Basic notions in math[1]

A course instructed by Amir Yehudayoff, Department of Mathematics, Technion-IIT

[1] An apology: this text probably contains errors.

Contents

Chapter 1

Preliminaries

High-level goals

This is course is meant to introduce basic notions and notation. Deeper ideas will be discussed in later courses.

Mathematics

Math is a language for thought. Mathematics consists of three main ingredients: definitions, theorems and proofs. Each serves a unique purpose in the thought process.

Chapter 2

Logic

Logic is a language that allows to build solid mental structures. It is fundamental in math, but also in science and technology.

It has two complementing levels. It is good to try to switch between them as you learn new things (in this course and in others).

— A *syntactic* level that deals with the "symbols on the page".

— A *semantic* level that is about the "meaning of these symbols".

The meaning is what is ultimately interesting but the syntax allows to be systematic and locate gaps in our understanding. The syntax is also helpful in communication and in technology.

Truth values

What is "truth"? In logic it can be represented syntactically:

TRUE and FALSE

1 and 0

T and F

...

Remark. *The "truth" in math is not philosophical, it says nothing about the "world". The connection to the real world (which is important and drives a lot of math) is made outside math.*

Notation. *We mostly use the "0,1" notation in this course. So for now we introduce the set* $\{0,1\}$*. This set contains the two elements* 0 *and* 1*. We write* $x \in \{0,1\}$ *to say "*x *is either* 0 *or* 1*". Next time, we shall talk about sets in more detail.*

Boolean functions

Remark. *George Bool was a logician in the 19th century.*

Notation (variables)**.** *Symbols like $x, y, z, \ldots$ denote variables. You know this from your previous education. In basic logic, variables are Boolean, they take values in $\{0, 1\}$.*

1**-variate.** There are two functions in one variable:

Example. *The identity*

$$f(x) = x.$$

Example. *Negation*

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

Remark. *There are many ways to represent functions. By formulas. Over the reals you have seen graphs. In the Boolean setting, the most straightforward way is via truth tables.*

Exercise 1. *Draw the two truth tables.*

2**-variate.** Some examples of functions in two variables:

Example. *The logical operation AND is denoted by $\wedge$. E.g. $0 \wedge 1 = 0$ and $1 \wedge 1 = 1$. Read these sentences out loud. If we think of the two possible truth values as variables, we get a function*

$$f(x, y) = x \wedge y.$$

The input to this function is one of the four options:

$$(0, 0), (0, 1), (1, 0), (1, 1).$$

The output is either 0 or 1. We can pictorially represent it in a truth table. Draw it.

Example. *A second similar option is OR which is denoted by $\vee$. Draw its truth table.*

Example. *A third function is XOR (exclusive or); it is denoted by by $\oplus$. Draw its truth table.*

Remark. *In day-to-day life, OR is typically confused with XOR. Examples?*

Example. *A fourth option is IMPLIES which is denoted by $x \to y$. Draw its truth table. There is something confusing about this function.*

Exercise 2. *How many Boolean functions in two variables are there?*

n-variate. Instead of two variables, we can have n variables $x_1, \ldots, x_n$. Each takes the value 0 or 1. We have a function $f(x_1, \ldots, x_n)$ that assign to each list of n 0s and 1s a truth value.

Example. *Multivariate AND*

$$\bigwedge_{i=1}^{n} x_i = x_1 \wedge x_2 \wedge \ldots \wedge x_n.$$

Remark. *The fact that there are no brackets requires explanation (associativity).*

Example. *Multivariate OR*

$$\bigvee_{i=1}^{n} x_i = x_1 \vee x_2 \vee \ldots \vee x_n.$$

Example. *Multivariate XOR*

$$\bigoplus_{i=1}^{n} x_i = x_1 \oplus x_2 \oplus \ldots \oplus x_n.$$

Remark. *Each function of this type has a truth table with 2^n rows and $n+1$ column (n columns for the input and one column is for the output).*

Exercise 3. *How many Boolean functions in n variables are there?*

Expressions

An expression of the form

$$x_1 \wedge (x_2 \vee x_3)$$

defines a boolean function. This is similar to how $x+1$ is defined over the real numbers.

Example. *Here's an expression:*

$$(\neg x) \vee (x \wedge y).$$

Draw the truth table.

Remark. *An important property that we just briefly mention is that $\wedge, \vee$ and $\neg$ are* complete*: If $f(x_1, \ldots, x_n)$ is a Boolean function then there is an expression using $\wedge, \vee$ and $\neg$ and the variables $x_1, \ldots, x_n$ that represents it.*

Remark. *The function $NAND(x, y) = \neg(x \wedge y)$ is also complete.*

Remark. *This "completeness" is useful; every computer can be built just from a single gate!*

Equivalence

Example. *Consider the two expressions:*

$$(\neg x) \vee (x \wedge y) \quad \textit{and} \quad x \to y.$$

They are syntactically different but semantically the same; they define the same function. They are equivalent*.*

Notation. *We denote this by*

$$(\neg x) \vee (x \wedge y) \;\equiv\; x \to y.$$

Example. *The contra-positive rule can be expressed as*

$$x \to y \;\equiv\; \neg y \to \neg x.$$

Exercise 4. *Verify correctness of the contra-positive equivalence.*

Quantifiers

There is one more important component in the language of logic: quantifiers. They allow to state universal statements, or the existence of some interesting object.

Notation. *There are two qualifiers: a "for all" quantifier denoted by* $\forall$ *and a "there exists" quantifier denoted by* $\exists$*.*

Remark. *We do not go into the full depth of the theory here. We focus on a few examples.*

Example ("x implies y")**.** *Consider the qualified formula:*

$$\forall x \in \{0,1\}, y \in \{0,1\}\ x \to y.$$

(Try to read this and the following statements out loud.) Is this statement true? Of course not: it is not true that always x implies y. Formally, the truth value of $1 \to 0$ *is* 0*.*

Example ("x and not x imply y")**.**

$$\forall x \in \{0,1\}, y \in \{0,1\}\ (x \wedge \neg x) \to y.$$

Is this statement true? Yes. "From a contradiction, we can prove what ever we want."

Example (repeated ANDs)**.** *We can have a statement of the form*

$$\forall x_1 \in \{0,1\}, x_2 \in \{0,1\}, \ldots, x_n \in \{0,1\}\ f(x_1, \ldots, x_n).$$

This statement can be true or false. To check the validity of such a statement we need to go over all quantified variables and check if each option the value of f is 1*. This formula is true iff the following is true*

$$f(0, \ldots, 0, 0) \wedge f(0, \ldots, 0, 1) \wedge f(0, \ldots, 1, 0) \wedge \ldots \wedge f(1, \ldots, 1, 1).$$

Example ("sometimes x implies y")**.**

$$\exists x \in \{0,1\}, y \in \{0,1\}\ x \to y.$$

Is this statement true? Yes. For example, $x = y = 1$ proves the validity.

Example (repeated ORs)**.** *We can have a statement of the form*

$$\exists x_1 \in \{0,1\}, x_2 \in \{0,1\}, \ldots, x_n \in \{0,1\}\ f(x_1, \ldots, x_n).$$

This statement is true if for some assignment to the variables f returns the value 1*. (The OR of all options is true.)*

Example (mixing quantifiers)**.**

$$\forall x \in \{0,1\} \exists y \in \{0,1\}\ x \to y.$$

Is this true? Yes. In this case, the value $y = 1$ always works.

Remark. *We can build complicated expressions using this language. Computing the truth value of such expressions is in general very difficult (there is a way to formalize this hardness).*

Example (operations)**.** *There are standard operations that preserve truth values. For example,*

$$\forall x_1, \ldots, x_n \in \{0,1\} f(x_1, \ldots, x_n) \ \equiv\ \neg(\exists x_1, \ldots, x_n \in \{0,1\} \neg f(x_1, \ldots, x_n)).$$

Summary. *Boolean functions, truth tables, variables, expression, equivalence, quantifiers, quantified formula.*

Math

Going back to intro to math. This formal language of logic is not used on a day-to-day basis in most of mathematics. However, the idea is that every theorem you see or prove can be translated back to this language.

Remark. *This language also allows to build "real-world" proofs in a way that makes error detection easy, both by the writer and the reader.*

Remark. *It is sometimes difficult to understand the "idea" from the formalism. The proof often starts with intuition (a picture or a vague idea) and the formalism is found in small steps.*

Remark. *I recommend drawing as many pictures as you can. Try to find meaningful pictures!*

Remark. *In all discussion above, the truth value of a statement was absolute. "There were no proofs."*

Axioms and proofs

In math, there is one more logical construct: **axioms**. Axioms are expressions (that encode statements) that are *assumed* to be true.

