Algebra and Analysis

Yuxin Gong

Imperial College London

Contents

1 **Mathematical logic** 2

1.1 Proposition and logic 2

1.2 More about propositions 2

1.2.1 AND (Conjunction) 2

1.2.2 OR (Disjunction) 3

1.2.3 NOT (Negation) 3

1.2.4 IMPLIES 4

1.2.5 Logical equivalence 4

1.3 Basic laws in logic 4

1.3.1 Properties in implication 5

1.4 Quantifiers 6

1.5 Examples and Practice: 6

2 **Sets** 7

2.1 Sets and Quantifiers 7

2.2 Operations between sets 9

3 **Functions** 12

3.1 Composition of Functions 12

4 **Groups** 12

4.1 Operations 12

4.2 Groups 15

5 **The Sylow Theorems** 16

5.1 Statements of Three Theorems 16

5.2 Preparation of Proofs 16

5.3 Proof of The Sylow's Theorems 17

6 **Metric Spaces and Convergence** 18

6.1 Metric Spaces 18

6.2 Convergence of Sequences 19

6.3 Subsequences 20

7 **Normed Vector Space** 21

7.1 Norms 21

1 Mathematical logic

"Wir müssen wissen. Wir werden wissen."

– David Hilbert

1.1 Proposition and logic

Before introducting the first definition, making sure what are common expressions, here are some examples:

- Mathematics is a good subject
- π
- $1 + 1$

The first expression is ambiguous, how to define what is a "**good**" subject? The second and third one are just some items in the Math. But in the world of Math, one can not deal with those "non-sense" things. We need definitions!!

Definition 1.1.1: Proposition

A proposition is a declarative statement that is either true or false but not both.

This means that if someone gives you a proposition in mathematics, it can be either true or false. Once you encounter expresssions like this:

This proposition is False.

You will realize this is no longer a proposition anymore. If it is true, then it is false. If it is false, then it is true. You should do this three times by yourself, you will realize that this is impossible a proposition.

1.2 More about propositions

There will also be some cases that the second proposition is related to the first proposition. Now a truth table is needed to record the T/F value of those two propositions. This note will not include formal definition of a truth table as it is a technical thing to do when combining several propositions.

1.2.1 AND (Conjunction)

Suppose that P and Q are propositions. In our real life, we always say both something and something are true. When it comes to mathematics, it becomes **AND**. Denoting **AND** as $\wedge$, $P \wedge Q$ becomes a new proposition. Now it's time to use truth table!!

Definition 1.2.1: AND $\wedge$

Truth table for conjunction is summarized as follows:

P	Q	$P \wedge Q$
T	T	T
T	F	F
F	T	F
F	F	F

1.2.2 OR (Disjunction)

Similarly, denoting **OR** as $\vee$. To connect those concepts with real life, you can actually think about a real life example. A familiy is expecting to have two children, the first proposition is that the elder child is boy, the second proposition is that the younger child is boy. Now if we combine those two propositions using an OR logic connective. The proposition will be one of the children is boy. Think about this yourself.

Definition 1.2.2: OR $\vee$

Truth table for disjunction is summarized as follows:

P	Q	$P \wedge Q$
T	T	T
T	F	T
F	T	T
F	F	F

1.2.3 NOT (Negation)

Negation is one of most famous logic connective you will use in your daily life. You can think every proposition you make in your daily life and work with its negation.

Definition 1.2.3: NOT $\neg$

Truth table for negation is summarized as follows:

P	$\neg P$
T	F
F	T

By definition, one can say that either a proposition is true or its negation is true.

1.2.4 IMPLIES

This is the connective that most people will get confused with at first. Just imagine, we will have four cases in this situation. How can we define $P \implies Q$ in each stage? First, we give the definition of it and explain it later.

Definition 1.2.4: IMPLIES $\implies$

Truth table for implies is summarized as follows:

P	Q	$P \implies Q$
T	T	T
T	F	F
F	T	T
F	F	T

Think carefully about this. Taking a mathematical[1] example, let P to be $x = y$ and let Q to be $a \times x = a \times y$, given all x, y, a are real numbers. In mathematical logic, one knows that $P \implies Q$ is True. From the table, this means Q must be True if P is True. To convince ourselves that the third row and forth row are correct, try $x = 1$, $y = 2$ and $a = 0$ and $x = 1$, $y = 2$ and $a = 1$.

1.2.5 Logical equivalence

Logical equivalence has symbol $\iff$. In below, ":=" means defined to be.

Definition 1.2.5: EQUIVALENCE $\iff$

$P \iff Q := (P \implies Q) \wedge (Q \implies P)$

You can construct a truth table for $\iff$.

P	Q	$P \implies Q$	$Q \implies P$	$Q \iff P$
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

This means once you know $P \iff Q$ is True, then P and Q are going to have same T/F value.

1.3 Basic laws in logic

You should be able to prove the following lemmas by yourself. Here, when we say "then Q", this means that the proposition Q is True.

[1] We haven't introduced what is x, what is y, even what is 1!!

Lemma 1.3.1:

Given P a proposition, then $\neg(\neg P) \iff P$.

Hint: Run all the possibilities in a truth table, check they have the same value.

Theorem 1.3.1: De Morgan's Law

Suppose that P and Q are propositions, then:

1. $\neg(P \wedge Q) \iff (\neg P) \vee (\neg Q)$
2. $\neg(P \vee Q) \iff (\neg P) \wedge (\neg Q)$

Hint: your proof should include a truth table with 8 columns.

