University of Toronto – MAT246H1-S – LEC0201/9201 – Winter 2021

Concepts in Abstract Mathematics

7 - Cardinality

Jean-Baptiste Campesato

Let's start with a short story.

A conference about singularity theory is going to take place in the lovely village of Tarski, and participants start to arrive. Most of them decided to be hosted at the Aleph Nought Hotel. It is a huge hotel, built especially for this occasion, with infinitely many rooms numbered using $\mathbb{N}$: 0, 1, 2, 3... Despite this large number of rooms, the sign *FULL* lights up over the front door indicating that there is no vacancy!

A group of 42 late mathematicians from Nice show up at the front desk and are received by the receptionist David H. who exclaims *"For Cantor's sake! I thought that we were not expecting new guests! No worries, I will find a solution"*. Then he uses the intercom of the hotel to send the following message to all the current guests: *"Sorry for the inconvenience, but I would need your cooperation in order to accomodate new guests. Please, if your current room is labeled n then could you move to the room $n+42$? Thank you so much and once again, sorry of the inconvenience"*. Therefore the rooms 0, 1, 2, ..., 41 are now available for the latecomers and the already hosted guests still have individual rooms.

Later a bus containing the canadian delegation reaches the hotel. The driver meets David H. and says *"Sorry for the delay, I got lost on the way. I have a bus full with infinitely many canadian mathematicians! We booked infinitely rooms at your hotel and for your conveniency we gave to each of our member a card with a natural number:* 0, 1, 2*..."*. Then David H. desperates *"Holy Dedekind! What a night! No worries, I will handle the situation!"*.

And once again, he uses the intercom of the hotel to send the following message: *"If you are currently in the room n, please could you move to the room $2n+1$? Sorry for the inconvenience."* This way the guests already in the hotel still have individual rooms and now there are infinitely many empty rooms (the even numbered ones) for the canadian participants. Next David H. asks the newcomers *"If your card shows the number m, please go to the room $2m$"*, and everyone gets an individual room[1].

Our friendly receptionist is later awakened by a terrible loudly noise outside. He shouts *"In Gödel's name, what's happening now?"* and then he reaches the front door to see infinitely many flying saucers[2] (numbered 0, 1, 2, ...). An extraterrestrial mathematician goes to meet David H. and tells *"Sorry for the delay, we come from Proxima Centauri and we got stuck in traffic jams. Each of our ships contains infinitely countably many participants!"*.

David H. doesn't seem particularly concerned and send the following message to the current guests using his intercom: *"Dear guest, if you are currently in the room n, please could you move to the room $2n$?"*. Therefore the odd-numbered rooms are now free. Then David H. asks the k^{th} passenger of the l^{th} ship to go to the room $3^k 5^l$, so that everyone gets an individual room[3].

This story highlights something interesting about the behaviour of infinite sets such as $\mathbb{N}$. First we were able to add 42 elements to $\mathbb{N}$ without changing its *size*. Even less intuitively, then we added a copy of $\mathbb{N}$ to $\mathbb{N}$ without changing its *size*. And finally, we were even able to add $\mathbb{N}\times\mathbb{N}$ (i.e. infinitely many copies of $\mathbb{N}$) to $\mathbb{N}$ without changing its *size*.

The goal of this chapter is to formally define the notion of *size* of a set (it will be called *cardinality*) and to study its properties (which may be counter-intuitive, as above, for infinite sets).

---

[1] The function $\mathbb{N}\ni n\mapsto 2n+1\in\mathbb{N}$ is one-to-one so the current guests keep individual rooms. Then $\mathbb{N}\ni m\mapsto 2m\in\mathbb{N}$ is also one-to-one, so two different newcomers are sent to two different rooms (and these rooms are empty since even numbered).

[2] Until now, the story was quite realistic...

[3] By uniqueness of the prime factorization, $\mathbb{N}\ni(k,l)\mapsto 3^k5^l\in\mathbb{N}$ is injective, so the newcomers are sent to different rooms, and these rooms are free since odd-numbered.

Contents

1 Reviews about functions

Definition 1 (Informal[4] definition of a function)**.** A *function* (or *map*) is the data of two sets A and B together with a "process" which assigns to each $x \in A$ a unique $f(x) \in B$:

$$f : \left\{ \begin{array}{ccc} A & \to & B \\ x & \mapsto & f(x) \end{array} \right.$$

Here, f is the name of the function, A is the *domain* of f, and B is the *codomain* of f.