Example. *"if x, y are numbers then $x + y$ is a number". Here the syntax also contains the symbol $+$.*

Remark. *We also did not formally define "numbers".*

Proofs show how to combine axioms using the rules of logic to deduce more complicated statements ("theorems"). Pictorially this looks like a directed graph – draw.

Remark. *There is a lot more structure and theory in logic, some of which is learned in the Logic course. Here we described basic notions and examples that will help you understand and use to the amazing language of mathematics.*

Chapter 3

Sets

A set is an intuitive object. It is a collection of "things". Naively, a set should have the property that we can say what is in it and what is not (this is the "operational" or "operational" meaning of sets).

Remark. *The basic property of a set S is that we should be able to answer a question of the form "is X in S?". This is a yes-no question, and we should always be able to answer it, no matter what X is.*

This approach was initiated by Cantor in the 19th century. He tried to build a mathematical formalism that capture "sets". Later on, Godel and others showed that things are not as simple as that. In this course, we follow the "naive" approach.

Basic examples

We already saw the set $\{0, 1\}$. In general, $\{\ldots\}$ denotes a set. Inside the brackets we describe the elements of the set. We can do this explicitly, as in

$$\{1, 5, 6, 12\}$$

or

$$\{1, 2, 3, 4, \ldots\}.$$

Remark. *The order of the elements in S is not important. For example,*

$$\{1, 5, 6\} = \{5, 6, 1\}.$$

Definition 5. *The size $|A|$ of a* finite *set A is the number of elements in it. For example,*

$$|\{1, 2, 4\}| = 3.$$

Remark. *We shall discuss size (cardinality) of infinite sets later on.*

Example. *A very important set is the empty set $\emptyset$. Its size is zero.*

Sets of numbers. We describe some basic sets of numbers.

Notation. *The natural numbers*

$$\mathbb{N} = \{1, 2, 3, 4, \ldots\}.$$

These numbers correspond to "counting items".

Notation. *For a natural number $n \in \mathbb{N}$, we denote by $[n]$ the set $\{1, 2, 3, \ldots, n\}$.*

Notation. *If we want the number to be closed under subtraction then we get the integers*

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, 3, \ldots\}.$$

Notation. *If we want the numbers to be closed under division then we get the rationals*

$$\mathbb{Q} = \{\tfrac{p}{q} : p, q \in \mathbb{Z} \wedge q \neq 0\} = \{\tfrac{p}{q} : p, q \in \mathbb{Z}, q \neq 0\}.$$

Here we already introduced new notation. We read it as the set of all things of the form $\frac{p}{q}$ so that p and q are integers and q is non zero. The AND can be either $\wedge$ or a comma (","); commas are more common.

Remark. *In fact, $\mathbb{Q}$ is not exactly this set. For example, syntactically $\frac{2}{3}$ and $\frac{4}{6}$ are different, but as numbers they are the same. More on this subtle issue later on.*

Notation. *More generally, we read*

$$\{A : B\}$$

as "the set of all A's that satisfy condition B".

Example. *We can now write the set of even numbers as*

$$\{2z : z \in \mathbb{Z}\}.$$

Remark. *We used the numbers $0, 1, \ldots$ even though we did not define them. One can define them using set theory, but we shall not do this here.*

Example. *A more complicated set is the set of real numbers $\mathbb{R}$. Pictorially, we think of $\mathbb{R}$ as the set of all points on the line. There are several ways to defined $\mathbb{R}$. In this course, we shall use the pictorial image and your prior knowledge. Intuitively, starting from the rationals $\mathbb{Q}$, if we want our numbers to be closed under then we get $\mathbb{R}$.*

Containments

Notation. *We denote the fact that X is in S by $X \in S$, and the fact that X is not in S by $X \notin S$.*

Example. $1 \in \mathbb{N}$ *but* $-3 \notin \mathbb{N}$.

Example. *The following statement is true*

$$\forall X \in \mathbb{N}\ X \notin \emptyset.$$

Notation (Containment)**.** *We write $B \subseteq A$ if A contains B; that is, if $x \in B \to x \in A$. If A does not contain B we write*

$$B \not\subseteq A.$$

Example. $\mathbb{N} \subset \mathbb{Z}$.

Example. *If S is a set then $\emptyset \subseteq S$. The statement*

$$X \in \emptyset \to X \in S$$

is true (in an "empty" sense, as we saw when we discussed logic).

Example. $\mathbb{R} \not\subseteq \mathbb{Q}$. *This is a basic and non trivial statement. In fact, we can point to a specific irrational number. The claim is that*

$$\sqrt{2} \in \mathbb{R} \quad \& \quad \sqrt{2} \notin \mathbb{Q}.$$

We shall not prove this claim. Proving that $\sqrt{2} \in \mathbb{R}$ requires knowing what $\mathbb{R}$ is. The standard proof of $\sqrt{2} \notin \mathbb{Q}$ uses that following fact: in every non empty set of natural numbers there is a minimal element. The property (as we shall see later on) is the basis for induction.

Notation. *Two sets A and B are equal (denoted by $A = B$) if $A \subseteq B$ and $B \subseteq A$. If A and B are not equal, we write $A \neq B$.*

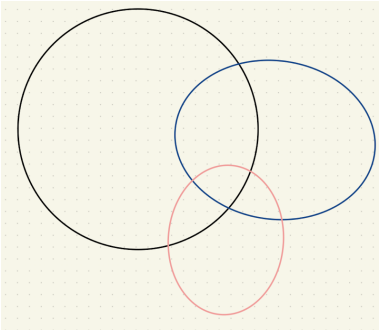
Exercise 6 (properties of containment)**.** *Let A, B, C be sets. Then the following hold.*
[reflexivity] $A \subseteq A$.
[transitivity] if $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.

Venn diagrams

Pictures are very helpful in understanding things. Venn was an English logician (lived around 19th century). He suggested a pictorial way to represent sets.

Example. *Imagine there is a universal set U that we are care about. Consider two sets $A, B \subseteq U$. Draw Venn diagram.*

Example. *The Venn diagram of three sets $A, B, C \subseteq U$:*



Remark. *Venn diagram can sometimes help to gain intuition, but they do not substitute proofs.*

Operations

Like addition and multiplication for numbers, two sets A, B naturally give rise to two more sets.

Definition 7. *The intersection of A and B is*

$$A \cap B = \{x : x \in A \wedge x \in B\}.$$

Definition 8. *The union of A and B is*

$$A \cup B = \{x : x \in A \vee x \in B\}.$$

Definition 9. *The set A minus B*

$$A \setminus B = \{x : x \in A \wedge x \notin B\}.$$

Example. *The operations applied to $A = \{1, 2, 3\}$ and $B = \{2, 3, 4\}$. Draw the Venn diagram.*

Definition 10. *If $A \subseteq B$ then the complement of A inside B is $A^c = B \setminus A$. Sometimes it is denote by $\bar{A}$.*

Remark. *The notation A^c is missing because it depends on B. In many cases, B is clear from the context. But in many other cases, we need to explicitly describe B.*

Properties

Just like addition and multiplication of numbers, these operation also satisfy some basic property. We prove one property (the rest are exercises).

1. $A \cup A = A$.
2. $A \setminus A = \emptyset$.
3. $A \cap B = B \cap A$ and $A \cup B = B \cup A$.
4. $(A \cap B) \cap C = A \cap (B \cap C)$. This allows to write $A \cap B \cap C$. A similar property holds for $\cup$.
5. $A \cap (B \cup C) = (A \cap B) \cup (A \cup C)$.

   Proof.

   $$\begin{aligned} & x \in A \cap (B \cup C) \\ & \Leftrightarrow (x \in A) \wedge (x \in B \cup C) \\ & \Leftrightarrow (x \in A) \wedge (x \in B \vee x \in C) \\ & \Leftrightarrow ((x \in A) \wedge (x \in B)) \vee ((x \in A) \wedge (x \in C)) \\ & x \in (A \cap B) \cup (A \cap C); \end{aligned}$$

   here we used the definitions and the rules of logic. □
6. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

General unions

If $A_1, \ldots, A_n$ are sets, then

$$\bigcup_{i=1}^{n} A_i = \{x : \exists i \in \{1, \ldots, n\}\ x \in A_i\}$$

and

$$\bigcap_{i=1}^{n} A_i = \{x : \forall i \in \{1, \ldots, n\}\ x \in A_i\}$$

More generally, if $\{A_i : i \in I\}$ is a set of sets whose "names" are elements of a set I, then we can take their union

$$\bigcup_{i \in I} A_i = \{x : \exists i \in I\ x \in A_i\}$$

and similarly for intersection.