Lemma 1.3.2: Equivalent definition

$(P \implies Q) \iff (\neg P) \vee Q$

$P \vee (Q \wedge \neg Q) \iff P$

$(P \iff Q) \iff \neg(P \vee Q) \vee (P \wedge Q)$

1.3.1 Properties in implication

You might use the following relation in your real life already, consider three propositions P, Q and R. If P can imply Q and Q can imply R, almost all the people will think that P can imply R. This is also true in mathematical logic, try to prove the following lemma.

Lemma 1.3.3: Transitivity in implication

$(P \implies Q) \wedge (Q \implies R) \implies (P \implies R)$

You now have at least two ways to prove this result when distributivity of $\vee$ and $\wedge$ are introduced later.

Lemma 1.3.4: Contrapositive

$(P \implies Q) \iff (\neg Q \implies \neg P)$

Think about how useful this result can be when we are doing math!! For example, if it is difficult to argue the statement forwards, this is a probably a way to think "backwards".

Lemma 1.3.5: Distributivity

Suppose that P,Q and R are propositions, now:

1. $P \wedge (Q \vee R) \iff (P \wedge Q) \vee (P \wedge R)$
2. $P \vee (Q \wedge R) \iff (P \vee Q) \wedge (P \vee R)$

It has a high degree of symmetry, you can remember the formula easily by recognizing it.

1.4 Quantifiers

Usually quantifiers are introduced with the logic, however, it would be better to define "set". Most common quantifiers are "$\exists$" and "$\forall$". $\exists$ means there exists at least, usually in a set. $\forall$ means for all, usually means for all elements in a set.

1.5 Examples and Practice:

Here will be an example to illustrate how the mathematical logic related to our daily proof.

Q1. Prove that if n is an integer, then n is even if and only if n^2 is even.(*Imperial IUM Course Example*)

Here is my explaination, let P be the proposition n is even and Q be the proposition that n^2 is even. Our goal is to prove that $P \iff Q$ is always true. By definition, that is to prove $P \implies Q$ and $Q \implies P$ are both true. Basic idea is to find some intermediate proposition, e.g. P_1, $P_2 \cdots P_n$. Then by transitivity $(P \implies P_1) \wedge (P_1 \implies P_2) \implies (P \implies P_2)$, if one knows that $P \implies P_1$ is true and $P_1 \implies P_2$ is true, then $P \implies P_2$ must be true. Follow this idea, proving that $P \implies Q$ is true, so is $Q \implies P$.

Q2. Prove that there is no rational number whose square is $\sqrt{2}$

Q3. Show that $P \vee Q \implies Q \vee P$, i.e $\vee$ is symmetry.(*Obvious but need proof!*)

2 Sets

Sets are the most basic concepts in mathematics. It is one of the most fundamental thing in Math.

2.1 Sets and Quantifiers

You may have already known the definition of a set, let's state again here.

Definition 2.1.1: Set

A set is a collection of **different** things.

Those things in the set are called elements of the set. For example, $\{a, b, c\}$ is a set with three elements. We usually denote those sets as capital letters. i.e. $A = \{a, b, c\}$. $\in$ is used to denote that whether something is in the set, we can say $a \in A$, $c \in A$ but $d \notin A$.

Definition 2.1.2: For all $\forall$

Suppose A is a set, E is a property.

$$\forall x \in A : E(x)$$

means that for all elements in A, it has the property E.

Pick an example for youself and be familiar with this symbol as it will simplify your work!

Definition 2.1.3: Exists $\exists$

Suppose A is a set, E is a property.

$$\exists\ x \in A : E(x)$$

means there exists an element in A, it has property E.

E is a property that can be almost everything you want, for example, if A is a set of different people, E can be "height $\geq$ 175cm". Or it can also be "weight $\leq$ 50kg".

Now I want to use the quantifiers to define something you might already seen before. A subset! The definition of this is very straightforward, just define that if a set is a subset of another set, then all the elements of this set should be elements of another set. Formally, we should define it like below:

Definition 2.1.4: Subset

Suppose A and B are two sets, A is a subset of B, or using the symbol $A \subseteq B$ if

$$\forall x \in A : x \in B$$

If $\exists\ b \in B$, $b \notin A$, we call A a **proper subset** of B, denoted by $A \subset B$. And equality of sets should then be defined as

Definition 2.1.5: Equality of Sets

Suppose A, B are sets, then $A = B$ is defined to be equivalent to

$$(A \subseteq B) \land (B \subseteq A)$$

Convince yourself with the following properties of subset:

1. $A \subseteq A$ (reflexivity)
2. $(A \subseteq B) \land (B \subseteq C) \implies (A \subseteq C)$ (transitivity)

Actually the definition of the equality of sets is exactly "antisymmetry", which you will be familiar with when studying relation. Until now, you may want to construct some subsets by yourself. For example, $\{x \mid x^2 < 2\}$, but you'd better to specify which set x are belong to, formally, it should be $\{x \in \mathbb{R} \mid x^2 < 2\}$. Otherwise, something strange will happen, you may hear about "Russell paradox". Consider the following set, $R = \{x \mid x \notin x\}$, my question is, will $R \in R$? Think about this, and the main problem here is that it uses R to construct R. Imagine in your real life, how can you construct something using the thing that does not exist!! Formal set theory(ZFC) is established to avoid this kind of questions.

Definition 2.1.6: Empty set

Suppose that A is a set, define $\emptyset_A$ as

$$\emptyset_A := \{x \in X \mid x \neq x\}$$

Below are some theorems about empty set which might be useful.

