suppose that* f *is differentiable in* (a, b). *There exists* $c \in (a, b)$ *such that* $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Proof: The first strategy takes the secant line of any two points in the interval. Then one is able to move the secant line up or down so that it just touches the function f. Applying Rolle's Theorem, one can show that at this point that the secant line just touches f there is $f(c) = f(a) = f(b)$.

The second strategy is to let $g(x) = f(x) - [m(x - a) + f(a)]$ where $m = \frac{f(a)-f(b)}{b-a}$. Then take $g(a) = g(b)$, and by Rolle's Theorem, $\exists c \in (a, b)$ such that $g'(c) = 0$, which implies $f'(c) - m = 0$. Thus, we have $f'(c) = m$.

Q.E.D.

A consequence is the following:

Theorem 5.29. *If a function is Lipschitz continuous, then the derivative of the function is bounded.*

Given a function f is Lipschitz continuous. We have $f : [a, b] \to \mathbb{R}$ such that $|f'| \leq M$, which implies that $\frac{|f(x)-f(y)|}{|x-y|} \leq M$ for all $x = y \in [a, b]$.

Assume for $x < y$. By M.V.T., there exists some $c \in \mathbb{R}$ such that $f'(c) = \frac{f(x)-f(y)}{x-y}$ for some $c \in (x, y)$. Thus, we have $f' \in (-M, M)$.

For $x > y$, we have $f'(x_0) = \lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h}$, which is in $[-M, M]$.

Q.E.D.

Definition 5.30. (Convex). Let f be a function defined on an interval $[a, b]$. The function f is convex on $[a, b]$ if for each interval $[\alpha, \beta] \subseteq [a, b]$, the inequality

$$f\Big(\theta\alpha + (1-\theta)\beta\Big) \leq \theta f(\alpha) + (1-\theta) f(\beta)$$

for some $\theta \in [0, 1]$.

Example 5.31. For an example, let us look at $f(x) = x^2$. Let $[c, d] \subseteq I$ and $0 \leq \theta \leq 1$. We have

$$\begin{aligned}
f\Big((1-\theta)c + \theta d\Big) &= (1-\theta)^2 c^2 + 2\theta(1-\theta)cd + \theta^2 d^2 \\
&= (1-\theta)c^2 - \theta(1-\theta)c^2 + 2\theta(1-\theta)cd + \theta^2 d^2 \\
&= (1-\theta)c^2 + \theta(1-\theta)c(2d-c) + \theta^2 d^2 - \theta d^2 + \theta d^2 \\
&= (1-\theta)c^2 - \theta(1-\theta)(c^2 - 2cd + d^2) + \theta d^2 \\
&= (1-\theta)c^2 - \theta(1-\theta)(c-d)^2 + \theta d^2 \\
&\leq (1-\theta)c^2 + \theta d^2
\end{aligned}$$

which shows that f is a convex function.

□

Remark 5.32. We have discussed convexity in previous sections. Please go to Theorem 1.17 on page 10 of this notes to recall the initial introduction of convexity.

□

5.3 L'Hôpital's Rule

Go back to Table of Contents. Please click TOC

Theorem 5.33. *Suppose f and g are real and differentiable in (a, b), and $g'(x) \neq 0$ for all $x \in (a, b)$, where $-\infty \leq a < b \leq +\infty$. Suppose*

$$\frac{f'(x)}{g'(x)} \to A \text{ as } x \to a.$$

If $f(x) \to 0$ and $g(x) \to 0$ as $x \to a$, or if $g(x) \to +\infty$ as $x \to a$, then $\frac{f(x)}{g(x)} \to A$ as $x \to a$.

Proof: First consider the case $-\infty \leq A < +\infty$. Choose $q \in \mathbb{R}$ such taht $A < q$, and then choose r such that $A < r < q$. Suppose we have $f'(x)/g'(x) \to A$ as $x \to a$.Then there is a point $c \in (a, b)$ such that $a < x < c$ implies

$$\frac{f'(x)}{g'(x)} < r$$

If $a < x < y < c$, then Theorem 5.9 from text [10] shows that there is a point $t \in (x, y)$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < r.$$

Suppose $f(x) \to 0$ and $g(x) \to 0$ as $x \to a$. Let $x \to a$, we have

$$\frac{f(y)}{g(y)} \leq r < q, \ a < y < c.$$

Next, suppose $g(x) \to +\infty$ as $x \to a$. We keep y fixed in $\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(t)}{g'(t)} < r$, and simply choose a point $c_1 \in (a, y)$ such that $g(x) > g(y)$ and $g(x) > 0$ if $a < x < c_1$. Multiplying $\frac{f(y)}{g(y)} \leq r < q, \ a < y < c$ by $[g(x) - g(y)]/g(x)$, we obtain

$$\frac{f(x)}{g(x)} < r - r\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)}, \ a < x < c_1.$$

As $x \to a$, $g(x) \to +\infty$ as $x \to 0$ tells us that there is a point $c_2 \in (a, c_1)$ such that

$$\frac{f(x)}{g(x)} < q, \ a < x < c_2.$$

Summing up everything,"$\frac{f(y)}{g(y)} \leq r < q, \ a < y < c$" and "$\frac{f(x)}{g(x)} < q, \ a < x < c_2$" tell us that for any q, subject only to the condition $A < q$, there is a point c_2 such that $f(x)/g(x) < q$ if $a < x < c_2$. In the same manner, if $-\infty < A \leq +\infty$, and p is chosen so that $p < A$, we can find a point c_3 such that

$$p < \frac{f(x)}{g(x)}, \ a < x < c_3$$

and $f(x)/g(x) \to A$ as $x \to a$ would follow. Note that this proof from text [10] would seem more complicated because this approach puts everything together. Later after an example, we will present an alternative proof of L'Hôpital's Rule, as explained in the lecture.

Q.E.D.

Remark 5.34. The analogous statement is of course also true if $x \to b$, or if $g(x) \to -\infty$.

□

Remark 5.35. Theorem 5.9 of text [10] (Theorem 5.21 on page 76 in this section) says that: if f and g are continuous real functions on $[a, b]$ which are differentiable in (a, b), then there is a point $x \in (a, b)$ at which $[f(b) - f(a)]g'(x) = [g(b) - g(a)]f'(x)$. Note that differentiability is not required at the endpoints.

□

Example 5.36. As an application, we see the following example. L'Hôpital's Rule helps us to compute indefinite limit. For example we want to evaluate $\lim\limits_{x\to 0} \frac{7x - sin(x)}{x^2 + sin(3x)}$ at $x = 0$. The denominator would be 0, and this calculation would seem meaningless. However, using L'Hôpital's Rule, we have

$$\begin{aligned} \left.\frac{7x-sin(x)}{x^2+sin(3x)}\right|_{x=0} &= \left.\frac{[7x-sin(x)]'}{[x^2+sin(3x)]'}\right|_{x=0} \\ &= \left.\frac{7-cos(x)}{2x+3cos(3x)}\right|_{x=0} \\ &= \frac{7-cos(0)}{2\times 0+3cos(3\times 0)} \\ &= 2 \end{aligned}$$

which is a lot easier and actually meaningful.

□

For a better understanding, let us present an alternative proof of the L'Hôpital's Rule.

Proof: We first prove a special case under a more restrict condition. Then we present generalized case.

For special case, suppose f and g are both continuous and differentiable at a $c \in \mathbb{R}$, and also suppose $f(c) = g(c) = 0$, and that $g'(c) \neq 0$. Then we have

$$\begin{aligned} \lim_{x\to c}\frac{f(x)}{g(x)} &= \lim_{x\to c}\frac{f(x)-0}{g(x)-0} = \lim_{x\to c}\frac{f(x)-f(c)}{g(x)-g(c)} \\ &= \lim_{x\to c}\frac{\frac{f(x)-f(c)}{x-c}}{\frac{g(x)-g(c)}{x-c}} = \frac{\lim\limits_{x\to c}\frac{f(x)-f(c)}{x-c}}{\lim\limits_{x\to c}\frac{g(x)-g(c)}{x-c}} \\ &= \frac{f'(c)}{g'(c)} = \lim_{x\to c}\frac{f'(x)}{g'(x)}, \end{aligned}$$

and hence $\lim\limits_{x\to c}\frac{f(x)}{g(x)} = \lim\limits_{x\to c}\frac{f'(x)}{g'(x)}$.

Now we show L'Hôpital's Rule in general case. Let f and g be two continuous and differentiable functions. Let $\mathcal{I}$ be the open interval in the hypothesis with c. Assuming $g'(x) \neq 0$, $\mathcal{I}$ can be chosen very small so that g is nonzero on $\mathcal{I}$. For any $x \in \mathcal{I}$, define $m(x) = inf\frac{f'(\xi)}{g'(\xi)}$ and $M(x) = sup\frac{f'(\xi)}{g'(\xi)}$ as ξ can be any value between x and c. Mean Value Theorem ensures that for any two distinct points x and y in $\mathcal{I}$ there exists a ξ between x and y such that $\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(\xi)}{g'(\xi)}$. Then, as a consequence,

$m(x) \leq \frac{f(x)-f(y)}{g(x)-g(y)} \leq M(x)$ for all choices of distinct x and y in the interval. The value $g(x)-g(y)$ is always nonzero for distinct x and y in the interval, for if it was not, the mean value theorem would imply the existence of a p between x and y such that $g'(p) = 0$.

Here we discuss two cases:

First case is for $\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0$. For any x in the interval $\mathcal{I}$, and any point y between x and c, there is

$$m(x) \leq \frac{f(x)-f(y)}{g(x)-g(y)} = \frac{\frac{f(x)}{g(x)} - \frac{f(y)}{g(x)}}{1 - \frac{g(y)}{g(x)}} \leq M(x)$$

and therefore as y approaches c, $\frac{f(y)}{g(x)}$ and $\frac{g(y)}{g(x)}$ will be zero, so we are left with

$$m(x) \leq \frac{f(x)-f(y)}{g(x)-g(y)} = \frac{\frac{f(x)}{g(x)} - 0}{1-0} \leq M(x)$$

and thus

$$m(x) \leq= \frac{f(x)}{g(x)} \leq M(x)$$

which completes the first case.

Second case we discuss $\lim_{x\to c} |g(x)| = \infty$. For every x in the interval $\mathcal{I}$, define $S_x = \{y | y \ between \ x \ and \ c\}$. For every point y between x and c, we have

$$m(x) \leq \frac{f(y)-f(x)}{g(y)-g(x)} = \frac{\frac{f(y)}{g(y)} - \frac{f(x)}{g(y)}}{1 - \frac{g(x)}{g(y)}} \leq M(x)$$

As y goes to c, both $\frac{f(x)}{g(y)}$ and $\frac{f(x)}{g(y)}$ go to zero, and therefore

$$m(x) \leq \underset{y\in S_x}{lim\ inf} \frac{f(y)}{g(y)} \leq \limsup_{y\in S_x} \frac{f(y)}{g(y)} \leq M(x).$$

The $\limsup$ and $\liminf$ are necessary since the existence of the limit of f/g has not yet been established. Then we recall the fact that

$$\lim_{x\to c} m(x) = \lim_{x\to c} M(x) = \lim_{x\to c} \frac{f'(x)}{g'(x)} = L$$

and

$$\lim_{x\to c} \left(\liminf_{y\in S_x} \frac{f(y)}{g(y)} \right) = \underset{x\to c}{lim\ inf} \frac{f(x)}{g(x)}$$

and

$$\lim_{x\to c} \left(\liminf_{y\in S_x} \frac{f(y)}{g(y)} \right) = \underset{x\to c}{lim\ sup} \frac{f(x)}{g(x)}$$

In first case, the squeeze theorem, establishes that $\lim_{x\to c} \frac{f(x)}{g(x)}$ exists and is equal to L. In second case, the squeeze theorem again asserts that $\liminf_{x\to c} \frac{f(x)}{g(x)} = \limsup_{x\to c} \frac{f(x)}{g(x)} = L$, and so the limit $\lim_{x\to c} \frac{f(x)}{g(x)}$ exists and is equal to L, which is proven.

Q.E.D.

5.4 Taylor's Theorem

Go back to Table of Contents. Please click TOC

Theorem 5.37. *Suppose f is a real function on $[a,b]$, n is a positive integer, $f^{(n-1)}$ is continuous on $[a,b]$, $f^{(n)}(t)$ exists for every $g \in (a,b)$. Let α, β be distinct points of $[a,b]$, and define*

$$P(t) = \sum_{k=1}^{n-1} \frac{1}{k!} f^{(k)}(\alpha)(t-\alpha)^k.$$

Then there exists a point x between α and β such that

$$f(\beta) = P(\beta) + \frac{1}{n!} f^{(n)}(x)(\beta-\alpha)^n.$$

Proof: The text [10] provided the following proof. Let M be the number defined by

$$f(\beta) = P(\beta) + M(\beta-\alpha)^n$$

and put

$$g(t) = f(t) - P(t) - M(t-\alpha)^n, \ a \le t \le b$$

that is, we have to show that $n!M = f^{(n)}(x)$ for some x between α and β. By

$$P(t) = \sum_{k=1}^{n-1} \frac{1}{k!} f^{(k)}(\alpha)(t-\alpha)^k.$$

and

$$g(t) = f(t) - P(t) - M(t-\alpha)^n, \ a \le t \le b$$

we have

$$g^{(n)}(t) = f^{(n)}(t) - n!M, \ a < t < b.$$

Hence, we just need to show that $g^{(n)}(x) = 0$ for some x between α and β. Since

$$P^{(k)}9\alpha) = f^{(k)}(\alpha), \ k = 0, ..., n-1,$$

we have

$$g(\alpha) = g'(\alpha) = \cdots = g^{(n-1)}(\alpha) = 0.$$

Our choice of M shows that $g(\beta) = 0$, so that $g'(x_1) = 0$ for some x_1 between α and β, by the M.V.T. Since $g'(\alpha) = 0$, we conclude that $g''(x_2) = 0$ for some x_2 between α and x_1. We repeat this n steps and conclude that $g^{(n)}(x_n) = 0$ for some x_n between α and x_{n-1}, that is, between α and β.