De Morgan rules

Claim 11. *If $A, B \subseteq U$ then*

$$(A \cup B)^c = A^c \cap B^c$$

and

$$(A^c)^c = A.$$

Exercise 12. *Draw the Venn diagrams.*

Exercise 13. *Prove the claim.*

Claim 14.

$$A^c \cup B^c = (A \cap B)^c.$$

Proof.

$$A^c \cup B^c = ((A^c \cup B^c)^c)^c = ((A^c)^c \cap (B^c)^c)^c = (A \cap B)^c. \qquad \square$$

Power set

Definition 15. *The power set of A is the set of all subsets of A:*

$$2^A = \{S : S \subseteq A\}.$$

Remark. *This is notation; it is not like exponentiations of numbers. But if A is finite then*

$$|2^A| = 2^{|A|}.$$

We shall not prove it now.

Cartesian products

Definition 16. *If x, y are two elements, then (x, y) is the ordered pair whose first element is x and second element is y.*

Remark. *When $x \neq y$ we have*

$$(x, y) \neq (y, x).$$

And if $x = y$ then

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

Definition 17. *If A, B are sets then we have the set*

$$A \times B = \{(a, b) : a \in A, b \in B\}.$$

Exercise 18. *Draw $\{1, 3\} \times \{-1, 2, 3\}$ in the plane.*

Remark. *If A, B are finite sets then*

$$|A \times B| = |A| \cdot |B|.$$

We shall not prove right now.

Notation. *A n-tuple (a.k.a. a vector of length n, or a sequence of length n) of elements is denoted by*

$$(a_1, a_2, \ldots, a_n).$$

This is an ordered list of things.

Definition 19. *If $A_1, \ldots, A_n$ are sets then*

$$A_1 \times A_2 \times \ldots \times A_n = \{(a_1, \ldots, a_n) : \forall i \in \{1, \ldots, n\}\ a_i \in A_i\}.$$

Notation. *For a single set A, we denote by A^n the cartesian product of A with itself n times.*

Example. *$\{0, 1\}^n$ is the set of all n-bit strings. The size of $\{0, 1\}^n$ is 2^n. There is a natural identification between $2^{[n]}$ and $\{0, 1\}^n$:*

$$S \subseteq [n] \Leftrightarrow x \in \{0, 1\}^n$$

where

$$i \in S \Leftrightarrow x_i = 1.$$

Partitions

Definition 20. *A partition of a set A is a collection of pairwise disjoint sets $S_1, \ldots, S_n \subseteq A$ that cover A. That is, for all $i \neq j$ in $[n]$ we have*

$$S_i \cap S_j = \emptyset.$$

And

$$\bigcup_{i=1}^{n} S_i = A.$$

Example. *Draw a partition of an abstract set.*

3.1 Relations

It is natural to groups things to pairs. For example, all pairs of person-child, or all pairs of people who like the same movies. The mathematical way to formalize this is via a relation.

Definition 21. *For two sets A and B, a set $R \subseteq A \times B$ is called a relation.*

Remark. *Every pair $(a, b) \in R$ "satisfies" the relation, and every other pair does not.*

Notation. *We sometimes denote $(a, b) \in R$ by aRb. This is natural if R is a "verb", like "likes" in which case aRb can be read "a likes b".*

Remark. *In general, relations are not symmetric. It could be that $(a, b) \in R$ but $(b, a) \notin R$. For example, when R is "parent".*

Example. *The empty relation $R = \emptyset$.*

Example. *The identity (diagonal) relation $R = \{(a, a) : a \in A\}$.*

Example. *The greater-than relation $R = \{(i, j) \in [n] \times [n] : i > j\}$.*

Remark. *We can represent R as a zero-one matrix; e.g. greater-than.*

Remark. *We can represent R as a directed bipartite graph. We draw the elements of A on the left as points or circles, and the element of B on the right. We draw an edge from a point a to a point b if and only if $(a, b) \in R$. Draw greater-than.*

Properties

Let $R \subseteq A \times A$ be a relation.

Definition 22. *It is reflexive if $(a, a) \in R$ for all $a \in A$.*

Definition 23. *It is anti-reflexive if $(a, a) \notin R$ for all $a \in A$.*

Definition 24. *It is symmetric if $(a, b) \in R$ implies $(b, a) \in R$.*

Definition 25. *It is anti-symmetric if $(a, b) \in R$ for $a \neq b$ implies $(b, a) \notin R$.*

Definition 26. *It is transitive if $(a, b) \in R$ and $(b, c) \in R$ implies $(a, c) \in R$.*

Exercise 27. *Think of some examples.*

Exercise 28. *What can you say about $R \subseteq A \times A$ that is both symmetric and anti-symmetric?*

3.2 Equivalence relations

There is a special type of relations that is important in many areas of math.

Definition 29. *A relation $R \subseteq A \times A$ is an equivalence relation if it is reflexive, symmetric and transitive.*

Exercise 30. *The identity relation is an equivalence relation.*

Example.

$$R_3 = \{(a, b) \in \mathbb{Z}^2 : \exists\ k \in \mathbb{Z}\ a = b + 3k\}.$$

Claim 31. *R_3 is an equivalence relation.*

Proof. Reflexivity and symmetry are left as exercises. Let us look at transitivity. Assume aR_3b and bR_3c. There are $k_1, k_2 \in \mathbb{Z}$ so that

$$a = b + k_1 n = (c + k_2 n) + k_1 n = c + (k_1 + k_2)n. \qquad \square$$

Definition 32. *The equivalence class of $a \in A$ in an equivalence relation $R \subset A^2$ is*

$$[a]_R = \{a' \in A : (a, a') \in R\}.$$

This is the set of all elements that are equivalent to a.

Example.

$$[0]_{R_3} = \{\ldots, -6, -3, 0, 3, 6, \ldots\}.$$

Exercise 33. *If R is an equivalence relation and $(a, a') \in R$ then $[a]_R = [a']_R$.*

Exercise 34. *If R is an equivalence relation and $(a, a') \notin R$ then $[a]_R \cap [a']_R = \emptyset$.*

Definition 35. *If $R \subseteq A^2$ is an equivalence relation, then the collection of equivalent class of R is*

$$A/R = \{[a]_R : a \in A\}.$$

This is a set of sets. It is sometimes called "A modulo R".

Theorem 36. *If $R \subseteq A^2$ is an equivalence relation, then A/R is a partition of A.*

Proof. The partition property holds because each $a \in A$ belongs to $[a]_R$ and using Exercises 33 and 34. $\square$

Modulo

The "modulo" operation is extremely important in mathematics. Here is a simple example from algebra. The set $\mathbb{Z}/R_3$ is "the integers modulo 3":

$$\mathbb{Z}/R_3 = \{[0]_{R_3}, [1]_{R_3}, [2]_{R_3}\}.$$

You already know what is $a + b$ for numbers a, b. We can define a similar operation on equivalent classes:

$$[a]_{R_3} + [b]_{R_3} := [a + b]_{R_3}.$$

Exercise 37. *Verify that the definition makes sense. That is, if aR_3a' and bR_3b' then*

$$[a + b]_{R_3} = [a' + b']_{R_3}.$$

Remark. *We can now add sets! This type of construction is extremely important, and you will see variants of it in many courses later on.*

Remark. *This example explains the term "modulo" which some of you already know. The integers modulo n is the set $\{0, 1, \ldots, n-1\}$ with the operation of "addition modulo n".*

Remark. *The set $\mathbb{Z}$ is a* group *with addition. The equivalence class $[0]_{R_3}$ is a* normal subgroup. *The set $\mathbb{Z}/R_3$ is also a group. This is a general construction in group theory.*

Remark. *The set $\{0, 1, 2\}$, which is isomorphic to $\mathbb{Z}/R_3$, is not only an additive group, its subset $\{1, 2\}$ is also a multiplicative group.*

Exercise 38. *What is the relation on $\mathbb{Z}$ that encodes this?*

Remark. *These two groups (additive and multiplicative) over $\{0, 1, 2\}$ share additional structure which make it a* field.

The rationals

Coming back to the rationals $\mathbb{Q}$. We wrote

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

But this is not accurate. Let

$$A = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}).$$

Define an equivalence relation $\sim$ on A:

$$\tfrac{p}{q} \sim \tfrac{p'}{q'} \Leftrightarrow pq' = p'q.$$

The correct definition of $\mathbb{Q}$ is

$$\mathbb{Q} = A/\sim .$$

Exercise 39. *Verify that $\sim$ is an equivalence relation.*

Remark. *This definition of $\mathbb{Q}$ satisfies all the properties we are used to, but this requires verification.*

Postlude: is everything a set?

Sets are the most basic structures of mathematics. We saw examples, discussed operations, and defined relations. There is an axiomatic definition of "sets". We did not talk about it.