Theorem 2.1.1: Empty set possesses every property

Let E be a property, then

$$\forall x \in A : x \in \emptyset_A \implies E(x)$$

Proof:

By lemma 1.3.2,

$$(x \in \emptyset_A \implies E(x)) \iff \neg(x \in \emptyset_A) \lor E(x)$$

Since $\neg(x \in \emptyset_A)$ is always true, theorem is true. $\square$

Next theorem ensures that there is only one empty set, there can not be multiple empty sets.

Theorem 2.1.2: Uniqueness of empty set

Suppose A and B are sets, then

$$\emptyset_A = \emptyset_B$$

Denote the symbol of empty set by crossing out the symbol at the right corner of each empty set, which is $\emptyset$.

2.2 Operations between sets

Definition 2.2.1: Complement

Suppose A and B are subsets of X, then $A\backslash B$ means the complement of B in A, defined as:

$$A\backslash B \coloneqq \{x \in X \mid (x \in A) \land (x \notin B)\}$$

When X contains A, A^c denotes $X\backslash A$. Notice that in our definition, B is not necessary a subset of A.

Definition 2.2.2: Intersection

Suppose A and B are subsets of X, the intersection of A and B is denoted by $A \cap B$, defined as:

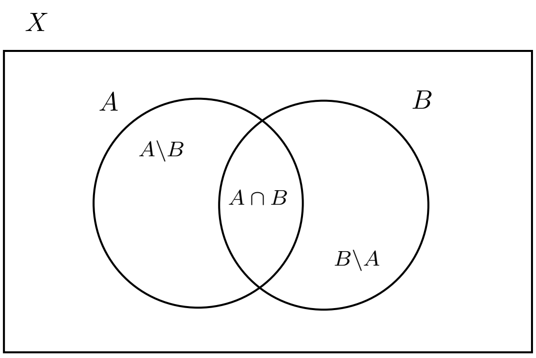
$$A \cap B \coloneqq \{x \in X \mid (x \in A) \land (x \in B)\}$$

Definition 2.2.3: Union

Suppose A and B are subsets of X, the union of A and B is denoted by $A \cup B$, defined as:

$$A \cup B \coloneqq \{x \in X \mid (x \in A) \lor (x \in B)\}$$

Venn diagrams are useful diagrams to represent the relations between sets. Following is an example to illustrate how you can use them.



Lemma 2.2.1:

(i) $X \cup Y = Y \cup X$, $X \cap Y = Y \cap X$ (Commutativity)

(ii) $X \cup (Y \cup Z) = (X \cup Y) \cup Z$, $X \cap (Y \cap Z) = (X \cap Y) \cap Z$ (Associativity)

(iii) $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$, $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$ (Distributivity)

To prove it rigorously, we need apply the definition of the equality of two sets, which is definition **2.1.5**. Here we are only going to prove the first equality. $\forall x \in X \cup Y$, $x \in X \lor x \in Y \iff x \in Y \lor x \in X$, which means $x \in Y \cup X$, hence $X \cup Y \subseteq Y \cup X$. Applying the same method will result in $Y \cup X \subseteq X \cup Y$, which implies that $X \cup Y = Y \cup X$.

Definition 2.2.4: Cartesian product

Suppose X, Y are two sets, then **Cartesian product** of X, Y are denoted by $X \times Y$, which is the set

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}$$

Two elements inside it are equal if and only if the elements in each component of the cartesian product are equal, i.e $(a, a') = (b, b')$ if and only if $a = b$ and $a' = b'$.

Proposition Let X and Y be sets,

$$X \times Y = \emptyset \iff (X = \emptyset) \vee (Y = \emptyset)$$

Proof:

To prove an if and only if statement, there are two directions for us to prove, let's first prove ' $\implies$ ' one. Prove by contradiction, suppose that $X \times Y = \emptyset$ but $(X \neq \emptyset) \wedge (Y \neq \emptyset)$. This means that $\exists\ x \in X, x = x$ and $\exists\ y \in Y, y = y$. Now $\exists\ (x, y) \in X \times Y, (x, y) = (x, y)$. Therefore $X \times Y \neq \emptyset$.

Let's that prove backwards, '$\Longleftarrow$'. Prove by contrapositive, suppose $X \times Y \neq \emptyset$, then $\exists\ (x, y) \in X \times Y, (x, y) = (x, y)$, this means that $\exists\ x \in X, x = x$ and $\exists\ y \in Y, y = y$, hence $(X \neq \emptyset) \wedge (y \neq \emptyset)$. $\square$

Definition 2.2.5: Families if Sets

Let A be a nonempty set, $\forall \alpha \in A$, let A_α be a set.

$$\mathcal{A} \coloneqq \{A_\alpha \mid \alpha \in A\}$$

is called a family of sets and A is the index set for this family.

Let X be the universal set of all sets in the family set. Intersection and union of those sets will be denoted as below:

$$\bigcap_\alpha A_\alpha \coloneqq \{x \in X \mid \forall \alpha \in A,\ x \in A_\alpha\}$$

$$\bigcup_\alpha A_\alpha \coloneqq \{x \in X \mid \exists\ \alpha \in A,\ x \in A_\alpha\}$$

Let $\{A_\alpha \mid \alpha \in A\}$ and $\{B_\beta \mid \beta \in B\}$ be families of subsets of a set X, then $(\bigcap_\alpha A_\alpha) \cap (\bigcap_\beta B_\beta) = \bigcap_{(\alpha,\beta)} A_\alpha \cap B_\beta$.