Q.E.D.

Theorem 5.38. *Suppose f is a continuous mapping of $[a,b]$ into R^k and f is differentiable in (a,b). Then there exists $x \in (a,b)$ such that*

$$|f(b) - f(a)| \le (b-a)|f'(x)|.$$

6 §Riemann-Stieltjes Integral§

Go back to Table of Contents. Please click TOC

6.1 Riemann Integral

Go back to Table of Contents. Please click TOC

The motivation is to obtain a solution for a differential equation, say $y' = f(t) \sim \frac{y(t+\triangle t)-y(t)}{\triangle t} \simeq f(t)$. That is, we want $h(t+\triangle t) \sim y(t) + f(t)\triangle t$. We have a map, f, defined on $[a,b]$, and suppose there are y_0 at $x = a$, y_1 at $x = t_1$ with $\triangle t = t_1 - a$. We iterate the following:

$$\begin{aligned} y_{k+1} &= f(t_k)\triangle t, \ with\ y_0 = y_0\ known \\ y_{k+1} &= y_k + f(t_k)\triangle t \\ &= (y_{k-1} + f(t_{k-1})\triangle t) + f(t_k)\triangle t \\ &\ \ \ \dots \\ &= y_0 + \sum_{i=0}^{k} f(t_i)\triangle t \end{aligned}$$

with $\sum_{i=0}^{k} f(t_i)\triangle t \sim \int_a^b f(t)dt$, the Riemann Sum.

Remark 6.1. Suppose we know before that $y' = g'$ and that $y(a) = g(a)$ at some a. The goal is to solve $y = g$, which is a unique solution. The reason is that from M.V.T., we know $y = g + c$, c being some constant, and this condition at a will give us $c = 0$.

□

Definition 6.2. Let $[a,b]$ be a given interval. By a partition P of $[a,b]$ we mean a finite set of points $x_0, x_1, ..., x_n$, where

$$a - x_0 \leq x_1 \leq \cdots \leq x_{n-1}x_n = b.$$

We write

$$\triangle x_i = x_i - x_{i-1},\ i = 1, ..., n.$$

Now suppose f is a bounded real function defined on $[a,b]$. Corresponding to each partition P of $[a,b]$ we put

$$M_t = \sup f(x),\ for\ x_{i-1} \leq x \leq x_i,$$

$$m_i = \inf f(x),\ for\ x_{i-1} \leq x \leq x_i,$$

$$U(P,f) = \sum_{i=1}^{n} M_i \triangle x_i,$$

$$L(P,f) = \sum_{i=1}^{n} m_i \triangle x_i,$$

and finally we have

$$\overline{\int_a^b} f dx = \inf U(P,f),$$

$$\underline{\int_a^b} f dx = \sup L(P,f),$$

where the *inf* and the *sup* are taken over all partitions P of $[a,b]$. The left members of $\overline{\int}_a^b f dx = inf\ U(P,f)$ and $\underline{\int}_a^b f dx = sup\ L(P,f)$ are called *upper* and *lower* Riemann integrals over $[a,b]$, respectively.

If the upper and lower integrals are equal, we say that f is *Riemann-integrable* on $[a,b]$, we write $f \in \mathcal{R}$ (that is, $\mathcal{R}$ denotes the set of Riemann-integrable functions), and we denote the common value of $\overline{\int}_a^b f dx = \inf U(P,f)$ and $\underline{\int}_a^b f dx = \sup L(P,f)$ by $\int_a^b f dx$, or by $\int_a^b f(x)dx$.

This is the *Riemann integral* of f over $[a,b]$. Since f is bounded, there exist two numbers, m and M, such that

$$m \leq f(x) \leq M, \ for\ a \leq x \leq b.$$

Hence, for every P, we have

$$m(b-a) \leq L(P,f) \leq U(P,f) \leq M(b-a),$$

so that the numbers $L(P,f)$ and $U(P,f)$ form a bounded set. This shows that the upper and lower integrals are defined for every bounded function f. The question of their equality, and hence the question of the integrability of f, is a more delicate one. From text, we have the following more general definition [10].

Definition 6.3. Let α be a monotonically icnreasing function on $[a,b]$ (since $\alpha(a)$ and $\alpha(b)$ are finite, it follows that α is bounded on $[a,b]$). Corresponding to each partition P of $[a,b]$, we write

$$\triangle\alpha_i = \alpha(x_i) - \alpha(x_{i-1}).$$

It is clear that $\triangle\alpha_i \geq 0$. For any real function f which is bounded on $[a,b]$, we write

$$U(P,f,\alpha) = \sum_{i=1}^{n} M_i \triangle\alpha_i,$$

$$L(P,f,\alpha) = \sum_{i=1}^{n} m_i \triangle\alpha_i,$$

where M_i, m_i have the same meaning as the previous definition, and we define

$$\overline{\int_a^b} f d\alpha = \inf U(P,f,\alpha),$$

$$\underline{\int_a^b} f d\alpha = \sup L(P,f,\alpha),$$

the *inf* and *sup* again being taken over all partitions.

If the left members of the two equations above are equal, we denote their common value by

$$\int_a^b f d\alpha$$

or by

$$\int_a^b f(x) d\alpha(x).$$

This is the *Riemann-Stieltjes integral* (or simply the Stieltjes integral) of f with respect to α, over $[a,b]$.

Definition 6.4. We say that the partition P^* is a refinement of P if $P^* \supset P$ (tat is, if every point of P is a point of P^*). Given two partitions, P_1 and P_2, we say that P^* is their common refinement if $P^* = P_1 \cup P_2$.

Theorem 6.5. *If P^* is a refinement of P, then*

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

and

$$U(P^*, f, \alpha) \leq U(P, f, \alpha).$$

Proof: To prove limit sum, L, suppose first that P^* contains just one point more than P¿ Let this extra point be x^*, and suppose $x_{i-1} < x^* < x_i$, where x_{i-1} and x_i are two consecutive points of P. Write

$$w_1 = \inf f(x),\ x_{i-1} \leq x \leq x^*,$$

$$w_2 = \inf f(x),\ x^* \leq x \leq x_i.$$

Observe $w_1 \geq m_i$ and $w_2 \geq m_i$, where $m_i = \inf f(x),\ x_{i-1} \leq x \leq x_i$. Hence, we have

$$\begin{array}{rl} & L(P^*, f, \alpha) - L(P, f, \alpha) \\ = & w_1[\alpha(x^*) - \alpha(x_{i-1})] + w_2[\alpha(x_i) - \alpha(x^*)] - m_i[\alpha(x_i) - \alpha(x_{i-1})] \\ = & (w_1 - m_i)[\alpha(x^*) - \alpha(x_{i-1})] + (w_2 - m_i)[\alpha(x_i) - \alpha(x^*)] \\ \geq & 0 \end{array}$$

If P^* contains k points more than P, we repeat this reasoning k times, and arrive

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

which completes proof for L.

Proof for U is similar.

Q.E.D.

Theorem 6.6. *Suppose f is defined on $[a, b]$. Then we have*

$$\underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha.$$

Proof: Let P^* be the common refinement of two partitions P_1 and P_2. By the theorem directly above, we have

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

Hence,

$$L(P_1, f, \alpha) \leq U(P_2, f, \alpha).$$

If P_2 is fixed and the *sup* is taken over all P_1, then we have

$$\underline{\int} f d\alpha \leq U(P_2, f, \alpha).$$

The theorem follows by taking the inf over all P_2 in $\underline{\int} f d\alpha \leq U(P_2, f, \alpha)$, which completes the theorem.

Q.E.D.

Theorem 6.7. *Suppose we have a function f. We say $f \in \mathcal{R}(\alpha)$ on $[a, b]$ if and only if for every $\epsilon > 0$ there exists a partition P such that*

$$U(P, f, \alpha) - L(P, f, \alpha) < \epsilon.$$

Proof: ($\Leftarrow$) For every P, we have

$$L(P, f, \alpha) \leq \underline{\int} f d\alpha \leq \overline{\int} f d\alpha < U(P, f, \alpha).$$

Thus, $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$, implies that

$$0 \leq \overline{\int} f d\alpha - \underline{\int} f d\alpha < \epsilon.$$

Hence, if $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$. can be satisfied for every $\epsilon > 0$, then we have

$$\overline{\int} f d\alpha = \underline{\int} f d\alpha,$$

that is $f \in \mathcal{R}(\alpha)$.

($\Rightarrow$) Conversely, suppose $f \in \mathcal{R}(\alpha)$, and let $\epsilon > 0$ be given. Then there exist partitions P_1 and P_2 such that

$$U(P_2, f, \alpha) = \int f d\alpha < \frac{\epsilon}{2},$$

$$\int f d\alpha - L(P_1, f, \alpha) < \frac{\epsilon}{2}.$$

We choose P to be the common refinement of P_1 and P_2. Then Theorem 6.5 above (which is Theorem 6.4 in text [10]), together with two equations above, shows

$$U(P, f, \alpha) \leq U(P_2, f, \alpha) < \int f d\alpha + \frac{\epsilon}{2} < L(P_1, f, \alpha) + \epsilon \leq L(P, f, \alpha) + \epsilon,$$

so $0 \leq \overline{\int} f d(\alpha) - \underline{\int} f d\alpha < \epsilon$ holds for this partition P.

Q.E.D.

Theorem 6.8. *This theorem has three parts:*

(a) *If $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$ holds for some P and some ϵ, then $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$ holds with the same ϵ for every refinement of P.*

(b) *If $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$ holds for $P = \{x_0, ..., x_n\}$ and if s_i, t_i are arbitrary points in $[x_{i-1}, x_i]$, then*

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \triangle \alpha_i < \epsilon.$$

Proof: Theorem 6.5 above (which is Theorem 6.4 in text [10]) implies (a). Under the assumptions in (b), both $f(s_i)$ and $f(t_i)$ lie in $[m_i, M_i]$, so that $|f(s_i) - f(t_i)| \leq M_i - m_i$. Thus

$$\sum_{i=1}^{n} |f(s_i) - f(t_i)| \triangle\alpha_i \leq U(P, f, \alpha) - L(P, f, \alpha),$$

which proves (b). The obvious inequalities

$$L(P, f, \alpha) \leq \sum f(t_i) \triangle\alpha_i \leq U(P, f, \alpha)$$

and

$$L(P, f, \alpha) \leq \int f d\alpha \leq U(P, f, \alpha)$$

prove (c).

Q.E.D.

Theorem 6.9. *If f is continuous on $[a, b]$, then $f \in \mathcal{R}(\alpha)$ on $[a, b]$.*

Proof: Let $\epsilon > 0$ be given. Choose $\eta > 0$ so that $[\alpha(b) - \alpha(b)]\eta < \epsilon$. Since f is uniformly continuous on $[a, b]$ (from Theorem 4.19 of text [10]), there exists a $\delta > 0$ such that

$$|f(x) - f(t)| < \eta$$

if $x \in [a, b]$, $t \in [a, b]$, and $|x - t| < \delta$.

If P is any partition of $[a, b]$ such that $\triangle x_i < \delta$ for all i, then $|f(x) - f(t)| < \eta$ implies that

$$M_i - m_i \leq \eta, \ i - 1, ..., n$$

and therefore

$$U(P, f, \alpha) - L(P, f, \alpha) = \sum_{i=1}^{n} (M_i - m_i) \triangle\alpha_i \leq \eta \sum_{i=1}^{n} \triangle\alpha_i = \eta[\alpha(b) - \alpha(a)] < \epsilon.$$

By Theorem 6.7 (Theorem 6.6 in text [10]), we have $f \in \mathcal{R}(\alpha)$.

Q.E.D.

Theorem 6.10. *Suppose that α is monotonic. If f is monotonic on $[a, b]$, and if α is continuous on $[a, b]$, then $f \in \mathcal{R}(\alpha)$.*

Proof: Let $\epsilon > 0$ be given. For any positive integer n, choose a partition such that

$$\triangle\alpha_i = \frac{\alpha(b) - \alpha(a)}{n}, \ i = 1, ..., n.$$

This is because α is continuous (one can refer to Theorem 4.23 from text [10]).

Suppose that f is monotonically increasing and assume w.l.o.g.. Then we have $M_i = f(x_i)$, $m_i = f(x_{i-1})$, for $i = 1, ..., n$ so that

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &= \tfrac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \\ &= \tfrac{\alpha(b) - \alpha(a)}{n} [f(b) - f(a)] \\ &< \epsilon \end{aligned}$$

We can take n sufficient large, and by Theorem 6.7 (Theorem 6.6 in text [10]), $f \in \mathcal{R}(\alpha)$.

Q.E.D.