A program for basing math on sets was initiated in the late 19th century by Cantor, Zermelo, Hilbert, Russell and others. This program raised fundamental questions, that we shall not discuss at length. We provide one important example

Example (Russell's paradox (discovered earlier by Zermelo))**.** *We saw many symbols above. We can use them to write expressions (that "compile"). Consider*

$$\{x : x \notin x\}.$$

The observation is that the above can not define a set. Because if it is a set S, then

$$S \in S \Leftrightarrow S \notin S.$$

Set theory provides axioms that forbid the above expression from being a set, thus resolving the problem. You will learn more about this is the course on Set Theory.

Chapter 4

Functions

Functions are natural objects, especially today when computers are everywhere. They have inputs and outputs, and they typically correspond to something we wish to do or know.

Remark. *Formally, functions are a special kind of sets, a special kind of relations.*

Definition 40. *A function f is a relation $R \subseteq A \times B$ so that for every $a \in A$ there is a unique $b \in B$ so that aRb. This element b is denoted by $f(a)$.*

Notation. *Such a function f is denoted by $f : A \to B$.*

Definition 41. *The set A is called the domain of f, and B its range. Inside B there is a special subset, the image of f*

$$f(A) = \{f(a) : a \in A\} = \{b \in B : \exists a \in A \ f(a) = b\}.$$

Exercise 42. *Draw this.*

Remark. *For every $a \in A$, there is a unique b of the form $b = f(a)$. But there could be many $a \in A$ so that $f(a) = b$. Sometimes we write*

$$f^{-1}(b) = \{a \in A : f(a) = b\}.$$

Example. *The identity map $id : A \to A$ defined by $id(a) = a$.*

Example. *An indicator of $X \subseteq A$ is the map $1_X : A \to \{0, 1\}$ defined by*

$$1_X(a) = \begin{cases} 1 & a \in X, \\ 0 & a \notin X. \end{cases}$$

This is new notation. The $\{$ denotes "two or more cases". Each row is a case. The left part L of a row is the value. The right part R of a row is a condition. It is read: "when R holds the value is L".

Example. *An indicator of $x \in A$ is sometimes called a* delta *function.*

Notation. *The set of all functions from A to B is denoted by*

$$B^A = \{f : A \to B\}.$$

Remark. *This is consistent with B^n which is the same as $B^{[n]}$.*

Remark. *If A, B are finite then*

$$|B^A| = |B|^{|A|}.$$

Types of functions

Definition 43. *A function $f : A \to B$ is injective or one-to-one if $a_1 \neq a_2$ implies $f(a_1) \neq f(a_2)$.*

Remark. *An injective function is like an "ID number".*

Definition 44. *A function $f : A \to B$ is surjective or onto if $f(A) = B$.*

Example. *The map $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = x + 1$ is one-to-one and onto.*

Example. *The map $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = x^2$ is not one-to-one and not onto.*

Operations

Definition 45. *The composition of $g : B \to C$ and $f : A \to B$ is $g \circ f : A \to C$ defined by*

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

Example. *The composition of $x \mapsto x^2$ and $x \mapsto x + 1$ is*

$$x \mapsto (x + 1)^2.$$

Remark. *Composition is associative*

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

Remark. *Composition is not always commutative. (Find an example.)*

Definition 46. *A function $f : A \to B$ is invertible if there is $g : B \to A$ so that*

$$g \circ f = \mathsf{id}_A$$

and

$$f \circ g = \mathsf{id}_B.$$

Notation. *This g is denoted by f^{-1}.*

Remark. *Just one of the conditions above does not suffice.*

Theorem 47. *If $f : A \to B$ has an inverse, then it is unique.*

Proof. If g, h are two inverses of f. For every $b \in B$,

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

□

Theorem 48. *The function $f : A \to B$ is invertible iff it is one-to-one and onto.*

Proof. The simple proof is left as an exercise. □

Remark. *If A is a finite set then*

$$\textit{there is an invertible } f : A \to [n] \textit{ iff the size of } A \textit{ is } n.$$

This allows to define the "size" of sets for infinite sets as well. We do this next.

Summary

Functions are fundamental in all areas of math. We defined them and saw several basic and important properties.

Chapter 5

Cardinality

What is the size of a set?

Definition 49. *If A is a finite set then its size $|A|$ is the number of elements in it.*

What is the size of an infinite set? This question turns out to be quite deep.

Comparison

First, we compare sets.

Remark. *If A, B are finite sets then*

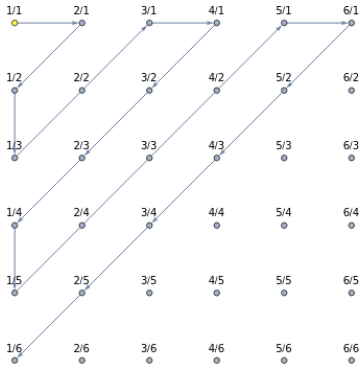
$$|A| \leq |B| \text{ iff there is a one-to-one map from } A \text{ to } B.$$

We can use this remark to *define* $\leq$ for general sets.

Definition 50. *For sets A, B, we write $|A| \leq |B|$ if there is a one-to-one map from A to B.*

Remark. *If A is a strict subset of a finite set B then it is not true that $|B| \leq |A|$.*

Example. *Although $\mathbb{N}$ is only the "diagonal" of $\mathbb{N}^2$, we have $|\mathbb{N}^2| \leq |\mathbb{N}|$. The injection $f : \mathbb{N}^2 \to \mathbb{N}$ is defined via the following drawing:*



Algebraically, it can be defined via

$$f((a,b)) = (a+b)^2 + a.$$

Indeed, for $(a,b) \neq (a',b')$*, if* $a+b = a'+b'$ *then* $a \neq a'$*, and if* $a+b > a'+b'$ *then*

$$(a+b)^2 + a > (a'+b'+1)^2 \geq (a'+b')^2 + a'.$$

Definition 51. *A set* A *is countable if* $|A| \leq |\mathbb{N}|$*.*

Remark. *A set* A *is countable if its elements can be "enumerated" as* $a_1, a_2, \ldots$

Remark. *The following are countable:*

$$\mathbb{N}, \mathbb{N}^2, \mathbb{Z}, \mathbb{Z}^2, \ldots$$

In fact, the following is countable:

$$\bigcup_{n\in\mathbb{N}} \mathbb{N}^n.$$

Question 52. *Are all sets countable?*

Theorem 53 (Cantor)**.** *The sets* $\mathbb{N}^{\mathbb{N}}$ *and* $2^{\mathbb{N}}$ *are not countable.*

Proof. This is called Cantor's diagonalization argument. Assume towards a contradiction that there is a one-to-one map from $2^{\mathbb{N}}$ to $\mathbb{N}$. This means that $2^{\mathbb{N}}$ can be ordered as $S_1, S_2, S_3, \ldots$ Draw this in a $\mathbb{N} \times \mathbb{N}$ table. Consider the following set $D \subset \mathbb{N}$. For every $n \in \mathbb{N}$,

$$n \in D \text{ iff } n \notin S_n.$$

The observation is that D does not appear in the list above. Indeed, if $D = S_k$ for some $k \in \mathbb{N}$ then

$$k \in D \Leftrightarrow k \notin S_k,$$

which means that $D \neq S_k$. This means that one of the sets in $2^{\mathbb{N}}$ is not in the list, a contradiction. □

Remark. *This powerful argument can be generalized and appears in CS as well.*

Equality of cardinalities

Remark. *The size of* A *is* n *if and only if there is a one-to-one and onto map from* A *to* $[n]$*.*

Remark. *The definition of "size" can be thought of as "all comparable sets have the same size."*

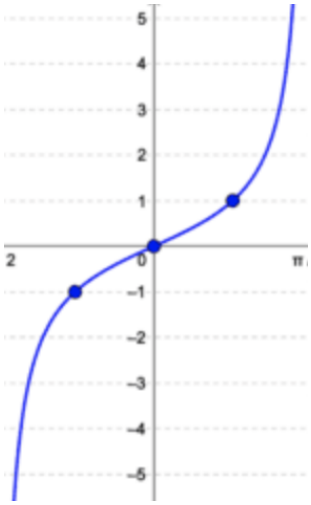
Definition 54. *Two sets A, B have the same size, denoted by $|A| = |B|$, if there is a one-to-one and onto map from A to B.*

Example. *Somewhat counterintuitively:*

$$|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = |\mathbb{Z}^2|.$$

Example. *Although $(0,1) = \{x \in \mathbb{R} : 0 < x < 1\}$ is a strict subset of $\mathbb{R}$, we have*

$$|(0,1)| = \mathbb{R}.$$



Example. *For every set A, we have $|2^A| = |\{0,1\}^A|$ via $X \mapsto 1_X$.*

Example. *$|(0,1)| = |2^{\mathbb{N}}|$ via binary representation; there are some details that we skip.*

Remark. *For every $B \in \{2, 3, \ldots\}$, we can express real numbers in base B. Each $x \in [0,1)$, can be written as*

$$x = \sum_{n=1}^{\infty} \frac{x_n}{B^n}$$

where $x_n \in \{0, 1, \ldots, B-1\}$ for all $n \geq 1$. Example: in base 3 write $1/3$ and $1/2$. We shall not prove this right now, but you know decimal representation ($B = 10$) and perhaps also binary representation ($B = 2$). This representation is not unique; for example, $1 = 0.9999999999\ldots$ in decimal.

