Introduction to Set Theory

Nordine Mir
Texas A&M University at Qatar

QF Math Circle
September 2025

So, what's a set?

If we put a family of (mathematical) objects together, we obtain a *set*.
You know a couple of sets: The set of natural numbers, the set of rational numbers, the set of real numbers, and so on.
If A is a set, for every given object x we need to be able to determine whether x belongs to A or not. We also say x is (not) an *element* or *member* of A. In notation:

$$x \in A, \qquad x \notin A$$

We use two different ways to describe sets:

Descriptive

$$A = \{x \colon x \text{ satisfies some property } \}$$

Enumerative

$$A = \{a_1, a_2, \dots\}$$

An example

The set of all integers is denoted by $\mathbb{Z}$. So:

$$\mathbb{Z} = \{m \colon m \text{ is an integer}\} = \{\dots, -2, -1, 0, 1, 2, \dots\}$$

And now you!

We say that an integer $m \in \mathbb{Z}$ is *divisible* by $k \in \mathbb{Z}$ if there exists an $\ell \in \mathbb{Z}$ such that $m = k\ell$. In that case, we write $k \mid \ell$.
Give a descriptive and an enumerative description of the set of all integers divisible by 3 (utilizing the symbol $\mid$ just introduced when you think it's useful).

A barber and his own beard

In a small village in France, there is a (male) barber. The barber shaves all of the men in the the village who don't shave themselves.

And now you!

Discuss with your neighbour to find an answer to the following question: Does the barber shave himself?

The Barber Paradox

If we are trying to form the set consisting of all sets which do not contain themselves,

$$X = \{A \colon A \notin A\},$$

can you answer the question whether X is an element of X?

"Solving" the barber paradox

We thus need rules which govern how we can form sets to avoid logical inconsistencies. Important theories are the following.

The Zermelo-Fränkel Axioms

The ZF axiom system is probably the the most widely used in all of mathematics. It gives rules which govern the formation of sets. You might hear about the *axiom of choice* later on.

Naive Set Theory

We take the approach that we are not going to deal with these issues of mathematical logic and assume that our sets are all contained in one large set (the "universe").

The empty set

A special set is the set which contains *no* elements. This set is denoted by

$$\emptyset = \{\}\,.$$

So for every a, we have $a \notin \emptyset$. Now, the set

$$\{\emptyset\}$$

is *not* the empty set, since $\emptyset \in \{\emptyset\}$.

Operations with sets

The union of two sets A and B is the set

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

The intersection of of the sets A and B is the set

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

Example

Given $A = \{-100, 2, e, \pi\}$, $B = \{\text{apple}, \text{Hakim}, \pi\}$ find $A \cup B$ and $A \cap B$.

Many rules such as $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.....

Operations with sets

The relative complement of B in A is

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

The cartesian product of A and B is

$$A \times B = \{(x, y) : x \in A, \text{ and } y \in B\},$$

where (x, y) denotes an ordered pair (not an interval!).

Example

Given $A = \{-100, 2, e, \pi\}$, $B = \{\text{apple}, \text{Hakim}, \pi\}$, find $A \setminus B$, $B \setminus A$, $A \times B$ and $B \times A$.

Subsets

A set A is a *subset* of a set X, written as $A \subset X$ (or $A \subseteq X$), if every element of A is also an element of X: $A \subset X \iff (a \in A \Rightarrow a \in X)$.

Example

Let $A = \{1, 2, \{1, 2\}, \{1, 3\}, 4\}$. Determine whether each of the following is an element of A, a subset of A, both an element and subset of A, or neither an element of A nor a subset of A.

1. 1
2. 3
3. $\{1\}$
4. $\{1,2\}$
5. $\{1,3\}$
6. $\{1,4\}$
7. $\{\{1,3\}\}$

The power set

Given a set X, the set of all subsets of X is called the power set of X and is denoted $\mathcal{P}(X)$.

And now you!

Find *all* of the subsets of the set $\{a, b\}$. How many are there?

The power set

And now you!

Find *all* of the subsets of the set $\{a, b, c\}$. List them in the following table:

# Elements	List of subsets with that many elements	# Subsets
0		
1		
2		
3		

Find *all* of the subsets of the set $\{a, b, c, d\}$.