6.2 Properties of the Integral

Go back to Table of Contents. Please click TOC

Theorem 6.11. *This theorem states four properties of integral.*

(a) *If $f_1 \in \mathcal{R}(\alpha)$ and $f_2 \in (\alpha)$ on $[a, b]$, then $f_1 + f_2 \in \mathcal{R}(\alpha)$, $cf \in \mathcal{R}(\alpha)$ for every constant c, and*

$$\int -a^b(f_1 + f_2)d\alpha = \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha,$$

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

(b) *If $f_1(x) \leq f_2(x)$ on $[a, b]$, then*

$$\int_a^b f_1 d\alpha \leq \int_a^b f_2 d\alpha.$$

(c) *If $f \in \mathcal{R}(\alpha)$ on $[a, b]$ and if $a < c < b$, then $f \in \mathcal{R}(\alpha)$ on $[a, c]$ and on $[c, b]$, and*

$$\int_a^c f d\alpha + \int_c^b f d\alpha = \int_a^b f d\alpha.$$

(d) *If $f \in \mathcal{R}(\alpha)$ on $[a, b]$ and if $|f(x) \leq M$ on $[a, b]$, then*

$$\left|\int_a^b f d\alpha\right| \leq M[\alpha(b) - \alpha(a)].$$

(e) *If $f \in \mathcal{R}(\alpha_1)$ and $f \in \mathcal{R}(\alpha_2)$, then $f \in \mathcal{R}(\alpha_1 + \alpha_2)$ and*

$$\int_a^b f d(\alpha_1 + \alpha_2) = \int_a^b f d\alpha_1 + \int_a^b f d\alpha_2;$$

if $f \in \mathcal{R}(\alpha)$ and c is a positive constant, then $f \in \mathcal{R}(c\alpha)$ and

$$\int_a^b f d(c\alpha) = c\int_a^b f d\alpha.$$

Theorem 6.12. *If $f \in \mathcal{R}(\alpha)$ and $g \in \mathcal{R}(\alpha)$ on $[a, b]$, then*

(a) *$fg \in \mathcal{R}(\alpha)$;*

(b) *$|f| \in \mathcal{R}(\alpha)$ and $\left|\int_a^b f d\alpha\right| \leq \int_a^b |f| d\alpha$.*

Definition 6.13. The unit step function I is defined by

$$I(x) = \begin{cases} 0 & , \quad x \leq 0, \\ 1 & , \quad x > 0. \end{cases}$$

6.3 Fundamental Theorems of Calculus

Go back to Table of Contents. Please click TOC

Theorem 6.14. *If f in $\mathcal{R}$ on $[a, b]$ and if there is a differentiable function F on $[a, b]$ such that $f' = f$, then*

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

Example 6.15. Consider $\ln : (0, \infty) \to \mathbb{R}$. Then we have $\ln x = \int_1^x \frac{1}{t} dt$. One observation is that $\ln x > 0$ if $x > 1$ and $\ln x < 0$ if $x \in (0, 1)$. Also note that $\ln 1 = 0$ and ln is increasing.

Figure 55: Graphic view for example $y = \ln x$.

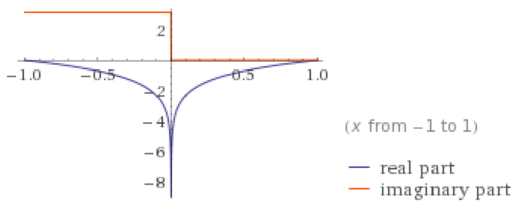
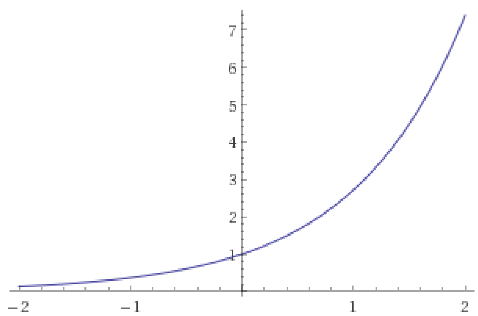


Figure 56: Graphic view for example $y = \exp(x) = e^x$.



$\square$

Definition 6.16. (Inverse). Consider function $\exp(x)$. There exists an inverse $\exp(x) = ln^{-1}(x)$. Moreover, $\exp(x)$ is an increasing function that is concave upward. Then the derivative of $\exp(x)$ is $\exp'(x) = \frac{1}{\ln'(\exp\ x)}$.

A question to ask is: is it the case that $\exp(x) \to 0$ as $x \to -\infty$? The answer is "yes" since $\exp(x)$ decreases monotonically as $x \to -\infty$. We consider the identities. Note that $\ln(xy) = \ln x + \ln y$. Note that $\exp(0) = 1$ and $\exp(x + y) = \exp(x)\exp(y)$.

Example 6.17. Consider the function $arc\tan x = \int_0^x \frac{dt}{1+t^2}$. We start with $y = \frac{1}{1+t^2}$, which is the following graph. We take integral of this function from 0 to x and we would obtain $arc\tan x$

Figure 57: Graphic view for example $y = \frac{1}{(1+t^2)}$.

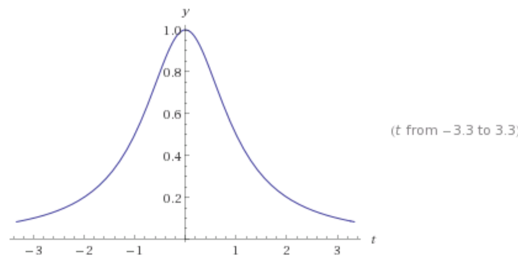
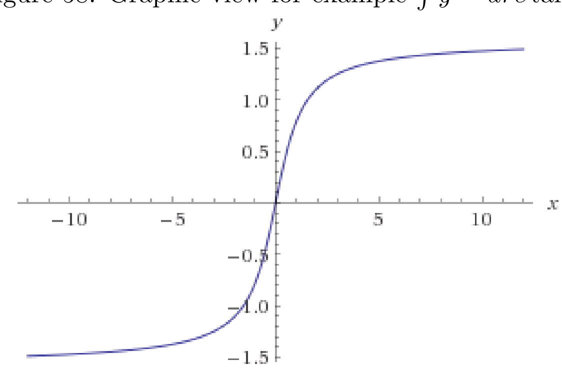


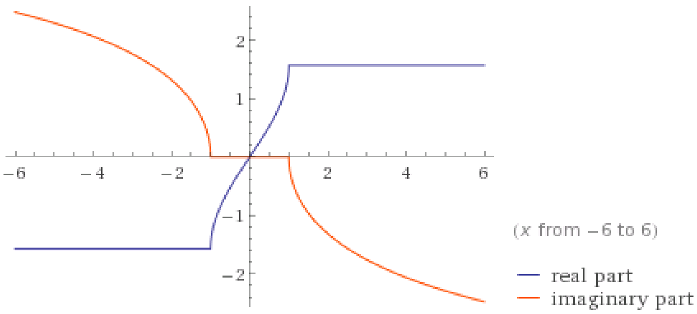
Figure 58: Graphic view for example $\int y = arc\tan x$.



$\square$

Example 6.18. For another example, consider $arc\sin x = \int_0^x \frac{dt}{\sqrt{1-t^2}}$ for $x \in (-1, 1)$.

Figure 59: Graphic view for example $\int y = arc \sin x$.



□

Consider a particular solution of

$$Y(t) = \int_{t_0}^{t} f'(s)dx.$$

we assume that we know $y = F(t)$ with $F(t_0) = 0$ is a solution of $y'(t) = f(t)$ with $y(t_0) = 0$. This gives us $f = F'$. For $f = F'$ is integrable in $[a, b]$, we prove that it implies $\int_a^b f(t)dt = F(b) - F(a)$.

Proof (Fundamental Theorem of Calculus): Given f is integrable. In other words, we have the following, $\forall \epsilon > 0$, $\exists P$ a partition such that $U(f, P) - L(f, P) < \epsilon$ for $t_0 = a \leq t \leq t_1 \leq \dots t_n \leq b\}$. The goal is to obtain

$$F(b) - F(a) = \sum_{i=1}^{n} (F(t_o) - F(_{i-1})).$$

From M.V.T., we know that $\exists t_i^* \in [t_{i-1}, t_i]$ such that $F(t_i) - F(t_{i-1}) = f(t_i^*)\triangle t_i = f(t_i^*)(t_i - t_{i-1})$. Then we have

$$\begin{cases} F(b) - F(a) & = & \sum_{i=1}^{n} f(t_i^*)\triangle t_i & \in & [L(f, P), U(f, P)] \\ & & \int_a^b f(t)dt & \in & [L(f, P), U(f, P)] \end{cases}$$

Adding them together, it gives us

$$\left| \int_a^b f(t)dt - (F(b) - F(a)) \right| < \epsilon$$

Q.E.D.

Remark 6.19. This is an application of telecopic sum, that is,

$$\sum_{i=1}^{n} \triangle a_i = a_n - a_0, \ \triangle a_i = a_i - a_{i-1}$$

□

Now we can look at the following examples.

Example 6.20. Consider $\int_0^x tdt = \frac{t^{n+1}}{n=1}\Big|_0^x = \frac{x^{n+1}}{n+1}$ for $n \in \mathbb{N}$.

□

Example 6.21. For $\alpha \in \mathbb{R} \setminus \{-1\}$, we have $\int_0^x t^\alpha dt = \frac{x^{\alpha+1}}{\alpha+1}$; and then $\int_0^x \frac{dt}{t} = \ln x$.

□

Remark 6.22. Consider x^α, consider $\alpha > 1$, $\alpha \in (0,1)$, and $\alpha < 0$.

Figure 60: Graphic view for example, considering x^α, of $\alpha > 1$.

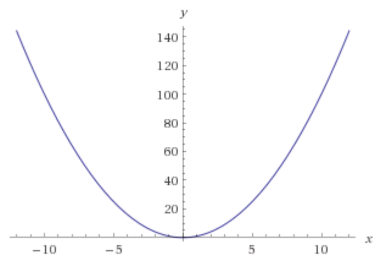


Figure 61: Graphic view for example, considering x^α, of $\alpha \in (0,1)$.

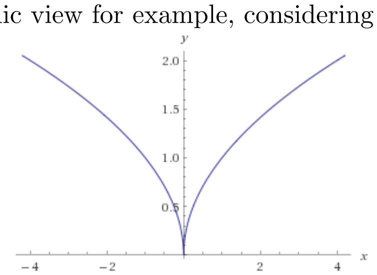
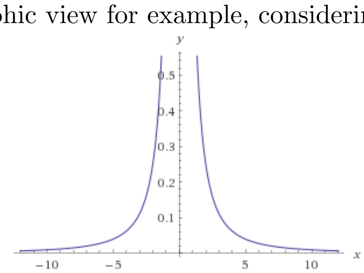


Figure 62: Graphic view for example, considering x^α, of $\alpha < 1$.



Moreover, we can consider $x|x|^{\alpha-1}$ for $\alpha > 1$, $\alpha \in (0,1)$, and $\alpha < 0$.

Figure 63: Graphic view for example, considering $x|x|^{\alpha-1}$, of $\alpha > 1$.

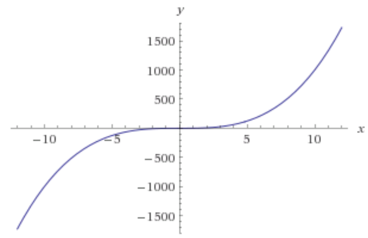


Figure 64: Graphic view for example, considering $x|x|^{\alpha-1}$, of $\alpha \in (0, 1)$.

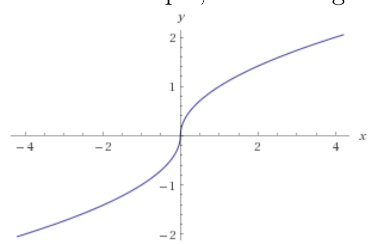
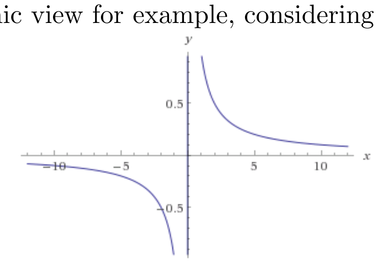


Figure 65: Graphic view for example, considering $x|x|^{\alpha-1}$, of $\alpha < 1$.



$\square$

Now we discuss integration by parts. Suppose we have f and g integrable. That is, we can assume that f' and g' are integrable. This implies that

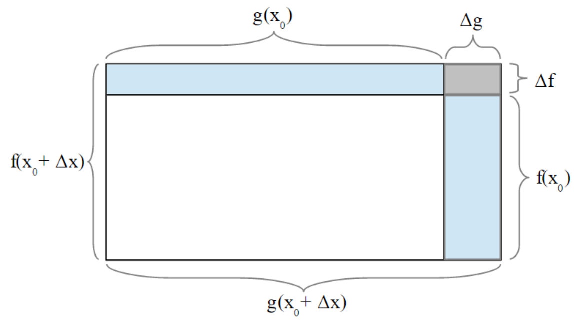
$$\int_a^b fg' = fg\Big|_a^b - \int_a^b f'g$$

which is proved as the following.

Proof: By *Product Rule*, we have $(fg)' = f'g + fg'$. Compute $\int_a^b dx$ on both sides. We obtain $fg|_a^b = \int_a^b (f'g + fg')$.

We can interpret this proof from the following geometric appearances, which is shown in the following graph.

Figure 66: Graphic view for the proof.



Note that $\triangle g \triangle f$ is $\triangle f \triangle g = f(x+\triangle x)g(x+\triangle x) - f(x)g(x)$ to be the area of the purple region, which tends to 0 as $\triangle$ partition tends to sufficiently small.

The sum of all the shaded region:

$$\begin{array}{rcl} \sum Area & = & \square_{Top\ Left\ Blue} + \square_{Bottom\ Right\ Blue} + \square_{Top\ Right\ Purple} \\ \Rightarrow \quad \triangle(fg) & = & (\triangle f)g + f(\triangle g) + \triangle f \triangle g \\ \Rightarrow \quad \frac{\triangle(fg)}{\triangle x} & = & \frac{(\triangle f)g}{\triangle x} + \frac{f(\triangle g)}{\triangle x} + \frac{\triangle f \triangle g}{\triangle x} \\ \Rightarrow \quad (fg)' & = & f'g + fg' + 0 \end{array}$$

which completes the proof.

Q.E.D.

Remark 6.23. The procedure is the same ofr integral. Consider function $y = f(x)$. We can take integral of y, that is, $\int_a^b f(x)dx$, to obtain some area under f. For α and β in the image of f, we have the inverse of y. Let it be $g = f^{-1}$. Then we can take the integral as well, that is, $\int_\alpha^\beta g(y)dy$.

In the situation that we do not know y but we want the integral of y, we can consider $g = y^{-1}$ and take the integral of g.

□

We discuss the following applications.