Definition 55. *We write $|A| < |B|$ if $|A| \leq |B|$ but $|A| \neq |B|$.*

Theorem 56 (Cantor)**.** *If A is a set then $|A| < |2^A|$.*

Proof. Diagonalization; you'll see again in the Set Theory course. □

Theorem 57 (Cantor-Bernstein)**.** *If* $|A| \leq |B|$ *and* $|B| \leq |A|$ *then* $|A| = |B|$.

Remark. *The proof is not trivial; you will see in the Set Theory course. From two one-to-one maps, we need to "construct" a single out-to-one and onto map.*

Exercise 58. *Let* X *be a set. The relation* $\sim \subset 2^X \times 2^X$ *defined by*

$$A \sim B \Leftrightarrow |A| = |B|$$

is an equivalence relation.

Remark. *For a set* X*, we can think of* $2^X/\sim$ *as the "sizes" of all subsets of* X*. Every equivalence class consists of sets of the same size. This allows to define "size".*

Definition 59. *We denote by* $\aleph_0$ *the cardinality of* $\mathbb{N}$ *and by* $\aleph_1$ *the cardinality of* $2^{\mathbb{N}}$*.*

Remark. *There are some subtleties here that we ignore; what is the "big set" that a set belongs to.*

Remark. *We saw* $\aleph_0 < \aleph_1$*.*

Question 60. *Are there sets of cardinalities between* $\aleph_0$ *and* $\aleph_1$*?*

Remark. *The continuum hypothesis (CH) states that there are no such sets. This leads to an interesting theory, and "problems" in the basis of math. It was proved by Godel and Cohen that CH is independent of the standard axioms of math; it can not be proved or disproved, unless the axioms lead to a contradiction.*

Summary

We defined "size" of sets. We saw that there are infinitely many different sizes of infinite sets. More on this theory in the Set Theory course.

Chapter 6

Order

In the previous chapter, we compared sizes of sets. We now move to the more complex notion of "order". The canonical example is the order

$$1 < 2 < 3 < \dots$$

Definition 61. *A partial order R on a set A is a relation that is reflexive, anti-symmetric, and transitive. Typically, denote this R by $\leq$. We write $(A, \leq)$ to denote a partial order on A. The set A is called partially ordered.*

Remark. *The notation $\leq$ may not be the "usual one".*

Example. $\mathbb{N}$ *with standard order is partially ordered.*

Example. *Let X be a set. Order 2^X by containment: $A \leq B$ if $A \subseteq B$.*

Remark. *Hasse diagrams are useful for visually representing partial orders.*

Example. *Draw $2^{\{1,2,3\}}$ with $\subseteq$.*

Remark. *We do not need to draw* all the arrows. We can deduce them via transitivity.

Remark. *The term "partial order" captured that not every two elements are comparable. E.g. for 2^X.*

Definition 62. *A full (or linear) order is a partial order in which every two elements are comparable. That is, for every $a, b \in A$ either $a \leq b$ or $b \leq a$.*

Example. $(\mathbb{N}, \leq)$ *is linear.*

Example. $(2^X, \subseteq)$ *is not linear for $|X| > 1$.*

Example. *Lexicographic order (as in dictionary) on $\mathbb{N}^2$ defined by $(a, b) \leq (c, d)$ if $a \leq c$ or ($a = c$ and $b \leq d$). This is a linear order.*

Example. *The Cartesian order on $\mathbb{N}^2$ is defined by $(a,b) \leq (c,d)$ if $a \leq c$ and $b \leq d$. Loosely speaking, if the point (a,b) is the the rectangle defined by (c,d) and the origin. This order is not linear.*

Definition 63. *For partially ordered set $(A, \leq)$, an element $a \in B$ is minimal in B if there is no $b \in B$ so that $b < a$.*

Remark. *Minimality does not implies that $a \leq b$ for all $b \in B$.*

Example. *1 is minimal in $(\mathbb{N}, \leq)$.*

Remark. *There is no maximal element in $(\mathbb{N}, \leq)$.*

Example. *In $((2^{\{1,2,3\}}, \subseteq)$, the set $\emptyset$ is minimal and $\{1,2,3\}$ is maximal. There are three minimal elements in $B = \{S \subseteq \{1,2,3\} : |S| > 0\}$.*

Example. *$(1,1)$ is minimal in the Cartesian order.*

Induction

Induction is a powerful method for proving claims.

Example. *Let first recall the standard setting you know. There is some set P of natural numbers that we care about. For example, the set of all $n \in \mathbb{N}$ so that*

$$\sum_{i=1}^{n} = \frac{n(n+1)}{2}.$$

The set P can a priori be empty, part of $\mathbb{N}$ or all of $\mathbb{N}$. The typical goal is to prove that $P = \mathbb{N}$. We can prove this by induction. First, we show that $1 \in P$. This is "the induction base". Then, we prove that if $n \in P$ then $n+1 \in P$. This is "the induction step".

Remark. *Here is a more abstract view. There is a set of objects A. Typically, A is $\mathbb{N}$. There is a subset P of A that we care about. Our goal is to prove that $P = A$. To do this, we use an additional structure on A. There is an operation $f : A \to A$ and a finite subset A_0 of A so that each $a \in A$ can be generating using f in a finite number of step from A_0. Typically, $A_0 = \{1\}$ and $f(a) = a + 1$. That is, for $n \geq 1$ let*

$$A_n = f(A_{n-1}).$$

Then

$$A = \bigcup_{n \geq 0} A_n.$$

To prove $P = A$ we need to prove the induction base $A_0 \subset P$, and the induction step $a \in P \Rightarrow f(a) \in P$.

Remark. *A better perspective is to view A as partially ordered. Assume that $(A, \leq)$ is partially ordered. Our goal again is to prove that $P = A$. What property of this order allows to prove things by induction?*

Definition 64. *The set A satisfies the* minimality condition *if every non empty subset B of A contains a minimal element.*

Example.

- *$\mathbb{N}$ with $\leq$ satisfies minimality (the minimal element is unique).*
- *$\mathbb{N}^2$ with Cartesian order satisfies minimality (the minimal element is not always unique).*
- *$\mathbb{Z}$ with $\leq$ does not satisfy minimality.*

Remark. *Assume that we know that A satisfies the minimality condition. A proof by induction consists of the following two claims.*

1. *Every minimal element $a \in A$ is in P.*
2. *For every a, if every $b < a$ is in P then $a \in P$.*

The observation is that this too implies that $P = A$. Indeed, assume towards a contradiction that $P \neq A$. Let a_* be a minimal element of $A \setminus P$. By (1), a_* is not a minimal element of A. By the minimality of a_*, every $b < a_*$ is in P. By (2) we see that $a_* \in P$.

Chains

Not all partial orders are linear. But some "local parts" can still be linear.

Definition 65. *A set $C \subseteq A$ is a chain if all of its elements are comparable.*

Example. $\emptyset \subseteq \{1\} \subseteq \{1, 2\} \subseteq \{1, 2, 3\}$.

On the other extreme:

Definition 66. *A set $C \subseteq A$ is an antichain if no two elements in it are comparable.*

Example. $\{1\}, \{2\}, \{3\}$ *in* $2^{\{1,2,3\}}$ *ordered by inclusion.*

Well orderings

Example. $(\mathbb{N}, \leq)$ *is well ordered.*

Example. $\mathbb{N}^2$ *with Cartesian order is not well ordered.*

Question 67. *What is the difference?*

Definition 68. *A set A is* well ordered *if every non empty subset contains a unique minimal element.*

Theorem 69. *A well ordered set is linearly ordered.*

Proof. The uniqueness of the minimal element in $\{a, b\}$ allows to compare a, b. □

Remark. *Not every linear order is a well ordering. For example $\mathbb{Z}$ with $\leq$.*

Remark. *Well orders are very important in set theory.*

Dense orders

What is difference between the order on $\mathbb{Z}$ and on $\mathbb{Q}$? The order on $\mathbb{Q}$ is "dense".