# Elements	List of subsets with that many elements	# Subsets
0		
1		
2		
3		
4		

And now you!

Question

We already know that the empty set has exactly one subset (itself), a one-element set has two subsets, and a two-element set has four subsets. How many did we get for a set with three and four elements?

Conjecture

A set with $k \in \mathbb{N}$ elements has 2^k subsets.

Any ideas for a proof?

Proof by induction

Let N_k be the number of subsets of a set with k elements. We know that $N_0 = 1$, $N_1 = 2$, $N_2 = 3$. Let us assume that for any fixed $k \in \mathbb{N}$ we have $N_k = 2^k$ (Induction Hypothesis).
Given a set X with $k+1$ elements, we fix one element $\star \in X$. For any subset $A \subset X$, either $\star \in A$ or $\star \notin A$. If we therefore split the subsets of X up into those which contain $\star$ and those which do not, we obtain a splitting of the subsets of X into two groups, each of which can be identified with the subsets of a set with k elements (the subsets if X with $\star$ removed, and then either leave them alone, or add $\star$ in).
So: $N_{k+1} = N_k + N_k = 2N_k = 2 \cdot 2^k = 2^{k+1}$. Since this is true for any fixed $k \in \mathbb{N}$, induction shows that $N_k = 2^k$ for all $k \in \mathbb{N}$.

Q.E.D.

Functions

If X and Y are sets, a *function* $f\colon X \to Y$ from X to Y is a rule associating to any $x \in X$ an element $f(x) \in Y$ (the value of x under f).
The set of all functions from X to Y is written as Y^X.
A simple example of a function is the *identity function* $\mathrm{id}_X\colon X \to X$ defined by $\mathrm{id}_X(x) = x$ for every $x \in X$.
Even though you are probably going to be used to functions given by *formulas*, e.g. $f(x) = x^2$, this does not make sense if X is not a set of numbers.
Alternatively, we give functions on finite sets as lists of pairs of points and their values.

Example

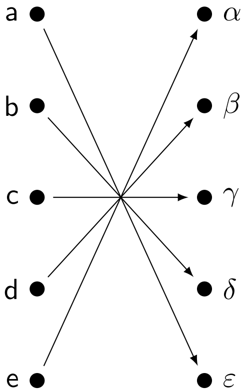
We can define a function $f\colon \{a, b, c, d, e\} \to \{\alpha, \beta, \gamma, \delta, \varepsilon\}$ by

x	a	b	c	d	e
f(x)	ε	δ	γ	β	α

Arrow Diagrams for functions

If we want to relate sets X and Y to each other, we use functions $f: X \to Y$ to do so.
For finite sets, functions can be represented by *arrow diagrams*, and for general sets, those are a good way to visualize the properties we want to talk about now. Here's our example:

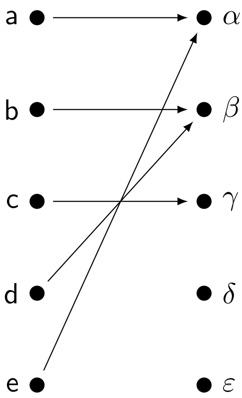
x	a	b	c	d	e
f(x)	ε	δ	γ	β	α



And now you!

Draw the arrow diagram for the function given in the table.

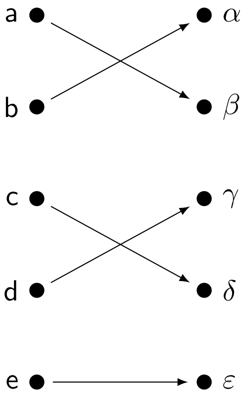
x	a	b	c	d	e
f(x)	α	β	γ	β	α



And now you!

Fill the table for the function described by the arrow diagram.

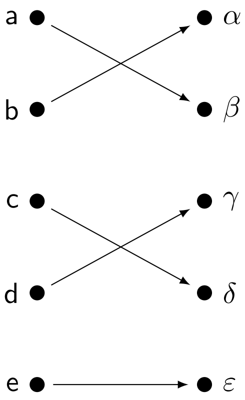
x	a	b	c	d	e
f(x)					



And now you!