Example 6.24. Consider $\int_a^b p(x)\exp(x)dx$ with $p(x)$ some polynomial, *Fundamental Theorem of Calculus* can help us obtain this integral.

□

Remark 6.25. Consider the integral of $\ln(t)$ from 1 to x. We compute

$$\begin{array}{rcl} \int_1^x \ln(t)dt & = & t(\ln(t)-1) + c\Big|_1^x \\ & = & x(\ln(x)-1) - (\ln(1)-1) \\ & = & x\ln(x) \end{array}$$

□

We can also compute integral by the change of variable. Consider *Chain Rule*, by Fundamental Theorem of Calculus, we have

$$\int_a^b f(g(y))g(y) = \int_{g(a)}^{g(b)} f(y)dy$$

which requires us the following aassumptions. We assume that g' is continuous, which is equivalent as g is continuous. We also assume f is continuous. We further aassume that $g([a,b]) \subseteq dom\ f$. First, consider the function $f(g(y))g'(y)$ to be continuous, which imply that it is integrable. Next, f is the derivative of $F(x) = \int_a^x f(t)dt$. This gives us

$$\begin{array}{rrcl} \Rightarrow & (f \circ g)' & = & (f \circ g)g' \\ \Rightarrow & \int_a^b (f \circ g)g' & = & F \circ g \Big|_a^b \\ & & = & F(g(b)) - F(g(a)) \\ & & = & \int_{g(a)}^{g(b)} \end{array}$$

Finally, as the end of this section, we summarize the following. The properties of derivations are interchangeable with properties of integrals through the work of *Fundamental Theorem of Calculus*. We have the following table for simple memory:

Property of Derivation	$\overset{F.T.C.}{\Leftrightarrow}$	**Properties of Integration**
Product Rule		Integration by Parts
Chain Rule		Substitution Rule or Change of Variables

7 §Sequences and Series of Functions§

Go back to Table of Contents. Please click TOC

Definition 7.1. Suppose $\{f_n\}$, $n = 1, 2, 3, ...$, is a sequence of functions defined on a set E, and suppose that the sequence of numbers $\{f_n(x)\}$ converges for every $x \in E$. We can define a function f by

$$f(x) = \lim_{n\to\infty} f_n(x),\ for\ x \in E.$$

In this case, we say that $\{f_n\}$ converges on E and that f is the limit, or the limit function, of $\{f_n\}$. Sometimes we shall use a more descriptive terminology and shall say that "$\{f_n\}$ converges to f pointwise on E" if the above euqation holds. Similarly, if $\sum f_n(x)$ converges for every $x \in E$, and if we define

$$f(x) = \sum_{n=1}^{\infty} f_n(x),\ x \in E,$$

the function f is called the sum of the series $\sum f_n$.

A natural question is to ask whether eseential properties of cuntions are preserved under the limit operations in the above equations.

Example 7.2. For $m = 1, 2, 3, ..., n = 1, 2, 3, ...,$ let

$$s_{m,n} = \frac{m}{m+n}.$$

Then for every fixed n, we have

$$\lim_{m\to\infty} s_{m,n} = 1,$$

so that

$$\lim_{n\to\infty}\lim_{m\to\infty} s_{m,n} = 1.$$

On the other hand, for every fixed m,

$$\lim_{n\to\infty} s_{m,n} = 0,$$

so that

$$\lim_{m\to\infty}\lim_{n\to\infty} s_{m,n} = 0.$$

$\square$

Example 7.3. Let

$$f_n(x) = \frac{x^2}{(1+x^2)^n}$$

for $x \in \mathbb{R}$ and $n = 0, 1, 2, ...,$ and consider

$$f_(x) = \sum_{n=0}^{\infty} f_n(x) = \sum_{n=0}^{\infty} \frac{x^2}{(1+x^2)^n}.$$

Since $f_n(0) = 0$, we have $f(0) = 0$. For $x \neq 0$, the last series in $f(x)$ is a convergent geometric series within sum $1 + x^2$ (see remark). Hence,

$$f(x) = \begin{cases} 0 & , \quad x = 0 \\ 1 + x^2 & , \quad x \neq 0 \end{cases}$$

so that a convergent series of continuous functions may have a discontinuous sum.

$\square$

Remark 7.4. Recall Theorem 3.26 from text [10]. If $0 \leq x < 1$, then $\sum\limits_{n=0}^{\infty} x^n = \frac{1}{1-x}$. if $x \geq 1$, the series diverges.

$\square$

Example 7.5. Let

$$f_n(x) = n^2 x (1 - x^2)^n$$

for $0 \leq x \leq 1$, $n = 1, 2, 3, ...$ Then, for $0 < x \leq 1$, we have

$$\lim_{n\to\infty} f_n(x) = 0,$$

by Theorem 3.20(d). Since $f_n(x) = 0$, we see that

$$\lim_{n\to\infty} f_n(x) = 0$$

for $0 \leq x \leq 1$.

Thus, we have

$$\int_0^1 x(1-x^2)^n dx = \frac{1}{2n+1}.$$

Thus,

$$\int_0^1 f_n(x)dx = \frac{n^2}{2n+2} \to \infty$$

as $n \to \infty$.

$\square$

7.1 Uniform Convergence

Go back to Table of Contents. Please click TOC

Definition 7.6. We say that a sequence of functions $\{f_n\}$, $n = 1, 2, 3, ...,$ converges uniformly on E to a function f if for every $\epsilon > 0$ there is an integer N such that $n \geq N$ implies

$$|f_n(x) - f(x)| \leq \epsilon$$

for all $x \in E$.

Definition 7.7. Let D be a subset of $\mathbb{R}$ and let $\{f_n\}$ be a sequence of functions defined on D. We say that $\{f_n\}$ converges pointwise on D if $\lim_{n\to\infty} f_n(x)$ exists for each point x in D. This means that $\lim_{n\to\infty} f_n(x)$ is a real number that depends only on x.

Example 7.8. Consider the sequence $\{f_n\}$ of functions defined by

$$f_n(x) = \frac{nx + x^2}{n^2}, \ \forall x \in \mathbb{R}.$$

Show that $\{f_n\}$ converges pointwise.

$\square$

Proof: For all $x \in \mathbb{R}$, we have

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{x}{n} + \frac{x^2}{n^2} = x(\lim_{n\to\infty} \frac{1}{n}) + x^2(\lim_{n\to\infty} \frac{1}{n^2}) = 0 + 0 = 0$$

Q.E.D.

Example 7.9. Let $\{f_n\}$ be the sequence of functions on $(0, \infty)$ defined by

$$f_n(x) = \frac{nx}{1 + n^2x^2}.$$

This function converges pointwise to zero. Indeed, $1 + n^2x^2) \sim n^2x^2$ as n gets larger and larger. Then

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{nx}{n^2x^2} = \frac{1}{x} \lim_{n\to\infty} \frac{1}{n} = 0.$$

For any $\epsilon < 1/2$, we have

$$\left| f_n(\frac{1}{n}) - f(\frac{1}{n}) \right| = \frac{1}{2} - 0 > \epsilon.$$

Hence, $\{f_n\}$ is not uniformly convergent.

Q.E.D.

$\square$

Thus, $\{f_n\}$ converges pointwise to the zero function on $\mathbb{R}$.

Every uniformly convergent sequence is pointwise convergent. Explicitly, the difference between the two concepts is this: if $\{f_n\}$ converges pointwise on E, then there exists a function f such that, for every $\epsilon > 0$, and for every $x \in E$, there is an integer N, depending on ϵ and on x such that the above equation holds for $n \geq N$. If $\{f_n\}$ converges uniformly on E, it is possible for $\epsilon > 0$, to find one integer N which the argument holds for all $x \in E$.

In other words, we say the series $\sum f_n(x)$ converges uniformly on E if the sequence $\{s_n\}$ of partial sums defined by

$$\sum_{i=1}^{n} f_i(x) = s_n(x)$$

converges uniformly on E, which leads to the following theorem.

Theorem 7.10. *The sequence of functions $\{f_n\}$, defined on E, converges uniformly on E if and only if for every $\epsilon > 0$ there exists an integer N such that $m \geq N$, $n \geq N$, $x \in E$ implies that*

$$|f_n(x) - f_m(x)| \leq \epsilon.$$

Proof: Suppose $\{f_n\}$ converges uniformly on E, and let f be the limit function. Then there is an itneger N such that $n \geq N$, $x \in E$ implies that

$$|f_n(x) - f(x)| \leq \frac{\epsilon}{2},$$

so that

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| \leq \epsilon$$

if $n \geq N$, $m \geq N$, $x \in E$.

For the opposite direction, suppose Cauchy holds. Then the sequence $\{f_n(x)\}$ converges, for every x, to a limit which we may call $f(x)$. Thus the sequence $\{f_n\}$ converges on E, to f. We have to prove that the convergence is uniform. Given $\epsilon > 0$, choose N such that $|f_n(x) - f_m(x)| \leq \epsilon$ holds. Since $f_m(x) \to f(x)$ as $m \to \infty$, this gives

$$|f_n(x) - f(x)| \leq \epsilon$$

for every $N \geq N$ and every $x \in E$, which finishes the proof.

Q.E.D.

Theorem 7.11. *Suppose*

$$\lim_{n \to \infty} f_n(x) = f(x), \; x \in E.$$

Let

$$M_n = \sup_{n \in E} |f_n(x) - f(x)|.$$

Then we have $f_n \to f$ uniformly on E if and only if $M_n \to 0$ as $n \to \infty$.

Theorem 7.12. *Suppose $\{f_n\}$ is a sequence of functions defined on E, and suppose that*

$$|f_n(x)| \le M_n,\ x \in E,\ n = 1, 2, 3, ...$$

Then $\sum f_n$ converges uniformly on E if $\sum M_n$ converges.

Proof: Suppose $\sum M_n$ converges as stated in the condition. For arbitrary $\epsilon > 0$, provided that m and n are sufficiently large, we have

$$\left|\sum_{i=n}^{m} f_i(x)\right| \le \sum_{i=n}^{m} M_i \le \epsilon,\ x \in E,$$

which shows unfirom convergence.

Q.E.D.

Example 7.13. Let $\{f_n\}$ be the sequence of functions defined by $f_n(x) = cos^n(x)$f for $-\pi \le x \le \pi/2$. Discuss the pointwise convergence of the sequence.

Solution: For $-\pi/2 \le x < 0$ and for $0 < x \le \pi/2$, we have

$$0 \le cos(x) < 1.$$

It follows that

$$\lim_{n\to\infty} (cos(x))^2 = 0,\ for\ x \ne 0.$$

Moreover, since $f_n(0) = 1$ for all $n \in \mathbb{N}$, one gets $\lim_{n\to\infty} f_n(0) = 1$. Therefore, $\{f_n\}$ converges pointwise to the function f defined by

$$f(x) = \begin{cases} 0 & , \quad if\ -\frac{\pi}{2} \le x < 0,\ or\ 0 < x \le \frac{\pi}{2} \\ 1 & , \quad if\ x = 0 \end{cases}$$

□

Example 7.14. Consider the sequence $\{f_n\}$ of functions defined by

$$f_n(x) = \frac{x}{3 + nx^2},\ \forall x \in \mathbb{R}.$$

Show that $\{f_n\}$ converges pointwise.

□

Proof: For every real number x, we have

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{x}{nx^2} = 0.$$

Hence, $f_n(x)$ converges pointwise to the zero function.

Q.E.D.

Example 7.15. Consider the sequence of functions defined by

$$f_n(x) = nx(1 - x)^n,\ on\ [0, 1]$$

Show that $\{f_n\}$ converges pointwise to the zero function.

□

Solution: Note that $f_n(0) = f_n(1) = 0$, for all $n \in \mathbb{N}$. Suppose $0 < x < 1$, we have

$$\lim_{n\to\infty} f_n(x) = 0$$

Therefore, the given sequence converges pontwise to zero.

Example 7.16. Let $\{f_n\}$ be the sequence of functions on $\mathbb{R}$ defined by

$$f_n(x) = \begin{cases} n^3 & , \quad if\ 0 < x \le \frac{1}{n} \\ 1 & , \quad otherwise \end{cases}$$

Show that $\{f_n\}$ converges pointwise to the constant function $f = 1$ on $\mathbb{R}$.

□

Solution: For any $x \in \mathbb{R}$, there is a natural number N such that x does not belong to the interval $(0, 1/N)$. The intervals $(0, 1/n)$ get smaller as $n \to \infty$. We see that $f_n(x) = 1$ for all $n > N$. Hence,

$$\lim_{n\to\infty} f_n(x) = 1,\ \forall x.$$

which leads to conclusion.

7.2 Uniform Convergence and Continuity

Go back to Table of Contents. Please click TOC

Theorem 7.17. *Suppose $f_n \to f$ uniformly on a set E in a metric space. Let x be a limit point of E, and suppose that*

$$\lim_{t\to x} f_n(t) = A_n,\ n = 1, 2, 3, ...$$

Then $\{A_n\}$ converges, and

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

That is, the conclusion is

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

Proof: Let $\epsilon > 0$ be given. By the uniform convergence of $\{f_n\}$, there exists N such that $n \ge N$, $m \ge N$, $t \in E$ imply

$$|f_n(t) - f_m(t)| \le \epsilon$$

Letting $g \to x$, we obtain

$$|A_n - A_m| \le \epsilon$$

for $n \ge N$, $m \ge N$, so that $\{A_n\}$ is a Cauchy sequence and therefore converges, say to A.

Next, we have

$$|f(t) - A| \le |f(t) - f_n(t)| + |f_n(t) - A_n| + |A_n - A|$$

Choose n such that $|f(t) - f_n(t)| \le \epsilon/3$ for all $t \in E$, and such that $|A_n - A| \le \epsilon/3$; then for this n, we choose a neighborhood V of x such that $|f_n(t) - A_n| \le \epsilon/3$ if $t \in V \cup E$, $t \ne x$. Hence, we have $|f(t) - A| \le \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon$, which completes the proof.