Remark 2. This process can be:

- A formula: define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = e^{x^2-\pi} + 42$.
- An exhaustive list: define $f : \{1, 2, 3\} \to \mathbb{R}$ by $f(1) = \pi$, $f(2) = \sqrt{2}$, $f(3) = e$.
- A property characterizing f uniquely: log is the unique antiderivative of $g : (0, +\infty) \to \mathbb{R}$ defined by $g(x) = \frac{1}{x}$ such that $\log(1) = 0$.
- By induction: we define the sequence $u_n : \mathbb{N} \to \mathbb{R}$ by $u_0 = 1$ and $\forall n \in \mathbb{N}$, $u_{n+1} = u_n^2 + 1$.
- The solution of a differential equation: the exponential function $\exp : \mathbb{R} \to \mathbb{R}$ is the unique differentiable function such that $\exp' = \exp$ and $\exp(0) = 1$.
- The solution of a functional equation: the exponential function with base $a \in \mathbb{R}$ denoted by $\exp_a : \mathbb{R} \to \mathbb{R}$ is the unique monotonic function such that $\exp_a(x + y) = \exp_a(x)_a \exp(y)$ and $\exp_a(1) = a$.
- ...

Remark 3. The domain and codomain are part of the definition of a function. For instance:

- $f : \left\{ \begin{array}{ccc} \mathbb{R} & \to & (0, +\infty) \\ x & \mapsto & e^x \end{array} \right.$ and $g : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & e^x \end{array} \right.$ are not the same function (the first one is surjective but not the second one, see below).
- $f : \left\{ \begin{array}{ccc} [0, +\infty) & \to & \mathbb{R} \\ x & \mapsto & x^2 + 1 \end{array} \right.$ and $g : \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^2 + 1 \end{array} \right.$ are not the same function (the first one is injective but not the second one, see below).

A function is not simply a "formula", you need to specify the domain and the codomain.

Definitions 4. Given a function $f : A \to B$.

- The *image of* $E \subset A$ *by* f is $f(E) \coloneqq \{f(x) : x \in E\} \subset B$.
- The *image of* f (or *range of* f) is $\mathrm{Range}(f) \coloneqq f(A)$.
- The *preimage of* $F \subset B$ *by* f is $f^{-1}(F) \coloneqq \{x \in A : f(x) \in F\}$.

[4] Formally, a function $f : A \to B$ is characterized by its graph $\Gamma_f \subset A \times B$ which needs to satisfy $\forall x \in A$, $\exists! y \in B$, $(x, y) \in \Gamma_f$.

- The *graph of* f is the set $\Gamma_f := \{(x,y) \in A \times B \ : \ y = f(x)\}$.
- We say that f is *injective* (or *one-to-one*) if $\forall x_1, x_2 \in A,\ x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$
  or equivalently by taking the contrapositive $\forall x_1, x_2 \in A,\ f(x_1) = f(x_2) \implies x_1 = x_2$
- We say that f is *surjective* (or *onto*) if $\forall y \in B,\ \exists x \in A,\ y = f(x)$
- We say that f is *bijective* if it is injective and surjective, i.e. $\forall y \in B,\ \exists! x \in A,\ y = f(x)$

Proposition 5. *Let $f : E \to F$ and $g : F \to G$ be two functions.*

1. *If f and g are injective then so is $g \circ f$.*
2. *If f and g are surjective then so is $g \circ f$.*
3. *If $g \circ f$ is injective then f is injective too.*
4. *If $g \circ f$ is surjective then g is surjective too.*

Proof.

1. Let $x, y \in E$ be such that $g(f(x)) = g(f(y))$. Then $f(x) = f(y)$ since g is injective. Thus $x = y$ since f is injective.
2. Let $z \in G$. Since g is surjective, it exists $y \in F$ such that $z = g(y)$. Since f is surjective, it exists $x \in E$ such that $y = f(x)$. Therefore $z = g(f(x))$.
3. Let $x, y \in E$ such that $f(x) = f(y)$. Then $g(f(x)) = g(f(y))$ and thus $x = y$ since $g \circ f$ is injective.
4. Let $z \in G$. Since $g \circ f$ is surjective, there exists $x \in E$ such that $z = g(f(x))$. Then $y = f(x) \in F$ satisfies $g(y) = z$.

■

Proposition 6. *$f : A \to B$ is bijective if and only if there exists $g : B \to A$ such that* $\begin{cases} \forall x \in A,\ g(f(x)) = x \\ \forall y \in B,\ f(g(y)) = y \end{cases}$.
Then g is unique, it is called the inverse of f *and denoted by $f^{-1} : B \to A$.*

Proof. $\Rightarrow$ Assume that f is bijective, then $\forall y \in B,\ \exists! x_y \in A,\ f(x_y) = y$. We define $g : B \to A$ by $g(y) = x_y$. Then g satisfies the required properties.
$\Leftarrow$ Assume that there exists g as in the statement. Then $g \circ f = id_A$ is injective, so f is too by Proposition 5. And $f \circ g = id_B$ is surjective, thus f is too, still by Proposition 5. Therefore f is bijective.