Fill the table for the function described by the arrow diagram.

x	a	b	c	d	e
f(x)	β	α	γ	δ	ε



Injectivity

Definition

We say that a function $f\colon X \to Y$ is *injective* if the value of a given x uniquely determines it, that is if

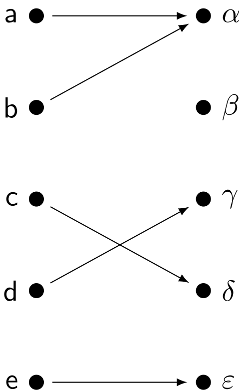
$$f(x_1) = f(x_2) \Rightarrow x_1 = x_2.$$

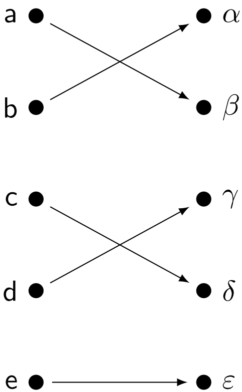
Example

$f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, is *not injective*, but $g\colon \mathbb{R}_+ \to \mathbb{R}$, $g(x) = x^2$, is. Strictly increasing functions are injective.

What does this definition mean in terms of arrow diagrams?

Pick the injective diagram!





The left diagram describes a non-injective function, because $f(a) = f(b)$.

Surjectivity

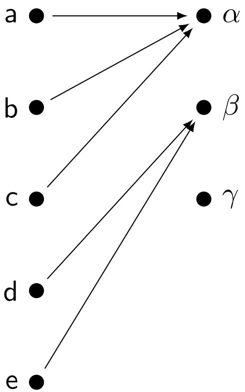
Definition

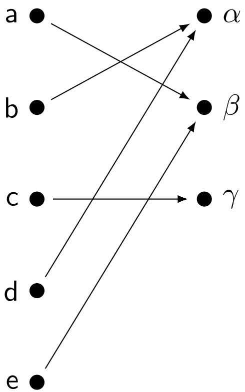
We say that a function $f\colon X \to Y$ is *surjective* if every $y \in Y$ is the value of some $x \in X$, that is if

$$\text{for every } y \in Y \text{ there exists an } x \in X \text{ such that } f(x) = y.$$

Surjectivity depends *heavily* on Y. If we replace Y by $f(X) = \{f(x)\colon x \in X\}$, then the function $\tilde{f}\colon X \to f(X)$, $\tilde{f}(x) = f(x)$, is surjective.

Pick the surjective diagram!





The left diagram describes a non-surjective function, because there is no x such that $f(x) = \gamma$.

Bijectivity and inverse functions

Definition

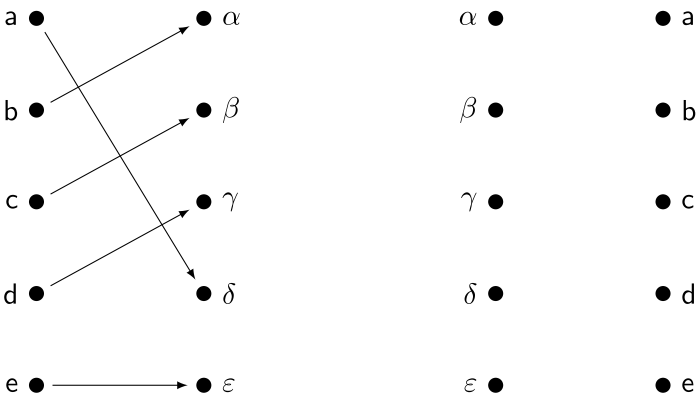
We say that a function $f\colon X \to Y$ is *bijective* if it is both injective and surjective. We say that a function $f\colon X \to Y$ is invertible if there exists a function $g\colon Y \to X$ which "undoes" f, i.e. if $g(f(x)) = x$ for every $x \in X$ and $f(g(y)) = y$ for every $y \in Y$.