Definition 70. *A partial order $\leq$ on A is dense if for every $a < b$ in A there is $c \in A$ so that $a < c < b$.*

Example. *The usual order on $\mathbb{Q}$ is dense and on $\mathbb{Z}$ is not.*

Exercise 71. *If A has a dense order then A is infinite.*

Remark. *The natural order on $\mathbb{Q}$ is dense and not bounded (there is not minimal element and no maximal element).*

Theorem 72 (Cantor)**.** *If $(A, \leq)$ is linear, dense, not bounded and countable then it is "isomorphic" to the order of $\mathbb{Q}$.*

Remark. *The term isomorphic is central in mathematics. Roughly speaking, it means "the same" or "equivalent". Sometimes we write $X \simeq Y$ to denote that X and Y are isomorphic. The meaning depends on the context!*

Definition 73. *Two partial orders $(A, \leq_A)$ and $(B, \leq_B)$ are isomorphic if there is a bijection $f : A \to B$ that respects order. That is, if $a \leq_A a'$ then $f(a) \leq_B f(a')$ and if $b \leq_B b'$ then $f^{-1}(b) \leq_A f^{-1}(b')$.*

We shall not prove Cantor's theorem, but we shall prove the following.

Theorem 74 (Universality of the rationals)**.** *If $(A, \leq_A)$ is a linear order with A countable then $(A, \leq_A)$ can be "embedded" in $\mathbb{Q}$ with the standard order $\leq$. That is, there is an injective map $f : A \to \mathbb{Q}$ so that if $a \leq_A a'$ then $f(a) \leq f(a')$.*

Proof. Enumerate $A = \{a_1, a_2, \ldots\}$. Construct the embedding f by induction on this enumeration. Define $f(a_1) = 1$, and for $n \geq 1$, define $f(a_{n+1})$ according to the position of a_{n+1} in $\{a_1, \ldots, a_n\}$. That is, let $A' = \{a_1, \ldots, a_n\}$ and let

$$A_- = \{a \in A' : a < a_{n+1}\}$$

and

$$A_+ = \{a \in A' : a_{n+1} < a\}.$$

If $A_- = \emptyset$ then set

$$f(a_{n+1}) = \min\{f(a) : a \in A'\} - 1.$$

If $A_+ = \emptyset$ then set

$$f(a_{n+1}) = \max\{f(a) : a \in A'\} + 1.$$

Otherwise, let $q \in \mathbb{Q}$ be so that

$$\max A_- < q < \min A_+$$

and set

$$f(a_{n+1}) = q.$$

It can be proved by induction that f satisfies the needed properties. The details are left as an exercise. □

Example. *Look at construction with $1 < 3 < 5 < \ldots < 2 < 4 < 6 < \ldots$ with enumeration*

$$\mathbb{N} = \{1, 2, 3, \ldots\}.$$

Chapter 7

Permutations

This chapter introduces a basic notion that captures linear orders on $[n]$, bijections from $[n]$ to $[n]$, and shuffling of n cards.

Definition 75. *A permutation on a (finite) set A is a one-to-one and onto map $\pi : A \to A$.*

Example. *The identity permutation.*

Example. *The identity permutation.*

Remark. *Every permutation π has an inverse π^{-1} and*

$$\pi\pi^{-1} = \pi^{-1}\pi = id,$$

where we removed the $\circ$ symbol.

Exercise 76. *If π, σ are permutation of A then $\pi\sigma$ is also a permutation.*

Remark. *The collection of permutation of A is thus a group. In general, it is not abelian $\sigma\pi \neq \pi\sigma$.*

Notation. *The set of permutations of $[n]$ is sometimes denoted by S_n.*

Remark. *The size of S_n is $n!$.*

Representations

Consider π on [5] defined by

$$\begin{aligned} 1 &\mapsto 2 \\ 2 &\mapsto 3 \\ 3 &\mapsto 1 \\ 4 &\mapsto 4 \\ 5 &\mapsto 4. \end{aligned}$$

It can be represented by

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 5 & 4 \end{pmatrix}.$$

It can also be represented by a union of disjoint "cycles" as

$$(123)(45) = (1,2,3)(4,5).$$

This is read from right to left, where in a cycle elements are mapped "rightward" modulo the circle. This can be achieved by looking at

$$(1, \pi(1), \pi^2(1), \ldots, \pi^k(1))(x, \pi(x), \ldots, \pi^\ell(x)) \ldots.$$

Remark. *In this "canonical" way to represent a permutation, all cycles are disjoint.*

Remark. *Every cycle by itself also represents a permutation.*

Remark. *Two disjoint cycles commute. For example,* $(123)(45) = (45)(123)$.

Remark. *The identity permutation can be written as*

$$id = (1)(2)(3)(4)(5).$$

Sometimes cycles of length one are "left out".

Transpositions

Definition 77. *A transposition is a cycle of length two:* (ij).

Example. $(12)(23)(13) = (23)$.

Remark. *Transpositions are involutions:* $(ij)^{-1} = (ij)$.

Example. $((12)(23)(34))^{-1} = (34)(23)(12)$.

Theorem 78. *Every permutation can be represented as a product of transpositions.*

Proof. First, write π as a product of cycles. Second, write a cycle as a product of transpositions via

$$(123\dots k) = (12)(23\dots k),$$

and use induction. □

Exercise 79. *The set S_n can be generated by the $n-1$ transpositions $(1,2),(2,3),\dots,(n-1,n)$.*

Exercise 80. *The set S_n can not be generated by $n-2$ transpositions.*

Theorem 81. *For every π, the parity of the number of transpositions in any representation of π is fixed. That is, if π is a product of k_1 and k_2 transpositions then k_1 and k_2 have the same parity.*

Remark. *The theorem is important and has several proofs.*

Definition 82. *The party of π is thus defined; e.g. call π even if it can be written as a product of an even number of transpositions.*

The proof of the theorem is based on the followings definition and claim:

Definition 83. *Define $N(\pi)$ as n minus the number of cycles in the "canonical" cycle decomposition of π, where we also write cycles of length one.*

Claim 84. *For every π and every transposition (i,j),*

$$N((i,j)\pi) = N(\pi) \pm 1.$$

Proof. It is enough to look at the cycles of π that contain i and j. There are at most two such cycles. When there is one such cycle, there is one additional cycle in $(i,j)\pi$:

$$(i,j)(i,a_1,\dots,a_t,j,a_{t+1},\dots,a_s) = (i,a_1,\dots,a_t)(j,a_{t+1},\dots,a_s).$$

When there are two such cycles, there is one less cycle in $(i,j)\pi$:

$$(i,j)(i,a_1,\dots,a_t)(j,a_{t+1},\dots,a_s) = (i,a_1,\dots,a_t,j,a_{t+1},\dots,a_s).$$ □

Proof of Theorem. The claim is that if π is a product of k transpositions then the parity of k is the same as the parity of $N(\pi)$. This is proved by induction. If $k=0$ then $\pi = id$ and both are zero. If π is a product of $k>0$ transpositions $\pi = t_1\cdots t_k$ then $t_1\pi$ is a

product of $k-1$ transpositions. By induction, the parity of $N(t_1\pi)$ is equal to that of $k-1$. By the claim above, $N(\pi) = N(t_1\pi) \pm 1$. The parity of $k = k-1+1$ is the same as the parity of $N(\pi)$. □

Remark. *An equivalent statement is that id can only be written as a product of an even number of transpositions. To see the (more complicated side of the) equivalence, write σ as a product of k_1, k_2 transpositions. The inverse of a product of k transpositions is the same k transpositions in reverse. Write $id = \sigma\sigma^{-1}$ and deduce that $k_1 + k_2$ is even.*

Definition 85. *The sign of π is*

$$sign(\pi) = \begin{cases} 1 & \pi \text{ is even,} \\ -1 & \pi \text{ is odd.} \end{cases}$$

Remark. *If S_n is the set of permutation on $[n]$ then*

$$sign : S_n \to \{1, -1\}$$

so that

$$sign(\pi\sigma) = sign(\pi)sign(\sigma).$$

This is a homomorphism from the group of permutations to the group of size two. Its kernel is the set of even permutations (and it is a (normal) sub-group).

Chapter 8

Sums

Remark. *Sums are basic; they allow to quantity things, measure errors, understand values, etc.*

Remark. *There are several ways to write sums:*

$$\begin{aligned} 1+2+3+4+5 &= \sum_{i=1}^{5} i \\ &= \sum_{n=1}^{5} n \\ &= \sum_{x\in\{1,2,3,4,5\}} x \\ &= \sum_{x\in X} x \end{aligned}$$

where $X = [5]$.