For the uniqueness: assume there exist two such functions $g_1, g_2 : B \to A$. Let $y \in B$. Then $f(g_1(y)) = y = f(g_2(y))$. So $g_1(y) = g_2(y)$ since f is injective. ■

2 Finite sets

Definition 7. We say that a set E is finite if there exists $n \in \mathbb{N}$ and a bijection $f : \{k \in \mathbb{N} \ : \ k < n\} \to E$. Then we write $|E| = n$.

Remark 8. Note that $\{k \in \mathbb{N} \ : \ k < n\} = \{0, 1, 2, \dots, n-1\}$.

We are first going to prove that if such a n exists, then it is unique.

Lemma 9. *Let $n, p \in \mathbb{N}$. If there exists an injective function $f : \{k \in \mathbb{N} \ : \ k < n\} \to \{k \in \mathbb{N} \ : \ k < p\}$ then $n \leq p$.*

Proof. We prove the statement by induction on n.

- *Base case at $n = 0$:* for any $p \in \mathbb{N}$ we have $n \leq p$.
- *Induction step.* Assume that the statement holds for some $n \in \mathbb{N}$.
  Let $p \in \mathbb{N}$. Assume that there exists an injective function $f : \{k \in \mathbb{N} \ : \ k < n+1\} \to \{k \in \mathbb{N} \ : \ k < p\}$.
  Define $g : \{k \in \mathbb{N} \ : \ k < n\} \to \{k \in \mathbb{N} \ : \ k < p-1\}$ as follows:

$$g(x) = \begin{cases} f(x) & \text{if } f(x) < f(n) \\ f(x) - 1 & \text{if } f(x) > f(n) \end{cases}$$

  Note that $f(x) \neq f(n)$ since f is injective.

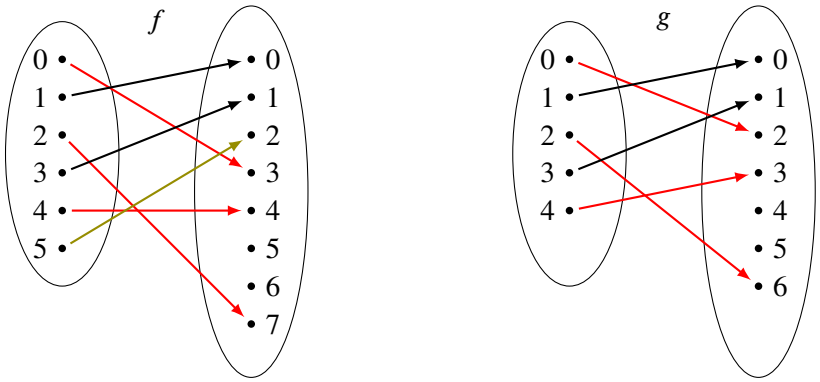


FIGURE: an example.

- Claim 1: g is well-defined, i.e. $\forall x \in \{k \in \mathbb{N} \,:\, k < n\},\ g(x) \in \{k \in \mathbb{N} \,:\, k < p-1\}$.
  Let $x \in \{k \in \mathbb{N} \,:\, k < n\}$.
  So either $f(x) < f(n)$, and then $g(x) = f(x) < f(n) < p$, therefore $0 \leq g(x) < p-1$.
  Or $f(x) > f(n)$, and then $g(x) = f(x) - 1 < p-1$, therefore $0 \leq g(x) < p-1$.
- Claim 2: g is injective.
  Let $x, y \in \{k \in \mathbb{N} \,:\, k < n\}$ be such that $g(x) = g(y)$.
  - First case: $f(x), f(y) < f(n)$.
    Then $g(x) = f(x)$ and $g(y) = f(y)$. So $f(x) = f(y)$ and thus $x = y$ since f is injective.
  - Second case: $f(x), f(y) \geq f(n)$.
    Then $g(x) = f(x) - 1$ and $g(y) = f(y) - 1$. So $f(x) = f(y)$ and thus $x = y$ since f is injective.
  - Third case: $f(x) < f(n)$ and $f(y) > f(n)$.
    Then $g(x) = f(x) < f(n)$ and $g(y) = f(y) - 1 > f(n) - 1 \geq f(n)$. Therefore, this case is impossible.
  - Fourth case: $f(y) < f(n)$ and $f(x) > f(n)$. Similar to the previous one.