Theorem

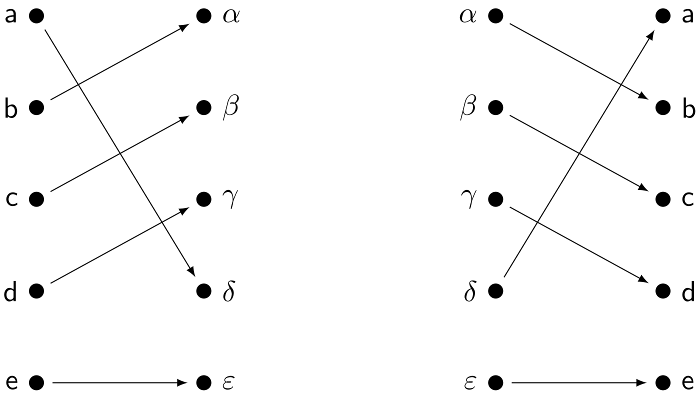
f is bijective if and only if it is invertible.

We skip a formal proof, but you can see this from the arrow diagram observation!

Arrow diagrams of the inverse function



Arrow diagrams of the inverse function



Some useful general statements

X and Y are two arbitrary sets (finite or not).

Proposition

Given two sets X and Y, if there exists an injective (surjective) function $f\colon X \to Y$, then there exists a surjective (injective) function $g\colon Y \to X$.

Theorem (Cantor-Bernstein)

If there is an injective function $f\colon X \to Y$ and a surjective function $h\colon X \to Y$ then there exists a bijective function $g\colon X \to Y$.

Cardinality of finite sets

The cardinality $|X|$ of a finite set X is the number of its elements.

Example

For $X = \{a, b, c, d\}$, we have $|X| = 4$.

Lemma

If X and Y are finite sets, then:

- *If there exists an injective function $f: X \to Y$, then $|X| \leq |Y|$.*
- *If there exists a surjective function $f: X \to Y$, then $|X| \geq |Y|$.*
- *If there exists a bijective function $f: X \to Y$, then $|X| = |Y|$.*

What is ...

Example

For finite sets X and Y, recall that the set of all function $f : X \to Y$ is written as Y^X. Can you determine the cardinality Y^X of this set?

Hint: How many choices do you have for $f(x)$ for each element $x \in X$?

Theorem

The cardinality of Y^X is given by

$$\left| Y^X \right| = |Y|^{|X|}.$$

Characteristic Functions

If X is a set, and $A \subseteq$ is a subset, the characteristic function $\chi_A \colon X \to \{0, 1\}$ is defined by

$$\chi_A(x) = \begin{cases} 0 & x \notin A \\ 1 & x \in A. \end{cases}$$

More interesting example

Here is an example, writing out the values of χ_A for some subsets of $\{a, b, c\}$.

A	a	b	c
$\emptyset$	0	0	0
$\{a\}$	1	0	0
$\{b, c\}$	0	1	1

And now you!

Write out the complete table of values of the characteristic function χ_A for *every* subset $A \subseteq \{a, b, c\}$.

Back to the power set

Definition

Given a set X, we define the power set $\mathcal{P}(X)$ of X to be the set consisting of all subsets of X.

Theorem

Given any set X (finite or not!), the mapping $F\colon \mathcal{P}(X) \to \{0,1\}^X$ defined by $F(A) = \chi_A$, is bijective.

Proof.

We have to show F is injective. If $F(A) = F(B)$, then $\chi_A(x) = \chi_B(x)$ for every $x \in X$, so $x \in A \Leftrightarrow x \in B$, and therefore $A = B$.
On the other hand, if $f\colon X \to \{0,1\}$ is any function, then $f = \chi_A$ for the set $A = \{x\colon f(x) = 1\}$. □

Cardinal Numbers–infinite sets

We are now considering sets that are not necessarily finite.

Definition

Let X and Y be sets. We then say that

- $|X| = |Y|$ if there exists a bijective function $f \colon X \to Y$;
- $|X| \leq |Y|$ if there exists an injective function $f \colon X \to Y$
- $|X| \geq |Y|$ if there exists a surjective function $f \colon X \to Y$.

This definition now works for *any* sets X and Y.

Remark

A set X is infinite if and only if there exists a proper subset $A \subsetneq X$ such that $|A| = |X|$.

Countable sets

A set X is called countable if $|X| = |\mathbb{N}|$ i.e. if there exists a bijection $X \to \mathbb{N}$.
A set is countable if its elements can be *enumerated* in some fashion, that is, we can list them $X = \{x_1, x_2, \dots\}$ where x_j is just the value of j under the bijective function $f : \mathbb{N} \to X$ that we know has to exist.
Sets X with a larger cardinality (i.e satisfying $|X| \geq |\mathbb{N}|$ and $|X| \neq |\mathbb{N}|$) are called *uncountable*.