Proof:

Let $S_l = (\bigcap_\alpha A_\alpha) \cap (\bigcap_\beta B_\beta)$ and $S_r = \bigcap_{(\alpha,\beta)} A_\alpha \cap B_\beta$. $\forall x \in S_l$, we can see $x \in \bigcap_\alpha A_\alpha$ and $x \in \bigcap_\beta B_\beta$, meaning $\forall \alpha \in A$ and $\forall \beta \in B$, $x \in A_\alpha$ and $x \in B_\beta$. Thus, $\forall (\alpha, \beta) \in A \times B$, $x \in A_\alpha \cap B_\beta$, which means $x \in S_r$. Similarly, if $x \in S_r$, then $\forall (\alpha, \beta) \in A \times B$, $x \in A_\alpha \cap B_\beta$. Fix β, running through A for α, $x \in \bigcap_\alpha A_\alpha$, similarly, $x \in \bigcap_\beta B_\beta$. $\square$

There are also some interesting things to consider about if we consider the following **Proposition**,

$$\left(\bigcap_\alpha A_\alpha\right) \times \left(\bigcap_\beta B_\beta\right) = \bigcap_{(\alpha,\beta)} A_\alpha \times B_\beta$$

Proof:

Very similar method to what we have applied before except decomposing an element into two components. $\square$

Distributivity and **de Morgan's law** are also true in family of sets.

- distributivity

$$(\bigcap_\alpha A_\alpha) \cup (\bigcap_\beta B_\beta) = \bigcap_{(\alpha,\beta)} A_\alpha \cup B_\beta$$
$$(\bigcup_\alpha A_\alpha) \cap (\bigcup_\beta B_\beta) = \bigcup_{(\alpha,\beta)} A_\alpha \cap B_\beta$$

- de Morgan's law

$$(\bigcap_\alpha A_\alpha)^c = \bigcup_\alpha A_\alpha^c$$
$$(\bigcup_\alpha A_\alpha)^c = \bigcap_\alpha A_\alpha^c$$

Let's prove the first one and the fourth one to give an example and basic idea.

Proof:(First One)

Denote the set on the right hand side as S_r and set on the left hand side as S_l. Then $\forall x \in S_r$, two cases needed to be consider. $\forall \alpha \in A$, $x \in A_\alpha$, then $x \in \bigcap_\alpha A_\alpha$. Otherwise, $\exists\ \alpha_1 \in A$ such that $x \notin A_{\alpha_1}$. Then $x \in \bigcap_\beta A_{\alpha_1} \bigcup B_\beta$, which implies that $x \in B_\beta$ for all $\beta \in B$. This means $x \in \bigcap_\beta B_\beta$. In both cases, $x \in S_l$. Thus, $S_l \subseteq S_r$. The other side is obvious which you can imply by yourself. $\square$

Proof:(Fourth One)

Using similar notations for sets in both sides. $\forall x \in S_r$, $x \notin \bigcup_\alpha A_\alpha$. That means $x \notin A_\alpha$ for all $\alpha \in A$. This is another way of saying $x \in A_\alpha^c$ for all α. So $x \in S_l$ by definition. Suppose now $x \in S_r$, $x \in A_\alpha^c$ for all α. $x \notin \bigcup_\alpha A_\alpha$ which means $x \in S_r$. $\square$

3 Functions

Definition 3.0.1: Function

Suppose X, Y are two sets. A function from X to Y is a rule which for each element x of X, it associates x with an element in Y, which we denote by $f(x)$.

To have a formal definition of function, we can define functions as triples (X, G, Y), but that would be more complex if you define function like that since you need prove everything from that, so here we just use our 'ambiguous' definition. They will be enough for understand. X is called **domain** of f. Y is called **codomain** of f. There are also some important terms which we need to specify:

$$\mathrm{im}(f) \coloneqq \{y \in Y \mid \exists\, x \in X,\ y = f(x)\}$$

$$\mathrm{graph}(f) \coloneqq \{(x, y) \in X \times Y \mid x \in X\}$$

3.1 Composition of Functions

Suppose $f : X \to Y$, $g : Y \to V$ be two functions. Denote a new function as $g \circ f : X \to V$. Define it as $g \circ f(x) = g(f(x))$. Now you can prove if there are three function f, g, h with appropriate domain and codomain. Then

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

This is **associativity** in composition of functions. We have already seen functions are between two sets. So sometimes we can draw a diagram to represent different sets and functions between them.

$$\begin{tikzcd} A \arrow[r, "\phi"] \arrow[dr, "f"'] & B \arrow[d, "g"] \\ & C \end{tikzcd}$$

This leads to a definition which is designed to discover the properties of this diagram.

4 Groups

4.1 Operations

Definition 4.1.1: Operation

A function $\odot : X \times X \to X$ is called an operation on X.

To be convenient, we write $x \odot y$ instead of $\odot(x, y)$. For nonempty subsets A and B of X, denoting $A \odot B$ as:

$$A \odot B = \{a \odot b \mid a \in A,\ b \in B\}$$

A nonempty subset A of X is **closed under the operation** if $A \odot A \subseteq A$. For example, let $\odot$ be $+$ in $\mathbb{R}$. Now $\mathbb{N}$ is closed under the operation.