Q.E.D.

Theorem 7.18. *If $\{f_n\}$ is a sequence of continuous functions on E, and if $f_n \to f$ uniformly on E, then f is continuous on E.*

Theorem 7.19. *Suppose K is compact, and*

(a) $\{f_n\}$ is a sequence of continuous functions on K,

(b) $\{f_n\}$ converges pointwise to a continuous function f on K,

(c) $\{f_n(x) \geq f_{n+1}(x)$ for all $x \in K$, $n = 1, 2, 3, ...$

Then $f_n \to f$ uniformly on K.

7.3 Uniform Convergence and Integration

Go back to Table of Contents. Please click TOC

Recall that we have discussed pointwise convergent versus uniform convergence. For *pointwise* convergence, $\forall x \in X$, there is $f_n(x) \to f(x)$ as $N \to \infty$, that is, we have $\forall x \in X$, for $\epsilon > 0$, there $\exists N = N(x, \epsilon) \in N$ sufficiently large such that $f_n(x) \to f(x)$. For *uniform* convergence, we have $f_n \to f$ as $n \to \infty$ uniformly convergent if $\forall \epsilon > 0$, there $\exists N = N(\epsilon) \in N$ sufficiently large such that $|f_n(x) - f(x)| < \epsilon$ for $n \geq N$ for all $x \in X$.

Figure 67: Graphic view for the pointwise convergence. For pointwise convergence, we first fix a value x_0. Then we choose an arbitrary neighborhood around $f(x_0)$, which corresponds to a vertical interval centered at $f(x_0)$. Finally, we pick N so that $f_n(x_0)$ intersects the vertical line $x = x_0$ inside the interval $(f(x_0) - \epsilon, f(x_0) + \epsilon)$.

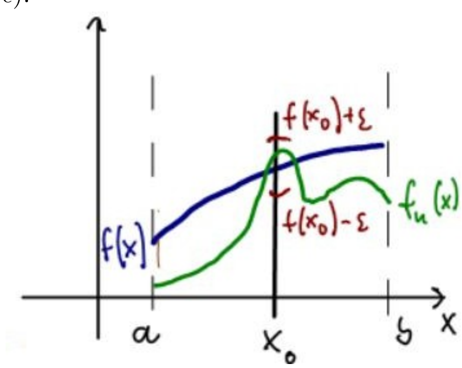
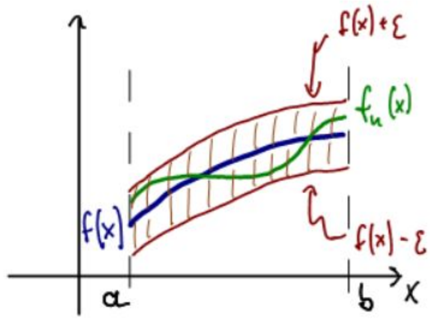


Figure 68: Graphic view for the uniform convergence. For uniform convergence, we draw an ϵ-neighborhood around the entire limit function f, which results in an "ϵ-stip" with $f(x)$ inthe middle (anywhere in the middle). Now we pick N so that $f_n(x)$ is completely inside that strip for all x in the domain.



Next, let us recall the definition of *Cauchy*. Let (a_n) be a sequence in $\mathbb{R}$. We say that (a_n) is a *Cauchy sequence* if, for all $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$m, n \geq N \Rightarrow |a_m - a_n| < \epsilon.$$

We already showed in previous sections that a *Cauchy* sequence is bounded. The proof is simple. Take $1 > 0$ so there exists N such that $m, n \geq N \Rightarrow |a_m - a_n| < 1$. For $m \geq N$, $|a_m \leq 1 + |a_N|$ be the triangle inequality. Thus, for all m, there is $|a_m \leq 1 + |a_n| + |a_2| + \cdots + |a_N|$.

We shall prove the following theorems.

Theorem 7.20. *Every convergent sequence is Cauchy. Moreover, every real Cauchy sequence is convergent.*

Proof: This proof has two statements and we prove them accordingly.

(*Convergence* $\Rightarrow$ *Cauchy*) Let $a_n \to l$ and let $\epsilon > 0$. Then there exists N such that $k \geq N \Rightarrow |a_k - l| < \epsilon/2$. For $m, n \geq N$, we have

$$|a_m - l| < \epsilon/2$$

$$|a_n - l| < \epsilon/2,$$

so

$$\begin{array}{rcl} |a_m - a_n| & \leq & |a_m - l| + |a_n - l|, \text{ by Triangle Inequality} \\ & < & \epsilon/2 + \epsilon/2 = \epsilon \end{array}$$

Now we prove the second part.

(*Cauchy* $\Rightarrow$ *Convergent*($\mathbb{R}$)) By the above, (a_n) is bounded. By *Bolzano-Weierstrass*, (a_n) has a convergent subsequence $(a_{n_k}) \to l$; so let $\epsilon > 0$. Then we have

$$\exists N_1 : r \geq N_1 \Rightarrow |a_{n_r} - l| < \epsilon/2$$

and

$$\exists N_2 : m, n \geq N_2 \Rightarrow |a_m - a_n| < \epsilon/2$$

Let $s = \min\{r|n_r \geq N_2\}$ and write $N = n_s$ (simple rewrite it). Then we have

$$\begin{array}{rcl} m, n \geq N & \Rightarrow & |a_m - a_n| \\ & \geq & |a_m - a_{n_s}| + |a_{n_s} - l| \\ & < & \epsilon/2 + \epsilon/2 = \epsilon \end{array}$$

Q.E.D.

Remark 7.21. The theorem is false if we take *uniform convergence* out of the condition.

□

Theorem 7.22. *If a sequence of functions $f_n(x)$ defined on D converges uniformly to a function $f(x)$, and if each $f_n(x)$ is continuous on D, then limit function $f(x)$ is also continuous on D.*

Proof: The idea is the following. Given f_n converges uniformly and that each f_n is continous. We want to prove that f is continuous on D. Thus, we need to pick an x_0 and show that

$$|f(x_0) - f(x)| < \epsilon \ if \ |x_0 - x| < \delta$$

as if it is stated in the definition.

Let us start by choosing an arbitrary $\epsilon > 0$. Because of uniform convergence, we can find an N such that

$$|f_n(x) - f(x)| < \epsilon/3 \ if \ n \geq N$$

for some N sufficiently large and for all $x \in D$. Because all f_n are continuous, we can find a particular $\delta > 0$ such that

$$|f_N(x_0) - f_N(x)| < \epsilon/3 \ if \ |x_0 - x| < \delta$$

Then, we sum up and have

$$\begin{array}{rcl} |f(x_0) - f(x)| & \leq & |f(x_0) - f_N(x_0)| + |f_N(x_0) - f_N(x)| + |f_N(x) - f(x)| \\ & \leq & \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon \end{array}$$

as long as $|x_0 - x| < \delta$; and that means that f is continuous at x_0 by definition, which completes the proof.

Q.E.D.

Recall that we have discussed metric spaces. A *metric* on X is a function $d : X \times X \to \mathbb{R}$, with the following properties: (a) $d(x, y) \geq 0$ for all $x, y \in X$, and $d(x, y) = 0$ if and only if $x = y$; (b) $d(x, y) = d(y, x)$, for all $x, y \in X$; and (c) $d(x, y) \leq d(x, z) + d(z, y)$, for all $x, y, z \in X$. Based on this understanding, we can move further and define *norm* as well as *formed spaces.*

Definition 7.23. A norm on a linear space X is a function $||\cdot|| : X \to \mathbb{R}$ with the following properties:

(a) $||x|| \geq 0$, for all $x \in X$ (nonnegative);

(b) $||\lambda x|| = |\lambda| ||x||$, for all $x \in X$ and $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) (homogenous);

(c) $||x + y|| \leq ||x|| + ||y||$, for all $x, y \in X$ (triangle inequality);

(d) $||x = 0||$ implies that $x = 0$ (strictly positive).

Definition 7.24. A normed linear space $(X, ||\cdot||)$ is a linear space X equipped with a nrom $||\cdot||$.

Definition 7.25. In the understanding of this analysis, a metric space M is called *complete* or *Cauchy space* if every Cauchy sequence of points in M has a limit hthat is also in M. Alternatively, we say a metric space M is complete if every *Cauchy* sequence in M converges.

Example 7.26. Consider $\{f_n\} \subseteq C(x)$ and $f_n \to f$ uniformly continuous in $C(x)$. That is, there is $\{x_m\} \subseteq X$ such that $x_m \to x_0$ as $m \to \infty$.

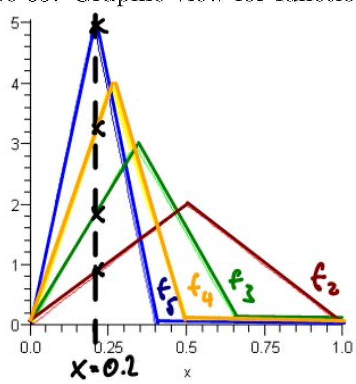
□

Example 7.27. We say that $\lim_{m,n\to\infty} c_{n,m} = L$ if $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$, we have $|c_{n,m} - L| < \epsilon$ if $n, m \in \mathbb{N}$.

□

Example 7.28. Consdier $f_n \to 0$ (pointwise) such that f_n takes the following form

Figure 69: Graphic view for function f_n.



Notice that $x_m = \frac{1}{m} \to 0$ given, that is, the area under the function is always 1. For example, take $x = 0.2$ and we would have $f_n(0.2) = 5$.

We discuss two aspects. First, $f_n(x_m) \to f_n(0) = 0$ as $m \to \infty$. Second, we have $_n(x_m) \to 0$ as $n \to \infty$. However, we have $\lim_{n,m\to\infty} f_n(x_m)$ would not work. Because if take $m = n$, we would have $f_n(x_m) = 1$. And also because if take $m = 2n$, we would have $f_n(x_m) = 0$. Thus, the limit does not exist.

□

Now we can discuss integration.

Theorem 7.29. *Let α be monotonically increasing on $[a, b]$. Suppose $f_n \in \mathcal{R}(\alpha)$ on $[a, b]$, for $n = 1, 2, 3, ...,$ and suppose $f_n \to f$ uniformly on $[a, b]$. Then $f \in \mathcal{R}(\alpha)$ on $[a, b]$, and*

$$\int_a^b f d\alpha = \lim_{n\to\infty} f \int_a^b f_n d\alpha.$$

Proof: It suffices to prove this for real f_N. Put $\epsilon_n = \sup |f_n(x) - f(x)|$, the supremum being taken over $a \leq x \leq b$. Then $f_n - \epsilon_n \leq f \leq f_n + \epsilon_n$, so that the upper and lower integrals of f satisfy

$$\int_a^b (f_n - \epsilon_n) d\alpha \leq \underline{\int} f d\alpha \leq \overline{\int} f d\alpha \leq \int_a^b (f_n + \epsilon_n) d\alpha.$$

Hence,

$$0 \leq \underline{\int} f d\alpha - \underline{\int} f d\alpha \leq 2\epsilon_n [\alpha(b) - \alpha(a)].$$

Since $\epsilon_n \to 0$ as $n \to \infty$, the upper and lower integrals of f are equal. Thus, $f \in \mathcal{R}(\alpha)$.

Q.E.D.

The rest of the section is for education purpose only.

Definition 7.30. A family $\mathcal{F}$ of complex functions f defined on a set E in a metric space X is said to be equicontinuous on E if for every $\epsilon > 0$ there exists a $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ whenever $d(x, y) < \delta$, $x \in E$, $y \in E$, and $f \in \mathcal{F}$. Here d denotes the metric of X.

It is clear that every member of an equicontinuous family is uniformly continuous.

Theorem 7.31. *If K is a compact metric space, if $f_n \in \mathcal{G}(K)$ for $n = 1, 2, 3, ...$, and if $\{f_n\}$ converges uniformly on K, then $\{f_n\}$ is equicontinuous on K.*

Proof: Let $\epsilon > 0$ be given. Since $\{f_n\}$ converges uniformly, there is an integer N such that $||f_n - f_N|| < \epsilon$ for $n > N$. Since continuous functions are uniformly continuous on compact sets, there is a $\delta > 0$ such that $|f_i(x) - f_i(y)| < \epsilon$ if $1 \leq i \leq N$ and $d(x, y) < \delta$. If $n > N$ and $d(x, y) < \delta$, it follows that $|f_n(x) - f_n(y)| \leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < 3\epsilon$, which completes the theorem.

Q.E.D.

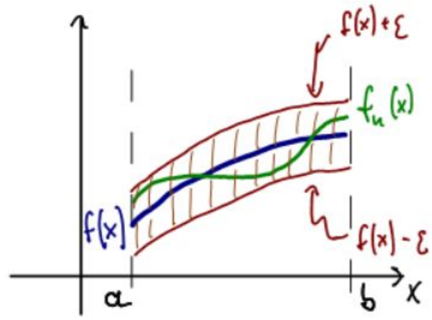
Thus, we have the following general understanding, which is an alternative approach to Theorem 7.29 above.

Consider $f_n \to f$ uniformly. We have $\{f_n\}$ and f are integrable in $[a, b]$, which implies that

$$\lim_{n \to \infty} \int_a^b f_n = \int_a^b f$$

The idea is the following. Referring to the figure of f_n that is uniformly continuous, we are essentially looking at the difference between f_n and f, which is bounded (always inside) by the dotted line.

Figure 70: Graphic view for the uniform convergence. For uniform convergence, we draw an ϵ-neighborhood around the entire limit function f, which results in an "ϵ-stip" with $f(x)$ inthe middle (anywhere in the middle). Now we pick N so that $f_n(x)$ is completely inside that strip for all x in the domain.