Example.

$$\sum_{i=1}^{n} i = \frac{1}{2}\Big(\sum_{i=1}^{n} i + \sum_{j=1}^{n} n + 1 - j\Big) = \frac{n(n+1)}{2}.$$

Example. *For* $x \neq 1$,

$$\sum_{i=0}^{n} x^i = \frac{1-x^{n+1}}{1-x}$$

because if multiply by $1-x$ *we get a telescopic sum.*

Example. *If* X *is a finite set*

$$\sum_{x\in X} 1 = |X|.$$

Example (How many fixed points does a typical permutation have?)**.** *For a permutation π on $[n]$, the number of fixed points of π is*

$$F(\pi) = |\{i \in [n] : \pi(i) = i\}|.$$

What is

$$\frac{1}{n!} \sum_{\pi \in S_n} F(\pi)?$$

Where S_n is the set of all permutations of $[n]$. For each π,

$$F(\pi) = \sum_{i \in [n]} 1_{\pi(i)=i}.$$

For each $i \in [n]$,

$$\sum_{\pi \in S_n} 1_{\pi(i)=i} = (n-1)!.$$

So,

$$\frac{1}{n!} \sum_{\pi \in S_n} F(\pi) = \frac{1}{n!} \sum_{\pi \in S_n} \sum_{i \in [n]} 1_{\pi(i)=i} = \frac{1}{n!} \sum_{i \in [n]} \sum_{\pi \in S_n} 1_{\pi(i)=i} = \frac{1}{n!} \sum_{i \in [n]} (n-1)! = 1.$$

In words, the expected number of fixed points is one.

Remark. *Changing order of summation is a powerful mechanism.*

Chapter 9

Pegionhole

Remark. *A basic goal in math: prove that X exists. Examples: equation has a solution, function has a maximum, system has an equilibrium, etc.*

> The pigeonhole principle is a mechanism for proving existence. *If $n+1$ pigeons fly to n holes then there is at least one hole with at least two pigeons.*

Or formally:

Theorem 86. *There is no injection from $[n+1]$ to $[n]$.*

This basic idea allows to prove some non trivial statements.

Example. *In a drawer with black and white socks, how many do we need to take out to get a pair with the same color?*

Example. *There are two people in Israel with the same amount of hairs on head.*

Example. *There is no way to cover a 6×6 square with two diagonal corners removed using a 2×1 dominos. (Color it like a chess board, and notice that removed corners have the same color.)*

Claim 87. *For every integer $n > 0$ there is $k > 0$ so that in the decimal representation of nk there are only ones and zeros.*

Proof. Look at $n+1$ numbers of the form $1, 11, 111, \ldots$. Find a collision modulo n, We get $X > Y$ in the list and $X = Y \mod n$, so that $X - Y = kn$ has the desired form. □

Remark. *In applications, emphasize what are the pigeons and what are the holes.*

Claim 88. *For every $A \subset [100]$ of size $|A| = 10$, there are $X, Y \subset A$ so that $X \neq Y$ and*

$$\sum_{x \in X} x = \sum_{y \in Y} y.$$

Proof. Map $X \to \sum_{x \in X} x$. There are 2^{10} different X's and only $10 \times 100 < 2^{10}$ options for sums. □

Remark. *Posteriori we can conclude that X and Y are disjoint (by removing their intersection).*

Theorem 89 (Erdos and Szekeres)**.** *In every sequence of $k = n^2 + 1$ distinct integers $i_1, \ldots, i_k$ there is a monotone sub-sequence of length at least n.*

Proof. For each $j \in [k]$, let u_j be the longest increasing subsequence starting at i_j (u for "up") and let d_j be the longest increasing subsequence starting at i_j (d for "down"). The claim is that if $j < j'$ then

$$(u_j, d_j) \neq (u_{j'}, d_{j'}).$$

Indeed, if $i_j < i_{j'}$ then $u_j > u_{j'}$ and similarly when $i_j < i_{j'}$. In other words, the map $j \mapsto (u_j, d_j)$ is injective. The image of this map is not contained in $[n] \times [n]$. □

Remark. *This is one of the first examples of Ramsey theory. This theory says something like "in every large enough system, there are small ordered regions".*

Remark. *One of the first applications of the pigeonhole principle is in the context of Diophantine approximation (i.e., approximations reals via integers). The "trivial" approximation: For every $x \in \mathbb{R}$ and $q \in \mathbb{N}$ there is $p \in \mathbb{Z}$ so that*

$$|x - \tfrac{p}{q}| \leq \tfrac{1}{q}.$$

Theorem 90 (Dirichlet)**.** *For every $x \in \mathbb{R}$ there are infinitely many $p, q \in \mathbb{Z}$ so that $q > 0$ and*

$$|x - \tfrac{p}{q}| \leq \tfrac{1}{q^2}.$$

Proof. Look at $0, x, 2x, 3x, \ldots, nx$, and consider the fractional parts of this sequence. There are $a > b$ in $\{0, 1, \ldots, n\}$ so that $|ax - k_1 - (bx - k_2)| \leq \frac{1}{n}$, because these are $n+1$ points in $[0, 1)$. For $p = k_1 - k_2$ and $q = (a - b)$,

$$|x - \tfrac{p}{q}| \leq \tfrac{1}{nq} \leq \tfrac{1}{q^2}.$$ □

Remark. *An approximation of this form (order $\frac{1}{q^2}$ accurate with denominator q) is the best possible for irrational algebraic numbers (Roth's theorem). You can try to prove that for the special case of $\sqrt{2}$; hint: look at $|\sqrt{2} - \frac{p}{q}||\sqrt{2} + \frac{p}{q}|$.*

Remark. *Injections correspond to "encoding of data". For example, $E : X \to \{0,1\}^* = \bigcup_{n=1}^{\infty} \{0,1\}^n$ where the encoding E must be an injection. If X is "large" then there must $x \in X$ whose description length $|E(x)|$ is "large".*

Chapter 10

Hall's theorem

Remark. *Every relation $R \subseteq A \times B$ can be represented by a bipartite graph. The elements of A and B are called vertices. If aRb then there is an edge between a and b denoted by $e = \{a, b\}$ (and these are all the edges). From now on, we denote R by E, and think about it as the set of edges.*

Remark. *Thinking of E as "worker a can perform job b", we would like to find a "perfect matching" from A to B, as assignment to all the workers to jobs. In other words, we would like to find a one-to-one map from A to B that is contained in R. In the graph language, we would like to find a collection of $|A|$ disjoint edges.*

Definition 91. *A* matching $M \subseteq E$ *in a bipartite graph is a collection of pairwise disjoint edges. A matching M is* perfect for A *if for every $a \in A$ there is $e \in E$ so that $a \in e$.*

Remark. *The graph perspective allows to think of relations in a "geometric" way. There are neighbors, neighbors of neighbors, etc. The perspective is useful.*

Definition 92. *The neighborhood of $U \subseteq A$ is*

$$N(U) = N_E(U) = \{b \in B : \exists a \in U \text{ so that } \{a, b\} \in E\}.$$

Remark. *The most informative type of theorem is an "iff condition"; showing that two a priori different situations situations are actually the same. We seek such a condition for a perfect matching.*

Definition 93. *A bipartite graph satisfies Hall's condition if for every $U \subseteq A$,*

$$|N(U)| \geq |U|.$$

Theorem 94. *A bipartite graph has a perfect matching for A iff it satisfies Hall's condition.*

Remark. *The "easy" direction is that if there is a perfect matching then Hall's condition holds: if M is a perfect matching for A then*

$$|N_E(U)| \geq |N_M(U)| = |U|.$$

The "hard" direction is the content of the proof.

Proof. We prove that if Hall's condition holds then there's a perfect matching. The proof is by induction on $|A|$. The induction base is $|A| = 1$, for which the theorem holds. For the inductive step, choose $a \in A$. The vertex a is connected to some neighbor $b \in B$, because $|N(\{a\})| \geq 1$. Let E' be the graph induced on $A' = A \setminus \{a\}$ and $B' = B \setminus \{b\}$. If E' satisfies Hall's condition, then there is a perfect matching for A' and B' and adding $\{a, b\}$ we are done. If E' does not satisfy Hall's condition, then there $U \subseteq A'$ so that

$$|N_{E'}(U)| \leq |U| - 1.$$

Example. *An example that yields this case:* $(1,1), (1,2), (2,1), (3,1), (2,3)$*. There is a perfect matching* $(3,1), (1,2), (2,3)$*, but after deleting* 1 *on both sides we are left with two left vertices and a single edge.*

It follows that

$$|N_E(U)| = |U|.$$

Now, let E' be the graph induced on $U \times N(U)$, and let E'' be the graph induced on $(A \setminus U) \times (B \setminus N(U))$. The graph E' satisfies Hall's condition. By induction there is a perfect matching for U in E'. It remains to verify that the graph E'' satisfies Hall's condition; for each $V \subseteq E''$, we have

$$|U| + |V| \leq |N_E(U \cup V)| \leq |N_E(U)| + |N_{E''}(V)| = |U| + |N_{E''}(V)|.$$

By induction, we also have a perfect matching for $A \setminus U$ in E''. □

Remark. *Hall's theorem has many applications. We'll see a sequence of thoughts that starts with Hall's theorem, goes back to posets and continues to linear equations.*