Definition 4.1.2: Associative

An operation $\odot$ on X is associative if

$$\forall\, x, y, z \in X,\ x \odot (y \odot z) = (x \odot y) \odot z$$

Pick the previous example here, we can see $+$ on $\mathbb{N}$ is associative. It is not hard to find some operations that are not associative since you can define operation as whatever you like. Consider the following example, let $X = \mathbb{N} \times \mathbb{N}$, define $\odot$ to be a function from $\mathbb{N}^2 \times \mathbb{N}^2 \to \mathbb{N}^2$ as following: $(a_1, b_1) \odot (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$ if $a_1 + a_2 = 1$, otherwise, define it to be $(a_1 + a_2, b_1 + b_2)$. Consider $(1,\ 0) \odot ((0,\ 1) \odot (2,\ 3))$ and $((1,\ 0) \odot (0,\ 1)) \odot (2,\ 3)$. The previous one is equal to $(3,\ 4)$ while the second one is equal to $(2,\ 3)$.

Definition 4.1.3: Commutative

An operation $\odot$ on X is commutative if

$$x \odot y = y \odot x$$

$\forall x,\ y \in X$, i.e. given any two elements in X.

The bracket matters when multiple operations taken in the same time, we should ask a question for ourselves, will they come out for the same result? Or they will have different result depending on how we put the bracket?

Don't be confused that in the following theorem, the parentheses have the same meaning as brackets.

Theorem 4.1.1: Arbitrary parentheses

Let $\odot$ be an associative operation on a set X. Then the value of any valid expression only involving $\odot$, elements of X and parentheses, is independent of the placement of the parentheses.

Proof:

One can define the result of one way recursively. Let $K_1 := a_1$ and $K_{n+1} = K_n \odot a_{n+1}$. Now K_n is something like

$$(\cdots((a_1 \odot a_2) \odot a_3) \odot \cdots) \odot a_{n-1}) \odot a_n$$

Denote any expression with length n to be A_n. The thing left for us is to prove that $A_n = K_n$. By definition of associative operation, we know that $A_3 = K_3$ and there is nothing to prove for $n = 2$ or $n = 1$. For A_{n+1}, by considering the last operation it should take, it can must be written as $A_l \odot A_m$, where $m + l = n + 1$.

- Case 1

  $m = 1$, now $A_{n+1} = A_n \odot a_{n+1}$. By induction, we know $A_n = K_n$, and because $K_{n+1} = K_n \odot a_{n+1}$, this tells us that $A_{n+1} = K_{n+1}$.

- Case 2

  $m > 1$, by induction, A_m has the form $A_{m-1} \odot a_{n+1}$. This gives that $A_{n+1} = A_l \odot A_m = A_l \odot (A_{m-1} \odot a_{n+1}) =^a (A_l \odot A_{m-1}) \odot a_{n+1}$. Since $A_l \odot A_{m-1} = A_{l+m-1} = K_n$. Hence $A_{n+1} = K_{n+1}$.

We finish this proof by claiming again that we make use of induction to prove this theorem. □

[a] Prove this equality by your own

To simplify how we write the expression, an expression of length n, is written as

$$a_1 \odot a_2 \odot \cdots \odot a_n$$

Without any parentheses.

Definition 4.1.4: Identity of an operation

Let $\odot$ be an operation on X. Any element e of X such that $\forall x \in X$

$$e \odot x = x \odot e = x$$

is called an identity element.

Lemma 4.1.1: Uniqueness of Identity

There is at most one identity element of one operation.

4.2 Groups

Definition 4.2.1: Group

A group is a nonempty set G and an operation $\odot$ on G with following properties hold:

(G1) $\odot$ is associative.

(G2) $\exists\ e \in G$, $\forall\ g \in G$, $e \odot g = g \odot e = g$.

(G3) $\forall\ g \in G$, $\exists\ g^{-1} \in G$, such that $g \odot g^{-1} = g^{-1} \odot g = e$.

When $\odot$ is commutative, we call such a group **Abelian** group. Here are some very basic properties of a group with some new definitions.

Definition 4.2.2: Order of a group

The **order** of a group G is the number of elements it contains, usually denoted by $|G|$.

If $|G|$ is finite, the group is called finite group, otherwise, it is called infinite group.

Lemma 4.2.1: Uniqueness of inverse

For every element $g \in G$, there exists a unique inverse, which is denoted by g^{-1}.

Proof:

Suppose g_1 and g_2 are both the inverses of g. Now

$$g_1 = g_1 \odot e = g_1 \odot (g \odot g_2) = (g_1 \odot g) \odot g_2 = g_2$$

where the second equality is obtained by G1. □

There are also some laws that you will be very familiar with when you are working with daily real number operations.

Lemma 4.2.2: Cancellation Law

Let g_1, g_2 and g_3 be elements of a group G, then

$$g_1 \odot g_2 = g_1 \odot g_3 \implies g_2 = g_3$$

You can verify easily that if g, h are elements of a group G, then

$$(g \odot h)^{-1} = h^{-1} \odot g^{-1}$$

Below is a list of groups and they properties.

- The $n \times n$ general linear group. This is the group of all invertible $n \times n$ matrices. And it is denoted by

$$GL_n = \{n \times n \text{ invertible matrices } A\}$$

- The group of permutations of the set of indices $\{\mathbf{1}, \mathbf{2}, \cdots, \mathbf{n}\}$ is called symmetric group, denoted by S_n.

- The set of integers with addition $+$, denoted by $\mathbb{Z}^+$.

5 The Sylow Theorems

5.1 Statements of Three Theorems

These are notes taken from the book Algebra written by Artin. A Sylow p-subgroups is a p group which is a subgroup of that group whose index is not divisible by p.

Theorem 5.1.1: First Sylow Theorem

A finite group whose order is divisible by a prime p contains a Sylow p-subgroup.