We can simply let $\epsilon > 0$ and $N \in \mathbb{N}$: $|f(x) - f_n(x)| < \epsilon$ if $n > N$ and $\forall x \in [a, b]$. This is equivalent as stating $f(x) - \epsilon < f_n(x) < f(x) + \epsilon$. By monotonicity, we have

$$\int_a^b f - \epsilon(b-a) < \int_a^b f_n < \int_a^b f + \epsilon(b-a)$$

which is equivalent as

$$|\int_a^b f - \int_a^b f_n| < \epsilon(b-a)$$

which completes the proof.

7.4 Uniform Convergence and Differentiation

Go back to Table of Contents. Please click TOC

Consider function $f_n(x) = \frac{1}{n}\sin(nx)$. We know that $f_n \to 0$ uniformly. However, f_n' does not converge. To understand this situation, we introduce the following theorem.

Theorem 7.32. *Suppose $\{f_n\}$ is a sequence of functions, differentiable on $[a, b]$ and such that $\{f_n(x_0)\}$ converges for some point x_0 on $[a, b]$. If $\{f_n'\}$ converges uniformly on $[a, b]$, then $\{f_n\}$ converges uniformly on $[a, b]$, to a function f, and*

$$f'(x) = \lim_{n\to\infty} f_n'(x), \ a \leq x \leq b.$$

Proof: Let $\epsilon > 0$ be given. Choose N such that $n \geq N$, $m \geq N$, implies that

$$f_n(x_0) - f_m(x_0)| < \frac{\epsilon}{2}$$

and

$$|f_n'(t) - f_m'(t)| < \frac{\epsilon}{2(b-a)}, \ a \leq t \leq b.$$

If we apply the M.V.T., to the function $f_n - f_m$, we have

$$|f_n(x) - f_m(x) - f_n(t) + f_m(t)| \leq \frac{|x-t|\epsilon}{2(b-a)} \leq \frac{\epsilon}{2}$$

for any x and t on $[a, b]$, if $n \geq N$, $m \geq N$. The inequality

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)|$$

would imply that

$$|f_n(x) - f_m(x)| < \epsilon, \; a \leq x \leq b, \; n \geq N, \; m \geq N,$$

so that $\{f_n\}$ converges uniformly on $[a, b]$. Let $f(x) = \lim_{n\to\infty} f_n(x)$, for $a \leq x \leq b$. Let us now fix a point x on $[a, b]$ and define

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t-x}, \; \phi(t) = \frac{f(t) - f(x)}{t-x}$$

for $a \leq t \leq b$, and $t \neq x$. Then we have

$$\lim_{t\to x} \phi_n(t) = f'_n(x), \; n = 1, 2, 3, ...$$

From above, we have

$$|\phi_n(t) - \phi_m(t)| \leq \frac{\epsilon}{2(b-a)}, \; n \geq N, \; m \geq N,$$

so that $\{\phi_n\}$ converges uniformly, for $g \neq x$. Since $\{f_n\}$ converges to f, we conclude from above that

$$\lim_{n\to\infty} \phi_n(t) = \phi(t)$$

uniformly for $a \leq t \leq b$, and for $t \neq x$. Thus, we have

$$\lim_{t\to x} \phi(t) = \lim_{n\to\infty} f'_n(x),$$

and this gives us $f'(x) = \lim_{n\to\infty} f'_n(x)$, which completes the proof.

Q.E.D.

7.5 Stoke's Theorem

Go back to Table of Contents. Please click TOC

In vector calculus and more general cases such as differential geometry, Stokes' theorem is a statement about the integration of differential forms on manifolds, which both simplifies and generalizes several theorems and vector calculus. The theorem provides understanding of the integral of a differential form ω over the boundary of some orientable manifold Ω is equal to the integeral of its exterior derivative $d\omega$ over the whole of Ω, i.e.,

$$\int_{\partial\Omega} \omega = \int_{\Omega} d\omega.$$

Theorem 7.33. *(Stokes' Theorem). Let S be an oriented smooth surface that is bounded by a smple, closed, smooth boundary curve C with positive orientation. Also let*

$$\int_C \vec{F} \cdot d\vec{r} = \iint_S curl\ \vec{F} \cdot d\vec{s}$$

Theorem 7.34. *(Kelvin-Stokes Theorem). Let $\gamma : [a, b] \to \mathbb{R}^2$ be a piecewise smooth Jordan plane curve. The Jordan curve theorem implies that γ divides $\mathbb{R}^2$ into two components, a compact one and another that is non-compact. Let D denote the compact part that is bounded by γ and suppose $\psi : D \to \mathbb{R}^3$ is smooth, with $S := \psi(D)$. If Γ is the space curve defined by $\Gamma(t) = \psi(\gamma(t))$ and $\mathbb{F}$ field on $\mathbb{R}^3$. then*

$$\oint_\Gamma \boldsymbol{F} \cdot d\Gamma = \iint_S \bigtriangledown \times \mathbb{F} \cdot d\boldsymbol{S}$$

One consequence of the Kelvin-Stokes theorem is that the field lines of a vector field with zero curl cannot be closed contours. The formula can be rewritten as:

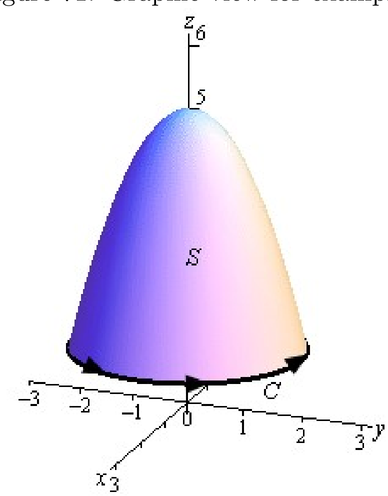
$$\iint_{\sum} \Big((\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z})dydz + (\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x})dzdx + (\frac{\partial Q}{\partial x}\frac{\partial P}{\partial y})dxdy \Big)$$

while $\sum$ represents a surface that the curl of a vector field is in.

Example 7.35. Use Stokes' Theorem to evaluate $\iint_S curl\vec{F} \cdot d\vec{S}$ where $\vec{F} = z^2\vec{i} - 3xy\vec{j} + x^3y^3\vec{k}$ and S is the part of $z = 50x^2 - y^2$ above the plane $z = 1$. Assume that S is oriented upwards.

We can start by the following graph.

Figure 71: Graphic view for example.



In this case the boundary curve C will be where the surface itnersets the plance $z = 1$ and so will be the curve $1 = 6 - x^2 - y^2$ and $x^2 + y^2 = 4$

at $z = 1$. So, the boundary curve will be the circle of radius 2 that is in the plance $z = 1$. The parameterization of this curve is,

$$\vec{r}(t) = 2\cos t\,\vec{i} + 2\sin t\,\vec{j} + \vec{k},\ -0 \leq t \leq 2\pi$$

The first two components give the circle and the third component makes sure that it is in the plane $z = 1$.

Using Stokes' Theorem we can write the surface integral as the following line integral.

$$\iint_S curl\vec{F} \cdot d\vec{S} = \int_C \vec{F} \cdot d\vec{r} = \int_0^{2s} \vec{F}\big(\vec{r}(t)\big) \cdot \vec{r}\,'(t)dt$$

Hence, it looks like we need a couple of quantities before we do this integral. Let us get the vector field evaluated on the curve. Remember

$$\begin{aligned} \vec{F}\big(\vec{r}(t)\big) &= (1)^2\,\vec{i} - 2(2\cos t)(2\sin t)\,\vec{j} + (2\cos t)^3(2\sin t)^3\,\vec{k} \\ &= \vec{i} - 12\cos t \sin t\,\vec{j} + 64\cos^3 t \sin^3 t\,\vec{k} \end{aligned}$$

Next, we need the derivative of the parameterization and the dot product of this and the vector field.

$$\vec{r}\,'(t) = -2\sin t\,\vec{i} + 2\cos t\,\vec{j}$$

$$\vec{F}\big(\vec{r}(t)\big) = -2\sin t - 24\sin t\cos^2 t$$

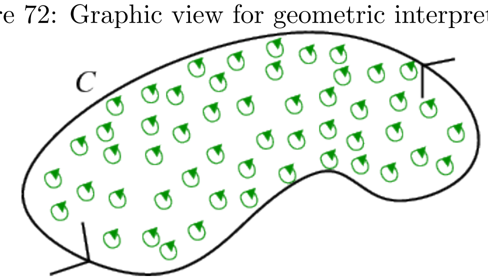
We can not do the integral

$$\begin{aligned} \textstyle\iint_S curl\vec{F} \cdot d\vec{S} &= \textstyle\int_0^{2\pi} -2\sin t - 24\sin t\cos^2 t dt \\ &= (2\cos t + 8\cos^3 t)\big|_0^{2s} \\ &= 0 \end{aligned}$$

□

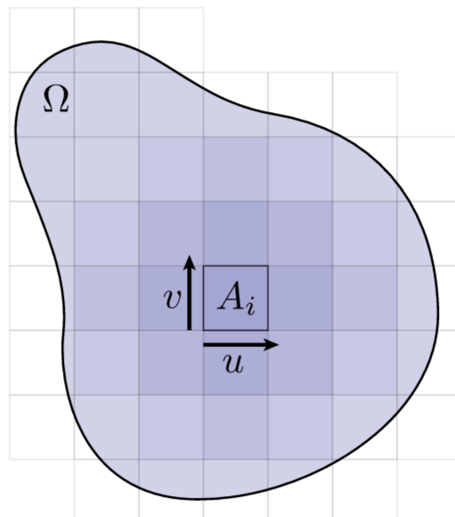
Stokes' Theorem is a generalization of Green's Theorem from circulation in a planar region to circulation along a surface. Given a continuously differentiable two-dimensional vector field $\mathbb{F}$, the integral of the "microscopic circulation" of $\mathbb{F}$ over the region D inside a simple closed curve C is equal to the total circulation of $\mathbb{F}$ around C, as suggested by the equation, that is, $\int_C \mathbb{F}ds$ is equivalent as $\iint_D$ "microscopic circulation of $\mathbb{F}$" dA. We often write $C = \partial D$ as notation meaning simply that C is the boundary of D. The "microscopic circulation" in Green's theorem is captured by the curl of the vector field and is illustrated by the green circles in the figure below.

Figure 72: Graphic view for geometric interpretation.



In other words, we can also consider the following cut. Suppose we want to compute the total area A_Ω of a region Ω in the following figure:

Figure 73: Graphic view for geometric interpretation.



The idea is to break up the domain in to a bunch of little pieces that are easy to measure (like squares) and add up their areas:

$$A_\Omega \approx \sum_i A_i.$$

As these squares get smaller and smaller, we get a better and better approximation, ultimately achieving the true area

$$A_\Omega = \int_\Omega dA.$$

7.6 The Arzela-Ascoli Theorem

Go back to Table of Contents. Please click TOC

Theorem 7.36. *If a sequence $\{f_n\}_1^\infty$ in $C(X)$ is bounded and equicontinuous then it has a uniformly convergent subsequence.*

IN this statement,

(a) *"$\mathcal{F} \subset C(X)$ is bounded" means that there exists a positive constant $M < \infty$ such that $|f(x)| \leq M$ for each $x \in X$ and each $f \in \mathcal{F}$, and*

(b) *"$\mathcal{F} \subset C(X)$ is equicontinuous" means that: for every $\epsilon > 0$ there exists $\delta > 0$ (which depends only on ϵ) such that for $x, y \in X$:*

$$d(x, y) < \delta \Rightarrow |f(x) - f(y)| < \epsilon \ \forall f \in \mathcal{F},$$

where d is the metric on X.

Proof:

STEP I. We show that the compact metric space X is separable, i.e., has a countable dense subset S.

Given a positive integer n and a point $x \in X$, let

$$B(x, 1/n) = \{y \in X : d(x, y) < 1/n\},$$

the open ball of radius $1/n$, centered at x. For a given n, the collection of all these balls as x runs through x is an open cover of x, so (because X is compact) there is a finite subcollection that also covers X. Let S_n denote the collection of centers of the balls in this finite subcollection. Thus S_n is a finite subset of X that is "$1/n$-dense" in the sense that every point of X lies within $1/n$ of a point of S_n. Clearly the union S of all the sets S_n is countable, and dense in X.

STEP II. We find a subsequence of $\{f_n\}$ that converges pointwise on S.

This is a standard diagonal argument. Let's list the (countably many) elements of S as $\{x_1, x_2, ...\}$. Then the numerical sequence $\{f_n(x_1)\}_{n=1}^{\infty}$ is bounded, so by Bolzano-Weierstrass it has a convergent subsequence, which we will wrtie using double subscripts: $\{f_{1,n}(x_1)\}_{n=1}^{\infty}$. Now the numerical sequence $\{f_{1,n}(x_2)\}_{n=1}^{\infty}$ is bounded, so it has a convergent subsequence $\{f_{2,n}(x_2)\}_{n=1}^{\infty}$. Note that the sequence of functions $\{f_{2,n}\}_{n=1}^{\infty}$, since it is a subsequence of $\{f_{1,n}\}_{n=1}^{\infty}$ is bounded, so it has a convergent subsequence $\{f_{2,n}(x_2)\}_{n=1}^{\infty}$. Note that the sequence of functions $\{f_{2,n}\}_{n=1}^{\infty}$, since it is a subsequence of $\{f_{1,n}\}_{n=1}^{\infty}$, converges at both x_1 and x_2. Proceeding in this fashion we obtain a countable collection of subsequences of our original sequence:

$$\begin{bmatrix} f_{1,1} & f_{1,2} & f_{1,3} & \cdots \\ f_{2,1} & f_{2,2} & f_{2,3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

where the sequence in the n-th row converges at the points $x_1, ..., x_n$, and each row is a subsequence of the one above it.

Thus the diagonal sequence $\{f_{n,n}\}$ is a subsequence of the original sequence $\{f_n\}$ that converges at each point of S.

STEP III. Completion of the proof.