Therefore, by the induction hypothesis, $n \leq p-1$, i.e. $n+1 \leq p$.

■

Corollary 10. *Let E be a finite set. If $|E| = n$ and $|E| = m$, then $m = n$.*
Then we say that $|E|$ is the cardinal *of E, which is uniquely defined.*

Proof. Assume there exists a bijection $f_1 : \{k \in \mathbb{N} \,:\, k < n\} \to E$ and a bijection $f_2 : \{k \in \mathbb{N} \,:\, k < m\} \to E$.
Then $f_2^{-1} \circ f_1 : \{k \in \mathbb{N} \,:\, k < n\} \to \{k \in \mathbb{N} \,:\, k < m\}$ is a bijection, so by the above lemma, $n \leq m$.
Similarly, $f_1^{-1} \circ f_2 : \{k \in \mathbb{N} \,:\, k < m\} \to \{k \in \mathbb{N} \,:\, k < n\}$ is a bijection and thus $m \leq n$.
Therefore $n = m$. ■

Remark 11. Informally, the cardinal of a finite set is its size, i.e. the number of elements it contains.

Remark 12. $|E| = 0 \Leftrightarrow E = \varnothing$
Indeed, if $E = \varnothing$ then $f : \{k \in \mathbb{N} \,:\, k < 0\} \to E$ is always bijective: injectiveness and surjectiveness are vacuously true. So $|E| = 0$.
Otherwise, if $E \neq \varnothing$ then $f : \{k \in \mathbb{N} \,:\, k < 0\} \to E$ is never surjective, so $|E| \neq 0$.

Proposition 13. *If $E \subset F$ and F is finite then E is finite too, besides, $|E| \leq |F|$.*

Proof. Let's prove by induction on $n = |F|$ that if $E \subset F$ then E is finite and $|E| \leq n$.

- *Base case at $n = 0$:* then $F = \varnothing$, so the only possible subset is $E = \varnothing$ and then $|E| = 0$.
- *Induction step.* Assume that the statement holds for some $n \in \mathbb{N}$.
  Let F be a set such that $|F| = n+1$.
  - *First case: $E = F$.* Then the statement is obvious.

– *Second case:* $E \neq F$. Then there exists $x \in F \setminus E$.
There exists a bijection $f : \{k \in \mathbb{N} : k < n+1\} \to F$.
Since f is bijective, there exists a unique $m \in \{0, 1, \dots, n\}$ such that $f(m) = x$.
Define $g : \{k \in \mathbb{N} : k < n\} \to F \setminus \{x\}$ by $g(k) = f(k)$ for $k \neq m$ and, if $m \neq n$, $g(m) = f(n)$.

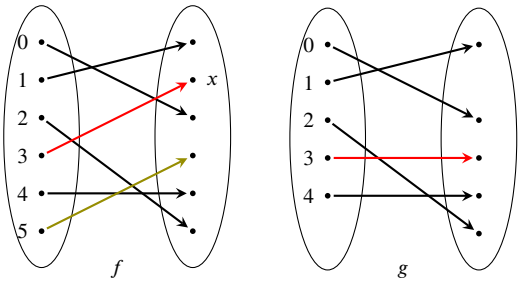


Figure 1: If $m \neq n$, i.e. $f(n) \neq x$

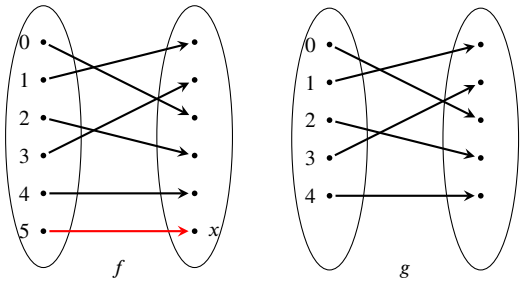


Figure 2: If $m = n$, i.e. $f(n) = x$

Then g is a bijection (*check it*), so $F \setminus \{x\}$ is finite and $|F \setminus \{x\}| = n$.
Since $E \subset F \setminus \{x\}$, by the induction hypothesis, E is finite and $|E| \leq n < n+1$. ■

Proposition 14. *Let $E \subset F$ with F finite. Then $|F| = |E| + |F \setminus E|$.*

Proof. Since $F \setminus E \subset F$ and $E \subset F$, we know that E and $F \setminus E$ are finite. Denote $r = |E|$ and $s = |F \setminus E|$.
There exist bijections $f : \{k \in \mathbb{N} : k < r\} \to E$ and $g : \{k \in \mathbb{N} : k < s\} \to F \setminus E$.
Define $h : \{k \in \mathbb{N} : k < r+s\} \to F$ by $h(k) = \begin{cases} f(k) & \text{if } k < r \\ g(k-r) & \text{if } k \geq r \end{cases}$.