Apply this theorem, we obtain following Corollary. A finite group whose order is divisible by p contains an element of order p. This can be proved by following argument: call this group G, then it contains a Sylow p-subgroup. Pick an element x which is different from identity, its order must be positive power of p. Suppose it's p^k, then x^{k-1} has order p.

Theorem 5.1.2: Second Sylow Theorem

Let G be a finite group whose order is divisible by a prime p.

- The Sylow p-subgorups of G are conjugate subgroups
- Every subgroup of G that is a p-group is contained in a Sylow p-subgroup.

The first point of this theorem is to say that, given $G_1, G_2 \trianglelefteq G$, there exists $g \in G$, s.t. $G_1 = gG_2g^{-1}$. Additionally, given a Sylow p-subgroup H of G, $\forall g \in G$, gHg^{-1} is also a Sylow p-subgroup.

Theorem 5.1.3: Third Sylow Theorem

Let G be a finite group whose order n is divisible by a prime p. Given that $n = p^e m$ where m is not divisible by p. Then if s is the number of Sylow p-subgroup, then $s|m$ and $s \equiv 1 \pmod{p}$.

The Third theorem is very strong so that it can help us to classify some groups whose order are not large enough. The things left to do is to prove those three theorems.

5.2 Preparation of Proofs

Before proving those theorems, let's first introduce a simple group action. Suppose that G acts on set S. Given a subset U of S, define

$$gU := \{gu \mid u \in U\}$$

One can easily verify that $|gU| = |U|$. Hence if we pick all the subsets of G with order $|U|$, G acts on this set with operation defined as $(g, U) \mapsto gU$. Axioms for group action can be checked easily.

Lemma 5.2.1:

G is a group and $U \subseteq G$. If G acts on the set of all subsets of G with order $|U|$ by left multiplication, then $|\text{Stab}(U)|$ divides both $|U|$ and $|G|$.

By counting formula $|\text{Stab}(U)||O_U| = |G|$, therefore the second divisibility is trivial. To prove the first one, we should define a relation of U where

$$R(u_1, u_2) \iff u_1 \in \text{Stab}(U)u_2$$

By checking reflexive, symmetry and transitivity, it can be proved that this is an equivalence relation. By some simple implications, the first divisibility can be proved.

Lemma 5.2.2:

Let n be an integer of form $p^e m$, where $e > 0$ and p does not divide m. The number N of subsets of order p^e in a set of order n is not divisible by p.

This is simply a number theory problem, since we can directly calculate the number N, which is

$$\frac{n(n-1)\cdots(n-k)\cdots(n-p^e+1)}{p^e(p^e-1)\cdots(p^e-k)\cdots 1}$$

Let's prove by contradiction that this number is not divisible by p. Suppose it's divisible by p, then

$$\frac{n(n-1)\cdots(n-k)\cdots(n-p^e+1)}{p^e(p^e-1)\cdots(p^e-k)\cdots 1} = pq$$

where q is an arbitrary natural number which is not 0. Hence

$$n(n-1)\cdots(n-k)\cdots(n-p^e+1) = p^e(p^e-1)\cdots(p^e-k)\cdots 1 \cdot p \cdot q$$

Conisder the term $n - k = p^e m - k$, if k is not divisible by p, then $n-k$ is not divisible by p. Thus, suppose we pick all the ks where k is divisible by p. k can be written as $p^{k_i} \cdot l$. Since $k_i < e$, $n-k$ is divisible by p^{k_i} but not p^{k_i+1}. So left hand side is divisible by $\sum_{\text{all } k} k_i$ but not $\sum_{\text{all } k} k_i + 1$. However right hand side is divisible by $\sum_{\text{all } k} k_i + 1$ because $p^e - k = p^e - p^{k_i} \cdot l$, which is divisible by p^{k_i}. Contradiction!

5.3 Proof of The Sylow's Theorems

Proof of the first Sylow's Theorem: Let's recall the first theorem needs to prove that a group G with order $n = p^e m, e > 0$ must have a Sylow p-subgroup, which is just a subgroup of order p^e. This construction is not straightforward, picking set S which consists of subsets of G of order p^e. We know orbits partition set S, so

$$|S| = \sum_{\text{orbits } O} |O|$$

According to lemma 4.2.2, $|S|$ is not divisible by p. This means at least one orbit's order is not divisible by p. Call this subset U, now $|\text{Stab}(U)|$ divides $|U|$, $|\text{Stab}(U)|$ must be some positive power of p. Since $|O_U|$ is not divisible by p, the only possibility is that $|O_U| = m$ and $|\text{Stab}(U)| = p^e$. $\square$

Proof of Sylow's Second Theorem: Suppose there is a p-subgroup K. The equivalent statement of Sylow's Second Theorem is that, given a Sylow p-subgroup H, there exists $g \in G$, such that $K \subseteq gHg^{-1}$. Construct a set $\mathcal{C}$ for which G acts on which has the folloing properties. p does not divide $|\mathcal{C}|$ and $\exists\ c \in \mathcal{C}$ such that $\mathrm{Stab}(c) = H$. The existence of this set can be proved easily since that the set of left cosets of H is one of examples. Since the order of this set is m and $\mathrm{Stab}(H) = H$. Additionally, the action is transitive. Now restrict our action for only K, then there must exists $c' \in \mathcal{C}$ which c' is a fixed point. That is $k(c') = c'$ for all $k \in K$. This meaning $K \subseteq \mathrm{Stab}(c')$. $c' = g(c)$, meaning $\mathrm{Stab}(c') = g\mathrm{Stab}(c)g^{-1}$. □