Remark. *There is an algorithmic aspect to this problem. Finding a perfect matching is important in applications. If $|A| = |B| = n$, then checking Hall's condition takes times exponential in n. The proof gives a recursive algorithm that also runs in time exponential in n. The Hopcroft–Karp algorithm solves it in time order $n^{2.5}$. There is a large body of research on this algorithmic question.*

Remark. *Counting the number of perfect matchings in a graph turns out to be a much harder problem. The best known algorithm for it takes exponential time (and there are good reasons to believe that there are no polynomial-time algorithms).*

Pushing sets to the middle

Remark. *This part deals with the poset $(2^{[n]}, \subseteq)$. Its Hasse diagram has $n+1$ layers. The vertices at layer $k \in \{0, 1, \ldots, n\}$ consists of all sets of size k. A set S of size k is connected to $n-k$ sets of size $k+1$, and to k sets of size $k-1$. We use Hall's theorem to start with a family of sets and "push it towards the middle without creating new containments".*

Claim 95. *For $k < \frac{n}{2}$, let*

$$A_k = \{S \subset [n] : |S| = k\}.$$

There is an injection

$$f_k : A_k \to A_{k+1}$$

so that for each $a \in A_k$ we have

$$f(a) \supset a.$$

Proof. Define a graph $E \subseteq A_k \times A_{k+1}$ where we connect $a \in A_k$ to $b \in A_{k+1}$ iff $a \subset b$. It remains to verify that Hall's condition holds; if $U \subseteq A_k$ then:

- Each $a \in A$ is connected to $n-k$ neighbors.
- Each $b \in A_{k+1}$ is connected to $k+1$ neighbors.

So,

$$\sum_{a \in U} \sum_{b \in N(U)} 1_{a \subset b} = |U|(n-k)$$

and switching order of summation:

$$\sum_{b \in N(U)} \sum_{a \in U} 1_{a \subset b} \leq |N(U)|k.$$

It follows that

$$|N(U)| \geq \frac{n-k}{k}|U| \geq |U|. \qquad \square$$

Remark. *The claim shows that we can push small sets "upward" while maintaining containment. A symmetric argument deals with large sets.*

Claim 96. *For $k > \frac{n}{2}$, There is an injection*

$$f_k : A_{k+1} \to A_{k+1}$$

so that for each $a \in A_k$ we have

$$f(a) \subset a.$$

Remark. *It follows that for every the sequence*

$$|A_0|, |A_1|, \ldots, |A_n|$$

is unimodal (it goes up and then down).

Sizes of antichains

Theorem 97 (Sperner)**.** *The maximum size of an antichain in* $(2^{[n]}, \subseteq)$ *is* $|A_m|$ *for* $m = \lceil \frac{n}{2} \rceil$.

Remark. *Each* A_k *is an antichain.*

Proof. Let $S \subset 2^{[n]}$ be an antichain. Let

$$\ell = \min\{k : S \cap A_k \neq \emptyset\}.$$

If $\ell < \frac{n}{2}$ then let

$$S_1 = (S \setminus A_\ell) \cup f_\ell(S \cap A_\ell).$$

Because S is an antichain, and by construction of f_ℓ, the size of S_1 is the same as that of S, and S_1 is also an antichain. We can similarly keep pushing S_1 towards the middle; if $\ell > \frac{n}{2}$ we push the sets "downward". Eventually, we get a sequence of sets $S_1, S_2, \ldots, S_t$ so that $S_t \subseteq A_m$ and $|S_t| = |S|$. □

Remark. *Sperner's theorem has many applications. Here's one in probability theory or number theory.*

Anti-concentration

Theorem 98 (Littlewood-Offord, Erdos)**.** *Let* $a_1, \ldots, a_n$ *positive integers and let* $b \in \mathbb{Z}$. *The number of solutions* $x_1, \ldots, x_n \in \{0, 1\}$ *to the equation*

$$\sum_{i \in [n]} a_i x_i = b$$

is at most $|A_m|$.

Remark. *The poset* $(2^{[n]}, \subseteq)$ *is isomorphic to the poset* $(\{0,1\}^n, \leq)$, *where*

$$x = (x_1, \ldots, x_n) \leq x' = (x'_1, \ldots, x'_n)$$

if $x_i \geq x'_i$ *for all* $i \in [n]$.

Proof. The set of solutions forms an antichain; if $x < x'$ then $\sum_i a_i x_i < \sum_i a_i x'_i$ so they can't both be solutions. □

Remark. *It follows that if $x \in \{0,1\}^n$ is chosen uniformly at random then the chance that $\sum_i a_i x_i = b$ is at most $\frac{1}{\sqrt{n}}$. Most vectors are not solutions for such an equation.*

Remark. *Starting from Hall's theorem we moved through several area ideas that are related to extremal set theory, combinatorics, and algebra.*

Vertex covers and duality

Remark. *Think of a bipartite graph in a different way. The vertices are computers and the edges are communication channels. The goal now is to find the minimum number of computers that "control" all channels.*

Definition 99. *A vertex cover in a bipartite graph $G = (A, B, E)$ is a set of vertices $C \subseteq A \cup B$ so that for each $e \in E$ we have $e \cap C \neq \emptyset$.*

Remark. *Our goal is to find the minimum vertex cover in a given graph.*

Remark. *Matching provide "obvious" barriers for vertex covers; if G has matching of size m then the minimum size of a vertex cover is m (need at least one vertex from each edge in matching). König's theorem shows that this is the "only barrier".*

Theorem 100. *In any bipartite graph, the maximum size of a matching is equal to the minimum size of a vertex cover.*

Remark. *Draw an example.*

Proof. Let C be a minimum vertex cover in $G = (A, B, E)$. We want to find a matching M of size $|M| \geq |C|$. Let G' be the induced graph on $A' = C \cap A$ and $B' = B \setminus C$. The minimality of C implies that G' satisfies Hall's condition; if $W \subset A'$ is so that $|N_{G'}(W)| < |W|$ then we can replace W be $N_{G'}(W)$ and get a smaller vertex cover. So there is a perfect matching for A' in G'. A similar argument shows that there is a perfect matching in G'' that is induced on $A'' = A \setminus C$ and $B'' = B \cap C$. The union of the two matchings is also a matching and it has the correct size. □

Remark. *The theorem is an example of linear programming (LP) duality: the maximum of problem X is equal to the minimum of the dual problem X^*. There is a lot of theory here, which involves geometry, convexity, etc. There are game theoretic interpretations as well (von Neumann minimax theorem).*

Posets: structure

Remark. *We describe an application for finding structure in general posets.*

Definition 101. *Let $(X, \leq)$ be a poset. A chain decomposition of X is a collection $X_1, \ldots, X_n$ of chains so that $\bigcup_{i \in [n]} X_i = X$. The size of the chain decomposition is n.*

Remark. *We can assume that the pieces of a chain decomposition are pairwise disjoint.*

Remark. *The goal is to find the smallest possible chain decomposition. An "obvious" barrier is an antichain; if $A \subseteq X$ is an antichain then the size of any chain decomposition is at least $|A|$. Dilworth's theorem shows that this is the "only" barrier.*

Theorem 102. *The minimum size of a chain decomposition of $(X, \leq)$ is equal to the maximum size of an antichain in X.*

Example. *Verify the theorem for the poset $(2^{[n]}, \subseteq)$.*

Proof. Define the bipartite graph $G = (A, B, E)$ by $A = B = X$ and $\{x, y\} \in E$ iff $x < y$. We can assume that there are no isolated vertices in G, because these correspond to elements that are incomparable to anything else. By König's theorem there is a matching M in G and a vertex cover C in G of the same size. The vertex cover C must contain at least one vertex in each edges in M. Let A be the set of $x \in X$ that do not "belong" to C.

Claim 103. *The set A is an antichain of size at least $|X| - |C|$.*

Proof. Each element in X has two copies, one in A and one in B. The size of A is therefore at least $|X| - |C|$. Because C is a vertex cover, the set A is an antichain. □

Define a collection of chains by putting x, y in the same chain if $\{x, y\} \in M$. Add to this collection of chain all singletons that remain as chains of size one. Denote this decomposition of chains by $X_1, \ldots, X_n$.

Claim 104. *The above decomposition to chains have size $n \leq |X| - |M| = |X| - |C|$.*

Proof. Because $X_1, \ldots, X_n$ partition X,

$$|X| = \sum_{i=1}^{n} |X_i|,$$

Because M is a matching,

$$\sum_{i=1}^{n} |X_i| - 1 = |M|.$$

So,

$$|X| = \sum_{i=1}^{n} |X_i| = \Big(\sum_{i=1}^{n} |X_i| - 1 \Big) + n = |M| + n. \qquad \square$$

$\square$

Remark. *All three theorems, Hall, König and Dilworth, are equivalent (proved by Fulkerson).*