- h is well-defined:
  Indeed, if $0 \leq k < r$ then $f(k)$ is well-defined and $f(k) \in E \subset F$.
  If $r \leq k < r+s$ then $0 \leq k - r < s$ so that $g(k-r)$ is well-defined and $g(k-r) \in F \setminus E \subset F$.
- h is a bijection:
  – h is injective: let $x, y \in \{0, 1, \dots, r+s-1\}$ be such that $h(x) = h(y)$.
  Either $h(x) = h(y) \in E$ and then $f(x) = h(x) = h(y) = f(y)$ thus $x = y$ since f is injective.
  Or $h(x) = h(y) \in F \setminus E$ and then $g(x-r) = h(x) = h(y) = g(y-r)$ thus $x - r = y - r$ since g is injective, hence $x = y$.
  – h is surjective: let $y \in F$.
  Either $y \in E$, and then there exists $x \in \{0, 1, \dots, r-1\}$ such that $f(x) = y$, since f is surjective.
  Then $h(x) = f(x) = y$.
  Or $y \in F \setminus E$, and then there exists $x \in \{0, 1, \dots, s-1\}$ such that $g(x) = y$ since g is surjective.
  Then $h(x+r) = g(x) = y$.

Therefore $|F| = r + s = |E| + |F \setminus E|$. ■

Proposition 15. *Let E and F be two finite sets. Then*
1. $|E \cup F| = |E| + |F| - |E \cap F|$
2. $|E \times F| = |E| \times |F|$

Proof.

1. Using Proposition 14 twice, we get
$$|E \cup F| = |E \sqcup (F \setminus (E \cap F))| = |E| + |F \setminus (E \cap F)| = |E| + |F| - |E \cap F|$$

2. We prove this proposition by induction on $n = |F| \in \mathbb{N}$.

- *Base case at $n = 0$:* then $F = \varnothing$ so $E \times F = \varnothing$ too and $|E \times F| = 0 = |E| \times 0 = |E| \times |F|$.

- *Case $n = 1$:* we will use this special case later in the proof.
  Assume that $F = \{*\}$ and that $|E| = p$. Then there exists a bijection $f : \{k \in \mathbb{N} : k < p\} \to E$.
  Note that $g : \{k \in \mathbb{N} : k < p\} \to E \times F$ defined by $g(k) = (f(k), *)$ is a bijection.
  Therefore $|E \times F| = p = p \times 1 = |E| \times |F|$.
- *Induction step.* Assume that the statement holds for some $n \in \mathbb{N}$.
  Let F be a set such that $|F| = n + 1$.
  Since $|F| > 0$, there exists $x \in F$ and $|F \setminus \{x\}| = |F| - |\{x\}| = n + 1 - 1 = n$. Then

$$\begin{aligned}
|E \times F| &= |(E \times (F \setminus \{x\})) \sqcup (E \times \{x\})| \\
&= |E \times (F \setminus \{x\})| + |E \times \{x\}| \\
&= |E| \times |F \setminus \{x\}| + |E| \text{ using the induction hypothesis and the case } n = 1 \\
&= |E| \times (|F| - 1) + |E| \\
&= |E| \times |F|
\end{aligned}$$

■

Proposition 16. *Assume that $E \subset F$ with F finite. Then $E = F \Leftrightarrow |E| = |F|$.*

Proof.
$\Rightarrow$ It is obvious.
$\Leftarrow$ Assume that $|E| = |F|$. Then $|F \setminus E| = |F| - |E| = 0$. Thus $F \setminus E = \emptyset$, i.e. $E = F$. ■

Proposition 17. *Let E a finite set. Then F is finite and $|E| = |F|$ if and only if there exists a bijection $f : E \to F$.*

Proof.
$\Rightarrow$ Assume that F is finite and that $|E| = |F| = n$.
Then there exist bijections $\varphi : \{k \in \mathbb{N} : k < n\} \to E$ and $\psi : \{k \in \mathbb{N} : k < n\} \to F$.
Therefore $f = \psi \circ \varphi^{-1} : E \to F$ is a bijection.
$\Leftarrow$ Assume that there exists a bijection $f : E \to F$.
Since E is finite there exists a bijection $\varphi : \{k \in \mathbb{N} : k < |E|\} \to E$.
Thus $f \circ \varphi : \{k \in \mathbb{N} : k < |E|\} \to F$ is a bijection. Therefore F is finite and $|F| = |E|$. ■

Proposition 18. *Let E, F be two finite sets such that $|E| = |F|$. Let $f : E \to F$. Then TFAE:*
1. *f is injective,*
2. *f is surjective,*
3. *f is bijective.*

Proof.
Assume that f is injective.
There exists a bijection $\varphi : \{k \in \mathbb{N} : k < |E|\} \to E$.
Then $f \circ \varphi : \{k \in \mathbb{N} : k < |E|\} \to f(E)$ is a bijection. Thus $|f(E)| = |E| = |F|$.
Since $f(E) \subset F$ and $|f(E)| = |F|$, we get $f(E) = F$, i.e. f is surjective.