Proof of Sylow's Third Theorem: Use the symbols in our statement of the third Sylow's Theorem. Let S be the set of all the Sylow p-subgroup. Now G acts on S by conjugation. Moreover, we know that this action is transitive by Second Sylow's Theorem. Pick one particular Sylow p-subgroup H. Call its stabilizer $N = N(H)$. By counting formula, $|G| = |N||O_H| = |N||S|$. Since $H \subseteq N$, then $|H|||N|$ as H is a subgroup of N. Using those two identities, you can actually deduce $s|m$. Now let's prove the second identity, where we focus on the group action of H on S. The orbit of H is just itself because H will be fixed by any element from H by conjugation. Any other set who is not fixed by H will have orbit of order greater than 2, and the order divides a positive power of p. Therefore we only need to prove the only set that is fixed by H is H itself. Suppose that there is another set H' which is fixed by H. Then the normalizer N' of H' contains H and H' as Sylow p-subgroups. $\exists\ n \in N$, such that $nH'n^{-1} = H$, however H' is normal in N'. Thus, $H = H'$. □

6 Metric Spaces and Convergence

6.1 Metric Spaces

It's important to abstract the model of $\mathbb{R}$ with respect to absolute value $|\cdot|$ and extend it into more general spaces. Metric spaces are spaces which have some key properties of $\mathbb{R}$.

Definition 6.1.1: Metric

Let X be a set. A function d is said to be a **metric** on X if $d : X \times X \to \mathbb{R}$ with three properties:

(M1) $d(x, y) = 0 \iff x = y$

(M2) $d(x, y) = d(y, x)$

(M3) $d(x, y) \leq d(x, z) + d(z, y)$

Such a tuple (X, d) is called a metric space. Similar inequalities hold in metric space, one of them is reverse triangle inequality. Let's prove that

$$d(x, y) \geq |d(x, z) - d(z, y)|$$

This is because that $d(x, y) + d(y, z) \geq d(x, z)$, that is $d(x, y) \geq d(x, z) - d(z, y)$. Since same result applies to $d(y, x)$, leading $d(y, x) \geq d(y, z) - d(z, x)$. □

Definition 6.1.2: Open Ball and Closed Ball

Let (X, d) be a metric space. $a \in X$ and $r > 0$, the open ball with center at a and radius r is defined to be

$$\mathbb{B}(a, r) := \{x \in X \mid d(a, x) < r\}$$

the closed ball with center a and radius $r > 0$ is

$$\bar{\mathbb{B}}(a, r) := \{x \in X \mid d(a, x) \leq r\}$$

6.2 Convergence of Sequences

In general, a sequence can be imaged as a row of elements from a set. In mathematics, we define it rigorously as a function.

Definition 6.2.1: Sequence

Let X be a set, a **sequence** is a function ϕ from $\mathbb{N}$ to X. Often denoted by

$$(x_n), (x_n)_{n\in\mathbb{N}}$$

where $x_n = \phi(n)$.

Almost all is a term which can be used to describe some properties in sequence. Let (x_n) be a sequence in X and let E be a property. We say that E holds for almost all terms of (x_n) if $\exists\ m \in \mathbb{N}$ such that $E(x_n)$ holds for $n \geq m$. Now back to the case where (X, d) is a metric space. $U \subseteq X$ is a **neighborhood** of $a \in X$ if $\exists\ r > 0$ such that $\mathbb{B}(a, r) \subseteq U$. Set of all neighborhood of the point a is denoted by $\mathcal{U}(a)$.

Let's think about a sequence in a space. For example $\mathbb{R}^2$,

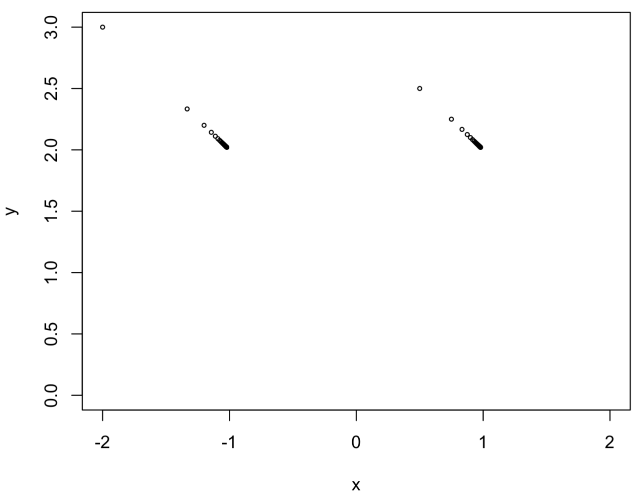


Figure 1: Sequence $\left((-1)^n - \frac{1}{n}, 2 + \frac{1}{n}\right)$

It can be seen from the diagram that $(-1, 2)$ and $(1, 2)$ are two points with special properties. We will define those points as cluster points. $a \in X$ is a **cluster point** of (x_n) if every neighborhood of a contains infinitely many terms of the sequence. This is equivalent to $\forall\, U \in \mathcal{U}(a), \forall\, m \in \mathbb{N}$, there is some $n \geq m$, such that $x_n \in U$. Now we are able to define the limit of a sequence.