Let $\{g_n\}$ be the diagonal subsequence produced in the previous step, convergent at each point of the dense set S. Let $\epsilon > 0$ be given, and choose $\delta > 0$ by equicontinuity of the original sequence, so that $d(x, y) < \delta$ implies $|g_n(x) - g_n(y)| < \epsilon/3$ for each $x, y \in x$ and each positive integer n. Fix $M > 1/\delta$ so that the finite subset $S_M \subset S$ that we produced in Step I is δ-dense in X. Since $\{g_n\}$ converges at each point of S_M, there exists $N > 0$ such that

$$(*)\ n, m > N \Rightarrow |g_n(s) - g_m(s)| < \epsilon/3\ \forall s \in S_m.$$

Fix $x \in X$. Then x lies within δ of some $s \in S_M$, so if $n, m > M$:

$$|g_n(x) - g_m(x)| \le |g_n(x) - g_n(s)| + |g_n(s) - g_m(s)| + |g_m(s) - g_m(x)|$$

The first and last terms on the right are $< \epsilon/3$ by our choise of δ (which was possible because of the equicontinuity of the original sequence), and the same estimate holds for the middle term by our choise of N in (*). In summary: given $\epsilon > 0$ we have produced N so that for each $x \in X$,

$$m, n > N \Rightarrow |g_n(x) - g_m(x)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon.$$

Thus on X the subsequence $\{g_n\}$ of $\{f_n\}$ is uniformly Cauchy, and therefore uniformly convergent. This completes the proof of the Arzela-Ascoli Theorem.

Q.E.D.

7.7 Wierstrass Approximation Theorem

Go back to Table of Contents. Please click TOC

In analysis, the Stone-weierstrass Theorem (*Weierstrass approximation theorem*) provides both practical and theoretical relevance, especially in polynomial interpolation.

Theorem 7.37. *(Stone-Weierstrass.) If f is a continuous complex function on $[a, b]$, there exists a sequence of polynomials P_n such that*

$$\lim_{n\to\infty} P_n(x) = f(x)$$

uniformly on $[a, b]$. If f is real, the P_n may be taken real.

Proof: Assume, w.l.o.g., that $[a, b] = [0, 1]$. We also assume that $f(0) = f(1) = 0$. For if the theorem is proved for this case, consider

$$g(x) = f(x) - f(0) - x[f(1) - f(0)], \ for \ 0 \le x \le 1.$$

Here $g(0) = g(1) = 0$, and if g can be obtained as the limit of a uniformly convergent sequence of polynomials, it is clear that the same is true for f, since $f - g$ is a polynomial.

Moreover, we define $f(x)$ to be zero for x outside $[0, 1]$. Then f is uniformly continuous on the whole line.

We put

$$Q_n(x) = c_n(1 - x^2)^n, \ n = 1, 2, 3, ...,$$

where c_n is chosen so that

$$\int_{-1}^{1} Q_n(x)dx = 1, \ n = 1, 2, 3, ...$$

We need some information about the order of magnitude of c_n. Since

$$\begin{aligned} \int_{-1}^{1}(1-x^2)^n dx = 2\int_0^1(1-x^2)^n dx &\ge 2\int_0^{1/\sqrt{n}}(1-x^2)^n dx \\ &\ge 2\int_0^{1/\sqrt{n}}(1-nx^2)dx \\ &= \tfrac{4}{3\sqrt{n}} \\ &> \tfrac{1}{\sqrt{n}} \end{aligned}$$

and it follows that $c_n < \sqrt{n}$.

The inequality $(1-x^2)^n \ge 1-nx^2$ which we used above is easily shown to be true by considering the function $(1-x^2)^n - 1 + nx^2$ which is zero at

$x = 0$ and whose derivative is positive in $(0, 1)$. For any $\delta > 0$, $c_n < \sqrt{n}$ implies that

$$Q_n(x) \leq \sqrt{n}(1 - \delta^2)^n, \ for \ \delta \leq |x| \leq 1,$$

so that $Q_n \to 0$ uniformly in $\delta \leq |x| \leq 1$.

Now we set

$$P_n(x) = \int_{-1}^{1} f(x+t)Q_n(t)dt, \ 0 \leq x \leq 1.$$

The the assumptions about f show, by a simple change of variable, that

$$P_n(x) = \int_{-x}^{1-x} f(x+t)Q_N(t)dt = \int_{0}^{1} f(t)Q_n(t-x)dt,$$

and the last integral is clearly a polynomial in x. Thus $\{P_n\}$ is a sequence of polnomials, which are real if f is real.

Given $\epsilon > 0$, we choose $\delta > 0$ such that $|y - x| < \delta$ implies that

$$|f(y) - f(x)| < \frac{\epsilon}{2}.$$

Let $M = \sup |f(x)|$. From above and the fact $Q_n(x) \geq 0$, we have for $0 \leq x \leq 1$,

$$\begin{array}{rcl} |P_n(x) - f(x)| & = & \left| \int_{-1}^{1} [f(x+t) - f(x)]Q_n(t)dt \right| \\ & \leq & \int_{-1}^{1} |f(x+t) - f(x)|Q_n(t)dt \\ & \leq & 2M \int_{-1}^{-\delta} Q_n(t)dt + \frac{\epsilon}{2} \int_{-\delta}^{\delta} Q_n(t)dt + 2M \int_{\delta}^{1} Q_n(t)dt \\ & \leq & 4M\sqrt{n}(1-\delta^2)^n + \frac{\epsilon}{2} \\ & < & \epsilon \end{array}$$

for large enough n, which proves the theorem.

Q.E.D.

Corollary 7.38. *For every interval $[-a, a]$ there is a sequence of real polynomials P_n such that $P_n(0) = 0$ and such that*

$$\lim_{n \to \infty} P_n(x) = |x|$$

uniformly on $[p-a, a]$.

Proof: By Theorem 7.26 from text [10], there exists a sequence $\{P_n^*\}$ of real polynomials which converges to $|x|$ uniformly on $[-a, a]$. In particular, $P_n^*(0) \to 0$ as $n \to \infty$. The polynomials

$$P_n(x) = P_n^*(x) - P_n^*(0) \ (n = 1, 2, 3, ...)$$

have desired properties.

Q.E.D.

7.8 Convolution

Go back to Table of Contents. Please click TOC

Let $f(x)$ and $g(x)$ be continuous rela-valued functions for $x \in \mathbb{R}$ and assume that f or g is zero outside some bounded set (this assumption can be relaxed a bit). Define the *convolution*

$$(f \star g)(x) := \int_{-\infty}^{\infty} f(x-y)g(y)dy$$

One preliminary useful observation is

$$f \star g = g \star f$$

To prove this make the change of variable $t = x - y$ in the integral of $f \star g$.

Remark 7.39. Note that if g is zero outside of the interval $[a, b]$, then $(f \star g)(x) = \int_a^b f(x-y)g(y)dy$, so only the values of f on the interval $[x-b, x-a]$ are used. Thus if $x \in [c, d]$, then the convolution only involves the values of f on $[c-b, d-a]$.

Remark 7.40. Similarly, if f is zero outside of the interval $[-\frac{1}{2}, \frac{1}{2}]$ and $x \in [c, d]$, then the convolution only involves the values of g on the interval $[c - \frac{1}{2}, d + \frac{1}{2}]$.

SMOOTHNESS OF $f \star g$.

Theorem 7.41. *If $f \in C^1(\mathbb{R})$ then $f \star g \in C^1(\mathbb{R})$. Better yet, if $f \in C^k(\mathbb{R})$ and $g \in C^l(\mathbb{R})$, then $f \star g \in C^{k+l}(\mathbb{R})$.*

Proof: This is clearer if we write $h(x) := f(\star g)(x)$. Then

$$\frac{h9x) - h(x_0)}{x - x_0} = \int_{-\infty}^{\infty} \frac{f(x-y) - f(x_0 - y)}{x - x_0} g(x)dx.$$

We can be done if we can show that $[f(x-y) - f(x_0-y)]/(x-x_0)$ converges uniformly to $f'(x_0 - y)$. To do this we use the integral form of the mean value theorem:

$$\begin{aligned} f(x-y) - f(x_0-y) &= \int_0^1 \frac{df(x_0 - y + t(x-x_0))}{dt} dt \\ &= \left[\int_0^1 f'(x_0 - y + t(x - x_0))dt\right](x - x_0). \end{aligned}$$

Then

$$\frac{f(x-y) - f(x_0-y)}{x - x_0} - f'(x_0 - y) = \int_0^1 [f'(x_0 - y + t(x - x_0)) - f'(x_0 - y)]dt$$

Since f' is assumed continuous and is zero outside of a bounded set, it is uniformly continuous. Thus, given any $\epsilon > 0$ there is a $\delta > 0$ so that if $|x - x_0| < \delta$ then

$$f'(z + t(x - x_0)) - f'(z)| < \epsilon$$

for all values of z. In our case $z = x_0 - y$. Thus the left side of $\frac{f(x-y) - f(x_0-y)}{x-x_0} - f'(x_0 - y)$ tends to zero uniformly for all choices of x_0 and y. Consequently, $h \in C^1(\mathbb{R})$.

Repeating this we conclude that if $f \in C^k$ then $h \in C^k$. Because of commutative property $f^{(k)} \star g = g \star f^{(k)}$, so we can repeat this reasoning to show that $g \star f^{(k)} \in C^l$. Thus $f \star g \in C^{k+l}$. note that alghough g might not be zero outside a bounded set, because f is zero outside a bounded set, the integration in $g \star f^{(k)}$ is only over a bounded set - in which the derivatives of g are uniformly continuous.

Theorem 7.42. *$f_n(x) \in C^\infty$ converges uniformly to $f(x)$ for all $x \in \mathbb{R}$. Thus, on a compact set any continuous function can be approximated arbitrarily closely in the uniform norm by a smooth function.*

Proof: The smoothness of the approximations f_n is an immediate consequence of Theorem 1. Since $f(x) = f(x)(\int_{-\infty}^{\infty} \varphi_n(t)dt) = \int_{-\infty}^{\infty} f(x)\varphi_n(t)dt$,

$$f_n(x) - f(x) = \int_{|t|\leq 1/n} [f(x-t) - f(x)]\varphi_n(t)dt.$$

Since f is uniformly continuous, given any $\epsilon > 0$ there is a $\delta > 0$ such that if $|t| < \delta$ then $|f(x-t) - f(x)| < \epsilon$ for all x. If $1/n < \delta$, then by (c) above we have

$$|f_n(x) - f(x)| < \epsilon \int_{|t|\leq 1/n} \varphi_n(t)dt = \epsilon.$$

Since the right side is independent of x this shows that in the uniform norm $||f_n - f||_\infty < \epsilon$.

Since the operators $f_n(f) := f \star \varphi_n \to f$, so in this sense T_n converges to the identity operator I, we sometime call the T_n (or the φ_n) *approximate identities.*

APPROXIMATE IDENTITIES

Let $\varphi_n(t)$ be a sequence of smooth real-valued functions with the properties

$$(a)\varphi_n(t) \geq 0,\ (b)\varphi_n(t) = 0\ for\ |t| \geq 1/n,\ (c) \int_{-\infty}^{\infty} \varphi_n(t)dt = 1$$

Note: because of (b), this integral is only over $-1/n \leq t \leq 1/n$.

Assume $f(x)$ is uniformly continuous for all $x \in \mathbb{R}$ and zero outside a bounded set. Define

$$f_n(x) := (f \star \varphi_n)(x) = \int_{-\infty}^{\infty} f(x-t)\varphi_n(t)dt.$$

7.9 The Gamma Function

Go back to Table of Contents. Please click TOC

This function is closely related to factorials and crops up in many unexpected places in analysis. Its origin, history, and development are very well described in an interesting article by P.J. Davis (Amer. Math. Monthly, vol. 66, 1959, pp. 849-869). Artin's book is another good elementary introduction [4].

Definition 7.43. For $0 < x < \infty$,

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.$$

The integral converges for these x. (When $x < 1$ both 0 and ∞ have to be looked at.

Theorem 7.44. *This theorem has the following parts:*

(a) The functional equation

$$\Gamma(x+1) = x\Gamma(x)$$

holds if $0 < x < \infty$.

(b) $\Gamma(n+1) = n!$ *for* $n = 1, 2, 3, ...$

(c) $\log \Gamma$ *is convex on* $(0, \infty)$.

Theorem 7.45. *If* f *is a positive function on* $(0, \infty)$ *such that*

(a) $f(x+1) = xf(x)$

(b) $f(1) = 1$

(c) $\log f$ *is convex,*

then $f(x) = \Gamma(x)$.

Proof: Since Γ satisfies (a), (b), and (c), it is enough to prove that $f(x)$ is uniquely determined by (a), (b), and (c), for all $x > 0$. By (a), it is enough to do this for $x \in (0, 1)$. Write $\varphi = \log f$. Then

$$\varphi(x+1) = \varphi(x) + \log x, \ 0 < x < \infty,$$

$\varphi(1) = 0$, and φ is convex. Suppose $0 < x < 1$, and n is a positive integer. By the equation above, we have $\varphi(n+1) = \log(n!)$. Consider the difference quotients of φ on the intervals $[n, n+1]$,$[n+1, n+1+x]$,$[n+1, n+2]$. Since φ is convex

$$\log n \leq \frac{\varphi(n+1+x) - \varphi(n+1)}{x} \leq \log(n+1).$$

Then we have

$$\varphi(n+1+x) = \varphi(x) + \log[x(x+1)\dots(x+n)].$$

Thus,

$$0 \leq \varphi(x) - \log\left[\frac{n!x^x}{x(x+1)\dots(x+n)}\right] \leq x \log\left(1 + \frac{1}{n}\right).$$

The last expression tends to 0 as $n \to \infty$. Hence, $\varphi(x)$ is determined, and the proof is complete.

Q.E.D.

Remark 7.46. As a by-product we obtain the relation

$$\Gamma(x) = \lim_{n\to\infty} \frac{n!n^x}{x(x+1)\dots(x+n)}$$

at least when $0 < x < 1$; from this one can deduce that the Gamma function holds for all $x > 0$ since $\Gamma(x+1) = x\Gamma(x)$.