Assume that f is surjective.
Then for every $y \in F$, $f^{-1}(y) \subset E$ is finite and non-empty, i.e. $\left|f^{-1}(y)\right| \geq 1$.
Assume by contradiction that there exists $y \in F$ such that $|f^{-1}(y)| > 1$.
Thus $|E| = \left|\bigsqcup_{y \in F} f^{-1}(y)\right| = \sum_{y \in F} \left|f^{-1}(y)\right| > |F| = |E|$.
Hence a contradiction. ■

Proposition 19. *Let E and F be two finite sets. Then $|E| \leq |F|$ if and only if there exists an injection $f : E \to F$.*

Proof.
$\Rightarrow$ Assume that $|E| \leq |F|$.
There exist bijections $\varphi : \{k \in \mathbb{N} : k < |E|\} \to E$ and $\psi : \{k \in \mathbb{N} : k < |F|\} \to F$.
Since $|E| \leq |F|$, $f = \psi \circ \varphi^{-1} : E \to F$ is well-defined and injective.

$\Rightarrow$ Assume that there exists an injection $f : E \to F$.
Then f induces a bijection $f : E \to f(E)$, so that $|E| = |f(E)|$.
And since $f(E) \subset F$, we have $|f(E)| \leq |F|$. ■

Corollary 20 (The pigeonhole principle or Dirichlet's drawer principle)**.**
Let E and F be two finite sets. If $|E| > |F|$ then there is no injective function $E \to F$.

Example 21. There are two non-bald people in Toronto with the exact same number of hairs on their heads.

Example 22. During a post-covid party with $n > 1$ participants, we may always find two people who shook hands to the same number of people.

Remark 23. Since the cardinal of a finite set is a natural number, we deduce that given two finite sets E and F, exactly one of the followings occurs:
- either $|E| < |F|$ (i.e. there is an injection $E \to F$ but no bijection $E \to F$),
- or $|E| = |F|$ (i.e. there is a bijection $E \to F$),
- or $|E| > |F|$ (i.e. there is an injection $F \to E$ but no bijection $E \to F$).

3 Generalization to infinite sets

Definition 24. We say that a set is *infinite* if it is not finite.

Theorem 25. *$\mathbb{N}$ is infinite.*

Proof. Assume by contradiction that $\mathbb{N}$ is finite. Then $\mathbb{N} \setminus \{0\} \subset \mathbb{N}$ so $\mathbb{N} \setminus \{0\}$ is finite too.
We define $f : \mathbb{N} \to \mathbb{N} \setminus \{0\}$ by $f(n) = n + 1$. Note that f is bijective with inverse $f^{-1} : \mathbb{N} \setminus \{0\} \to \mathbb{N}$ defined by $f^{-1}(n) = n - 1$.
Thus $|\mathbb{N}| = |\mathbb{N} \setminus \{0\}| = |\mathbb{N}| - |\{0\}| = |\mathbb{N}| - 1$, i.e. $0 = 1$.
Hence a contradiction. ■

So, we need to find a way in order to generalize the notion of cardinal from finite sets to all sets.

Definition 26. We say that two sets E and F have same *cardinality*, denoted by $|E| = |F|$, if there exists a bijection $f : E \to F$.
We also say that E and F are *equinumerous* or *equipotent*.

Proposition 27.
1. *If E is a set then $|E| = |E|$.*
2. *Given two sets E and F, if $|E| = |F|$ then $|F| = |E|$.*
3. *Given three sets E, F and G, if $|E| = |F|$ and $|F| = |G|$ then $|E| = |G|$.*

Proof.
1. $id : E \to E$ is a bijection.
2. Assume that $|E| = |F|$, i.e. that there exists a bijection $f : E \to F$.
   Then $f^{-1} : F \to E$ is a bijection, so $|F| = |E|$.
3. Assume that $|E| = |F|$ and $|F| = |G|$, i.e. that there exist bijections $f : E \to F$ and $g : F \to G$.
   Then $g \circ f : E \to G$ is a bijection, thus $|E| = |G|$.