Definition 6.2.2: Limit of Sequence

A sequence (x_n) coverges with limit a if each neighborhood of a contains almost all the terms of the sequence. It is denoted by

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

Some statements are more useful in proving the limit of the sequence. Let's prove the following one is equivalent to the definition of the limit.

$$\forall\, \epsilon > 0, \exists\, N \in \mathbb{N}, \forall\, n \geq N, x_n \in \mathbb{B}(a, \epsilon).$$

Since $\mathbb{B}(a, \epsilon)$ is a neighborhood, so implication from $\lim x_n$ to this is obvious. Now suppose this holds, then $\forall\, U \in \mathcal{U}(a)$, $\exists\, r > 0$, such that $\mathbb{B}(a, r) \subseteq U$. $\exists\, N \in \mathbb{N}$, $\forall\, n \geq N$, $x_n \in \mathbb{B}(a, r) \subseteq U$. $\square$

It's worth to pick an example of metric space and limit of the sequence. Let $(X_i, d_i), 1 \leq i \leq n$ to be metric spaces. $X := X_1 \times \cdots \times X_n$. Define

$$d(x, y) := \max_{1\leq i\leq n} d(x_i, y_i)$$

where $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$. d is called a **product metric**. Let's show that d is a metric on X.

We first verify M1. If $d(x, y) = 0$, then $\max\limits_{1\leq i\leq n} d(x_i, y_i) = 0$, this implies that $d(x_i, y_i) = 0$ for all i. Therefore $x = y$. For M2, since $d(x_i, y_i) = d(y_i, x_i)$, $d(x, y) = d(y, x)$. To prove M3, pick another $z \in X = (z_1, \ldots, z_n)$. $d(x, y) = \max d(x_i, y_i) \leq \max(d(x_i, z_i) + d(z_i, y_i)) \leq \max d(x_i, z_i) + \max d(z_i, y_i) = d(x, z) + d(z, y)$. Thus it follows that (X, d) is a metric space. $\square$

Associate this metric space with limit of sequence. Let's prove sequence $(x_m) = ((x^1_m, \ldots, x^n_m))_{m\in\mathbb{N}}$ converges to point $a = (a^1, \ldots, a^n)$ iff the sequence $(x^j_m)_{m\in\mathbb{N}} \to a^j$ for all j. This is quite obvious if you write down the definitions.

There are also some important definitions for a subset of metric space.

Definition 6.2.3: Bounded Sets

A subset $Y \subseteq X$ is called d-bounded in X if there is some $M > 0$ such that $d(x, y) < M$ for all x, y in Y.

6.3 Subsequences

We already identify a sequence as a function from $\mathbb{N}$ to X. It would be better if we can define a subsequence of a sequence as a function too.

Definition 6.3.1: Subsequence

Suppose $\psi : \mathbb{N} \to \mathbb{N}$ is a strictly increasing function. And ϕ is a sequence. Then $\phi \circ \psi$ is called a subsequence of ϕ. Denote this subsequence as

$$(x_{n_k})_{k\in\mathbb{N}}$$

where $n_k = \psi(k)$.

Notice that a subsequence is itself a sequence. Therefore, we say a subsequence converges to a if

$$\forall\ \epsilon > 0, \exists\ N \in \mathbb{N}, \forall\ k \geq N, x_{n_k} \in \mathbb{B}(a, \epsilon)$$

which is just rewriting the definition of convergence of sequence. There are some lemmas which are useful for subsequences which proofs are left for readers.

Lemma 6.3.1:

If (x_n) is a convergent sequence with limit a, then all its subsequences converge to a

Lemma 6.3.2:

A point a is a cluster point of a sequence (x_n) iff there is some subsequence $(x_{n_k})_{k\in\mathbb{N}}$ converges to a.

7 Normed Vector Space

So far we are looking at the 'distance' in only one set. What about 'distance' between vectors? Similarly, we can define some good things in a vector space.

7.1 Norms

Definition 7.1.1: Norm

Let E be a vector space over $\mathbb{K}$. A function $\|\cdot\| : E \to \mathbb{R}^+$ is called a norm if the following holds:

(N1) $\|x\| = 0 \iff x = 0$

(N2) $\|\lambda x\| = |\lambda|\,\|x\|$, $x \in E$ and $\lambda \in \mathbb{K}$

(N3) $\|x + y\| \leq \|x\| + \|y\|$

A pair $(E, \|\cdot\|)$ is called a normed vector space if $\|\cdot\|$ is a norm on E. A normed vector space is also a metric space since a metric can be defined as $d(x, y) := \|x - y\|$. It has a special name: **metric induced from the norm**. You can check M1 - M3 one by one. The **reversed triangle inequality** in norm can be expressed as:

$$\|x - y\| \geq |\|x\| - \|y\||$$

proof of this can be derived directly from metric space property by considering $\|x\| = d(x, 0)$ and $\|y\| = d(0, y)$. Some similar terms can be defined also in normed vector spaces. An open ball with center a and radius $r > 0$ is defined to be

$$\mathbb{B}(a, r) \coloneqq \{x \in E \mid \|x - a\| < r\}$$

and closed ball is defined to be $\bar{\mathbb{B}}(a, r) \coloneqq \{x \in E \mid \|x - a\| \leq r\}$. We then use the definition of bounded sets in metric space to define what is meant to be bounded in a normed vector space.

Definition 7.1.2: Bounded Set in Normed Vector Space

A subset X of E is called bounded in E if it is bounded in the induced metric space.

Some equivalent statements will also be used to describe that a set is bounded in normed vector space. One is the most familiar one you will encounter in many different notes:

$$X \subseteq E \text{ is bounded if and only if there is some } r > 0 \text{ such that } X \subseteq r\mathbb{B}$$