We are going to see examples of countable and uncountable sets.

Of course the first one is $\mathbb{N}$!

Examples of countable sets

And now you!

Show that $\mathbb{Z}$ is countable i.e. find an injective function $\mathbb{Z} \to \mathbb{N}$.

$0 \to 0$; $1 \to 1$; $-1 \to 2$; $2 \to 3$; $-2 \to 4$....summarized as

$$f(n) = \begin{cases} -2n \text{ if } n \leq 0 \\ 2n - 1 \text{ if } n \geq 1 \end{cases}$$

f is bijective!

And now you!

Show that $\mathbb{Z} \times \mathbb{Z}$ is countable by finding for example a surjective function $\mathbb{N} \to \mathbb{Z}^2$ or by finding an injective function $\mathbb{N}^2 \to \mathbb{N}$.

Examples of countable sets

And now you!

Deduce from the above that $\mathbb{Q}$ is countable.

Proposition

The set of all real numbers satisfying polynomial equations with integer coefficients is countable.(These are called algebraic numbers.)

$\sqrt{2}$, $\sqrt[3]{7}$....

Remark

The set of all real numbers that are not algebraic consists of the so-called transcendental numbers: π, e.... We will see later that such a set is not countable!

Cantor's trick

The first infinite sets we encounter are countable sets. What about power sets of countable sets–are they still countable?
In other words, is $|\mathcal{P}(\mathbb{N})| = |\mathbb{N}|$?
Recall that we can identify $\mathcal{P}(\mathbb{N})$ with functions on $\mathbb{N}$ valued in $\{0, 1\}$.
Now imagine that you enumerate these functions, say $f_1, f_2, \ldots$.

	1	2	3	4	5	6	...
f_1	0	0	0	0	0	0	...
f_2	0	1	0	0	0	0	...
f_3	0	1	0	1	0	0	...
$\vdots$							

Can you find a characteristic function of some subset $A \subseteq \mathbb{N}$ that does *not* appear in the list?

Cantor's trick

Yes, that's right! Define

$$g(j) = \begin{cases} 0 & f_j(j) = 1 \\ 1 & f_j(j) = 0 \end{cases}$$

then $g \neq f_j$ for every j.

Hence $|\mathcal{P}(\mathbb{N})| > |\mathbb{N}|$.

$\mathbb{R}$ is uncountable

We are going to show that $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})| = |\{0,1\}^{\mathbb{N}}|$ (assuming that $|\mathbb{R}| = |[0,1]|$).

Lemma

The function $f\colon \{0,1\}^{\mathbb{N}} \to [0,1]$ *given by* $(a_j)_{j\in\mathbb{N}} \mapsto \frac{a_0}{10} + \frac{a_1}{100} + \ldots\ldots = \sum_{j=0}^{\infty} \frac{a_j}{10^{j+1}}$ *is injective.*

Lemma (Binary expansion)

The function $g\colon \{0,1\}^{\mathbb{N}} \to [0,1]$ *given by* $(a_j)_{j\in\mathbb{N}} \mapsto \frac{a_0}{2} + \frac{a_1}{2^2} + \ldots\ldots = \sum_{j=0}^{\infty} \frac{a_j}{2^{j+1}}$ *is surjective.*

Using Cantor-Bernstein, we get $|[0,1]| = |\mathcal{P}(\mathbb{N})|$.

The continuum hypothesis

More generally, one can show that given any set X, one always has $|X| < |\mathcal{P}(X)|$. This provides a way to produce infinitely many infinite sets with all different cardinality.

One can also ask whether there is a set X such that

$$|\mathbb{N}| < |X| < |\mathcal{P}(\mathbb{N})|.$$

Many people have tried to find such a set (starting with Cantor himself), but it was still one of Hilbert's problems to determine whether such a set X exists or not.
One says that the *continuum hypothesis* holds if there is no such set X.
It turns out that the continuum hypothesis is *independent* of the Zermelo-Fränkel Axiom system (Proof in the 1960s). That means that there are versions of set theory in which the continuum hypothesis holds and others in which it doesn't hold.