■

Remark 28. At first glance, it seems that equipotence is an equivalence relation since it satisfies reflexivity, symmetry and transitivity. Nonetheless, recall that an equivalence relation is a binary relation on a set.
If equipotence were an equivalence relation, then it would be a binary relation on the set of all sets, which doesn't exist (See Theorem 54).

Theorem 29. *A set E is infinite if and only if for every $n \in \mathbb{N}$ there exists $S \subset E$ such that $|S| = n$.*

Proof.
⇒ Assume that E is infinite.
We are going to prove by induction that for every $n \in \mathbb{N}$ there exists $S \in \mathcal{P}(E)$ such that $|S| = n$.

- *Base case at $n = 0$:* $\varnothing \subset E$ satisfies $|\varnothing| = 0$.
- *Induction step.* Assume that for some $n \in \mathbb{N}$ there exists $T \subset E$ such that $|T| = n$.
  Note that $E \setminus T \neq \varnothing$ (otherwise $E = T$, which is impossible since E is infinite).
  Therefore there exists $x \in E \setminus T$. Define $S \coloneqq T \sqcup \{x\}$, then $S \subset E$ is finite and $|S| = |T| + 1 = n + 1$.
  Which ends the induction step.

⇐ Let E be a set such that for every $n \in \mathbb{N}$ there exists $S \subset E$ such that $|S| = n$.
Assume by contradiction that E is finite. Then there exists $k \in \mathbb{N}$ such that $|E| = k$.
Since $k + 1 \in \mathbb{N}$, there exists $S \subset E$ such that $|S| = k + 1$.
Since $S \subset E$, we get $k + 1 = |S| \leq |E| = k$. Hence a contradiction. ■

Corollary 30. *A set E is infinite if and only if for all $n \in \mathbb{N}$ there exists an injective function $\{k \in \mathbb{N} \,:\, k < n\} \to E$.*

Definition 31. Given two sets E and F, we write $|E| \leq |F|$ if there exists an injective function $f : E \to F$.

Theorem 32 (Cantor–Schröder–Bernstein theorem)**.**
Given two sets E and F, if $|E| \leq |F|$ and $|F| \leq |E|$ then $|E| = |F|$.

Remark 33. The above theorem states that if there exist injections $E \to F$ and $F \to E$ then there exists a bijection $E \to F$. It is less trivial than it seems at first glance when you look at the above statement.

It was first stated in 1887 by Cantor who didn't provide a proof. The first known proof is due to Dedekind on the same year, but he did not publish his proof (which was only found after he passed away). In 1895 Cantor published a proof relying on the trichotomoy principle (see Theorem 56), but Tarski later proved that the latter is actually equivalent to the axiom of choice.

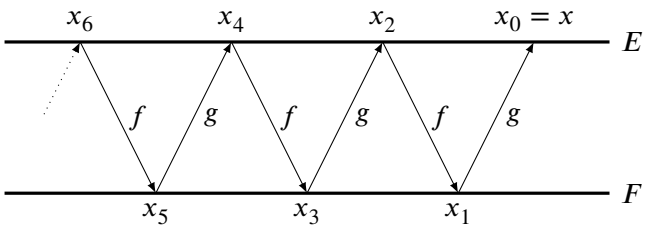
Around 1897, Bernstein, Schröder and Dedekind independently found proofs of the theorem (another one for Dedekind). But Schröder's proof later appeared to be incorrect. Several mathematicians subsequently gave alternative proofs, including Zermelo (1901, 1908) and König (1906).

Cantor–Schröber–Bernstein theorem is a little bit tricky to prove, so first I would like to informally explain the strategy of the proof before actually proving it.

We are given two injective functions $f : E \to F$ and $g : F \to E$.

Let's fix $x \in E$. We construct a *chain* $x_0, x_1, x_2, \ldots$ of elements which are alternatively in E and F as follows. First we set $x_0 = x \in E$ and then we define the next terms inductively by

- if $x_n \in E$ then we define $x_{n+1} \in F$ as the unique antecedant of x_n by g (if it exists, otherwise we stop the construction at x_n),
- if $x_n \in F$ then we define $x_{n+1} \in E$ as the unique antecedant of x_n by f (if it exists, otherwise we stop the construction at x_n).

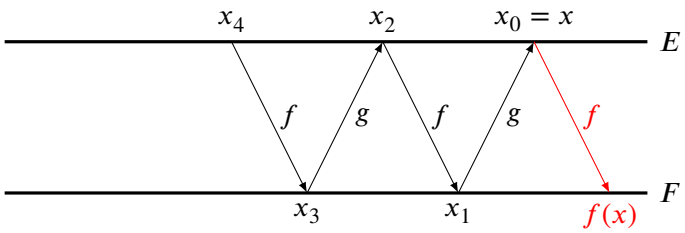


Then we face three possible cases:

1. Either the chain ends with an element in E, and then we put x in E_E,
2. or the chain ends with an element in F, and then we put x in E_F,
3. or the inductive definition of the chain doesn't stop, and then we put x in E_∞.