□

8 §Some Special Functions§

Go back to Table of Contents. Please click TOC

8.1 Power Series

Go back to Table of Contents. Please click TOC

In this section we shall derive some properties of functions which are represented by power series, i.e., functions of the form

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

or, more generally,

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n.$$

These are called *analytic functions*.

Theorem 8.1. *Suppose the series*

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

converges for $|x| < R$, and define

$$f(x) = \sum_{n=0}^{\infty} c_n x^n \ (|x| < R).$$

Then $\sum\limits_{n=0}^{\infty}$ converges uniformly on $[-R+\epsilon, R-\epsilon]$, no matter which $\epsilon > 0$ is chosen. The function f is continuous and differentiable in $(-R, R)$, and

$$f'(x) = \sum_{n=1}^{\infty} n c_n x^{n-1} \ (|x| < R).$$

Proof: Let $\epsilon > 0$ be given. For $|x| \leq R - \epsilon$, we have

$$|c_n x^n| \leq |c_n (R-\epsilon)^n|;$$

and since

$$\sum c_n (R-\epsilon)^n$$

converges absolutely (every power series converges absolutely in interior of its interval of convergence, by the root test), Theorem 7.10 from text [10] shows the uniform convergence of $\sum\limits_{n=0}^{\infty} c_n x^n$ on $[-R+\epsilon, R-\epsilon]$.

Since $\sqrt[n]{n} \to 1$ as $n \to \infty$, we have

$$\limsup_{n\to\infty} \sqrt[n]{n|c_n|} = \limsup_{n\to\infty} \sqrt[n]{|c_n|},$$

so that the series $f(x)$ and $f'(x)$ have the same interval of convergence.

Since $f'(x)$ is a power series, it converges uniformly in $[-R+\epsilon, R-\epsilon]$, for every $\epsilon > 0$, and we can apply Theorem 7.17 from text [10]. It follows that $f'(x)$ holds if $|x| \leq R - \epsilon$.

But, given any x such that $x| < R$, we can find an $\epsilon > 0$ such that $|x| < R - \epsilon$. This shows that $f'(x)$ holds for $|x| < R$. Continuity of f follows from the existence of f'.

Q.E.D.

Corollary 8.2. *Under the hypotheses of Theorem 8.1 (theorem above), f has derivatives of all orderes in $(-R, R)$, which are given by*

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)c_n x^{n-k}.$$

In particular,

$$f^{(k)}(0) = k!c_k \ (k = 0, 1, 2, ...).$$

(Here $f^{(0)}$ means f, and $f^{(k)}$ is the kth derivative of f, for $k = 1, 2, 3, ...$).

Proof: From $f^{(k)}(x)$, it follows that if we apply Theorem 8.1 to f, f', f'', ... Setting $x = 0$ for them, we obtain $f^{(k)}(0) = k!c_k$.

Q.E.D.

Proposition 8.3. *We have the following properties (summarized from lecture):*

(1) *Real analytic function form a vector space (close by superposition)*

(2) *Real analytic functions have k-derivatives for any $k \geq 0$, then, for $a \in E^0$, we have*

$$f(x) = \sum c_n(x-a)^n \Rightarrow f^{(k)}(a) = k!c_k$$

(3) *Uniqueness. For $a \in E^0$, we have $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n = \sum_{n=0}^{\infty} c'_n(x-a)^n \Rightarrow c_n = c'_n$ and*

$$f(x) \begin{cases} e^{-1/x^2}, & x \neq 0 \\ 0, & \end{cases}$$

Theorem 8.4. *(Abel's Theorem). Let*

$$G(x) = \sum_{k=0}^{\infty} a_k x^k$$

be a power series with real coefficients with radius of convergence 1. Suppose that the series $\sum_{k=0}^{\infty} a_k$ converges. Then $G(x)$ is continuous from the left at $x = 1$, i.e.

$$\lim_{x=1^-} G(x) = \sum_{k=0}^{\infty} a_k.$$

An alternative approach is the following:

Let $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ *and let* R *to be radius of convergence such that* $R \notin \{0, \infty\}$. *Then we have*

$$\sum_{n=0}^{\infty} c_n R^n = 1 \Rightarrow \lim_{x \to a+r} f(x) = 1$$

Example 8.5. Suppose we have

$$\log(1+x) = \sum_{n\geq 1} (-1)^{n-1}\frac{x^n}{n}, \ \sqrt{1+x} = \sum_{n \geq 0} \frac{(-1)^{n-1}(2n)!}{2^{2n}n!^2(2n-1)}$$

at $x = 1$. Set $g(x) = \sum_{n\geq 1} (-1)^{n-1}x^n/n$ for $|x| < 1$. Then $g(x) = \log(1+x)$ for $|x| < 1$. The series $g(1)$ converges since it is alternating, so by Abel's theorem

$$g(1) = \lim_{x \to 1^-} g(x) = \lim_{x \to q^-} \log(1+x) = \log 2$$

since the logarithm is a continuous function.

□

Example 8.6. Again, suppose we have

$$\log(1+x) = \sum_{n\geq 1} (-1)^{n-1}\frac{x^n}{n}, \ \sqrt{1+x} = \sum_{n \geq 0} \frac{(-1)^{n-1}(2n)!}{2^{2n}n!^2(2n-1)},$$

and we can set $g(x) = \sum_{n\geq 0} \frac{(-1)^{n-1}(2n)!}{2^{2n}n!^2(2n-1)}x^n$ for $|x| < 1$, so $g(x) = \sqrt{1+x}$ here. The series $g(1)$ is absolutely convergent, so by Abel's theorem and the continuity of $\sqrt{1+x}$,

$$g(1) = \lim_{x \to 1^-} g(x) = \lim_{x \to 1^-} \sqrt{1+x} = \sqrt{2}.$$

□

Example 8.7. Let $g(x) = 1/(1+x^2)$, which is differentiable for all real x. When $|x| < 1$, $g(x) = \sum_{n\geq 0} (-1)^n x^{2n}$ by expanding a geometric series. While $g(x)$ has a limit as $x \to 1^-$ (namely $1/2$), the power series does not converge at $x = 1$.

□

Now we can prove Abel's Theorem.

Proof: We are given that $\sum_{n\geq 0} c_n x^n$ converges for $|x| < 1$ and at $x = 1$. Our goal is to prove

$$\lim_{x \to 1^-} \sum_{n \geq 0} c_n x^n = \sum_{n\geq 0} c_n.$$

For $-1 < x < 1$ we will work with the trancated sums $\sum_{n=0}^{N} c_n x^n$ and $\sum_{n=0}^{N} c_n$. Set

$$s_n = c_0 + c_1 + \cdots + c_n$$

for $n \geq 0$. Note that $x_n - s_{n-1} = c_n$ for $n \geq 1$. Then

$$\begin{aligned}
\sum_{n=0}^{N} c_n x^n &= c_0 + \sum_{n=1}^{N} x^n (x_n - s_{n-1}) \\
&= c_0 + \sum_{n=1}^{N} u_n (s_n - s_{n-1}) \text{ where } u_n = x^n \\
&= c_0 + u_N s_N - u_1 s_0 - \sum_{n=1}^{N-1} s_n (u_{n+1} - u_n) \text{ by summation by parts} \\
&= c_0 + x^N s_N - x c_0 - \sum_{n=1}^{N-1} s_n (x^{n+1} - x^n) \\
&= (1-x) c_0 + x^N s_N + \sum_{n=1}^{N-1} s_n (x^n - x^{n+1}) \\
&= (1-x) c_0 + x^N s_N + \sum_{n=1}^{N-1} s_n (1-x) x^n.
\end{aligned}$$

Since $s_0 = c_0$, we can absorb the first term into the sum as the $n = 0$ term, and then pull a $1 - x$ out of each term in the sum:

$$\sum_{n=0}^{N} c_n x^n = x^N s_N + (1-x) \sum_{n=0}^{N-1} s_n x^n.$$

By hypothesis, the left side of the equation converges as $N \to \infty$. Also $x^N s_N \to 0$ as $N \to \infty$ since $x^N \to 0$ and s_N is bounded (in fact s_N converges since we are assuming $\sum_{n=1}^{\infty} c_n$ converges). Therefore when $-1 < x < 1$ the partial sums on the right side of the equation converge as $N \to \infty$ and we get

$$\sum_{n \geq 0} c_n x^n = (1-x) \sum_{n \geq 0} s_n x^n.$$

Let $x = \sum_{n \geq 0} c_n$. We want to show $\sum_{n \geq -} c_n x^n \to s$ as $x \to 1^-$. We subtract s from both sides of $\sum_{n \geq 0} c_n x^n = (1-x) \sum_{n \geq 0} s_n x^n$ and we have

$$\sum_{n \geq 0} c_n x^n - s = (1-x) \sum_{n \geq 0} (s_n - s) x^n,$$

where on the right side we used the formula $(1-x) \sum_{n \geq 0} x^n = 1$. Our goal is to show the right side of $\sum_{n \geq 0} c_n x^n - s = (1-x) \sum_{n \geq 0} (s_n - s) x^n$ tends to 0 as $x \to 1^-$.

By assumption, $s_n \to s$ as $n \to \infty$. Pick a positive number ϵ. For all large n, say $n \geq M$, $|s_n - s| \leq \epsilon$. Then we break up the right side of $\sum_{n \geq 0} c_n x^n - s = (1-x) \sum_{n \geq 0} (s_n - s) x^n$ into two sums

$$\sum_{n \geq 0} c_n x^n - s = (1-x) \sum_{n=0}^{M-1} (s_n - s) x^n + (1-x) \sum_{n \geq M} (s_n - s) x^n$$

and estimate

$$\begin{aligned}
\left|\sum_{n\geq 0} c_n x^n - s\right| &\leq |1-x|\sum_{n=0}^{M-1}|s_n - s||x|^n + |1-x|\sum_{n\geq M}|s_n - s||x|^n \\
&\leq |1-x| oversetM - 1\sum_{n=0}|s_n - s||x|^n + |1-x|\sum_{n\geq M}\epsilon|x|^n \\
&= |1-x|\sum_{n=0}^{M-1}|s_n - s||x|^n + |1-x|\epsilon\tfrac{|x|^M}{1-|x|} \\
&< 1-x|\sum_{n=0}^{M-1}|s_n - s||x|^n + |1-x|\epsilon\tfrac{1}{1-|x|}.
\end{aligned}$$

Taking $0 < x < 1$, $|1-x| = 1-x$ so this upper bound becomes

$$\left|\sum_{n\geq 0} c_n x^n - s\right| < |1-x|\sum_{n=0}^{M-1}|s_n - s| + \epsilon.$$

When $x \to 1^-$, the first term on the right side of $\left|\sum_{n\geq 0} c_n x^n - s\right| < |1 - x|\sum_{n=0}^{M-1}|s_n - s| + \epsilon$ tends to 0 on account of the $1-x$ there. (Note that the upper index of summation $M-1$ has nothing to do with x, so it does not change as $x \to 1^0$.) When x is close enough to 1, we can make the first term on the right side at most ϵ, so

$$\left|\sum_{n\geq 0} c_n x^n - s\right| \leq \epsilon + \epsilon = 2\epsilon$$

as $x \to 1^-$. Since ϵ is an arbitrary positive number, the left side of $\left|\sum_{n\geq 0} c_n x^n - s\right| \leq \epsilon + \epsilon = 2\epsilon$ must go to zero as $x \to 1^-$.

Q.E.D.

8.2 The Exponential and Logarithmic Functions

Go back to Table of Contents. Please click TOC

Define the following

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

and as an alternative notation we also write

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

Theorem 8.8. *Let e^x be defined on A^1 by $E(x) = e^x$ and $E(z) = \sum_{n=0}^{\infty}\frac{z^n}{n!}$.*

(a) e^x is continuous and differentiable for all x;

(b) $(e^x)' = e^x$;

(c) e^x is a strictly increasing function of x, and $e^x > 0$;

(d) $e^{x+y} = e^x e^y$;

(e) $e^x \to +\infty$ as $x \to +\infty$, $e^x \to 0$ as $x \to -\infty$;

(f) $\lim_{x\to+\infty} x^n e^{-x} = 0$, for every n.

9 §Functions of Several Variables§

Go back to Table of Contents. Please click TOC

10 §Integration of Differential Forms§

Go back to Table of Contents. Please click TOC

References

[1] Aigner, Martin, and Ziegler, Gunter M., *Proofs from The Book.*

[2] Apostol, Tom M., *Calculus*, c1967-c1969.

[3] Armstrong, M.A., *Basic Topology*

[4] Artin, E. (1964) "The Gamma Function," *Holt, Rinehart and Winston, Inc. New York,.*

[5] Dodson, C.T., Parker, P.E., Parker, Philip E., *A Users Guide to Algebraic Topology (Mathematics and Its Applications)*. 1997th Ediction.

[6] Emch, Arnold. *Some properties of closed convex curves in a plane.* Amer. J. Math, 35:407-412, 1913.

[7] Eriksson, Kenneth, Estep, Donald, and Johnson, Claes, *Applied Mathematics-Body and Soul.*

[8] Gilbarg, David, Trudinger, Neil S., *Elliptic Partial Differential Equations of Second Order.*

[9] Polya, George, Szego, Gabor, *Problems and Theorems in Analysis I*

[10] Rudin, walter, *Principles of Mathematical Analysis, 3rd Edition*, 1976.

[11] Tao, Terence, *Analysis I*, Hindustan Book Agency (India), 2006.

[12] Toeplitz, Otto. *Ueber einige Aufgaben der Analysis situs.* Verhandlungen der Schweizerischen. Naturforschenden Gesellschaft in Solothurn, 4:197, 1911.

[13] Winkler, Peter, *Mathematical Puzzles: A Connoisseur's Collection.*