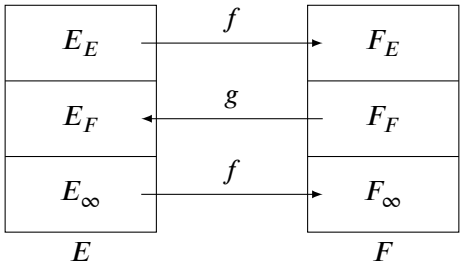
We have a partition $E = E_E \sqcup E_F \sqcup E_\infty$. We perform the same construction with $x_0 = x \in F$ in order to obtain $F = F_E \sqcup F_F \sqcup F_\infty$.

Now assume that $x \in E_E$, for instance the chain stops at x_4 as below. Then $f(x) \in F_E$, since its chain continues at x_0 and stops at $x_4 \in E$.



Therefore the function $f_{|E_E} : E_E \to F_E$ is well-defined. Besides, it is injective since f is, and it is surjective by definition of F_E. Therefore $f_{|E_E} : E_E \to F_E$ is a bijection.

Similarly $g_{|F_F} : F_F \to E_F$ and $f_{|E_\infty} : E_\infty \to F_\infty$ are bijections. Finally, we glue them in order to obtain a bijection $h : E \to F$.



Proof of Cantor–Schröder–Bernstein theorem.
Let $f : E \to F$ and $g : F \to E$ be two injective functions. Set

- $E_E = \{x \in E \ : \ \exists n \in \mathbb{N}, \exists r \in E \setminus \mathrm{Im}(g), \ x = (g \circ f)^n(r)\}$
- $E_F = \{x \in E \ : \ \exists n \in \mathbb{N}, \exists s \in F \setminus \mathrm{Im}(f), \ x = g\left((f \circ g)^n(s)\right)\}$
- $E_\infty = E \setminus \left(E_E \sqcup E_F\right)$
- $F_E = \{y \in F \ : \ \exists n \in \mathbb{N}, \exists r \in E \setminus \mathrm{Im}(g), \ y = f\left((g \circ f)^n(r)\right)\}$
- $F_F = \{y \in F \ : \ \exists n \in \mathbb{N}, \exists s \in F \setminus \mathrm{Im}(f), \ y = (f \circ g)^n(s)\}$
- $F_\infty = F \setminus \left(F_E \sqcup F_F\right)$

Note that if $x \in E_E$ then $f(x) \in F_E$. So $f_{|E_E} : E_E \to F_E$ is well-defined. It is injective since f is injective. And it is surjective by definition of the sets. Thus it is bijective.
Similarly, $g_{|F_F} : F_F \to E_F$ is well-defined and bijective and $f_{|E_\infty} : E_\infty \to F_\infty$ is well-defined and bijective.
We define $h : E \to F$ by $h(x) = \begin{cases} f(x) & \text{if } x \in E_E \\ g^{-1}(x) & \text{if } x \in E_F \\ f(x) & \text{if } x \in E_\infty \end{cases}$.
Then h is clearly a bijection (*I can use "clear", but you can't, and the same holds for "trivial" and "obvious" :-p*). ■

Proposition 34.
1. *If E is a set then $|E| \leq |E|$.*
2. *Given two sets E and F, if $|E| \leq |F|$ and $|F| \leq |E|$ then $|E| = |F|$.*
3. *Given three sets E, F and G, if $|E| \leq |F|$ and $|F| \leq |G|$ then $|E| \leq |G|$.*

Proof.
1. $id : E \to E$ is an injective function.
2. It is Cantor–Bernstein–Schröder theorem.
3. Assume that $|E| \leq |F|$ and $|F| \leq |G|$, i.e. that there exist injections $f : E \to F$ and $g : F \to G$. Then $g \circ f : E \to G$ is injective, thus $|E| \leq |G|$.

■

Remark 35. Comparison of cardinals shares the characteristic properties of an order. Nonetheless, it is not an order since it is not a binary relation on a set (as for equipotence).

Proposition 36. *If $E \subset F$ then $|E| \leq |F|$.*

Proof. Indeed, $f : E \to F$ defined by $f(x) = x$ is injective. ■

Proposition 37. *If $|E_1| = |E_2|$ and $|F_1| = |F_2|$ then $|E_1 \times F_1| = |E_2 \times F_2|$.*

Proof. Assume that $|E_1| = |E_2|$ and $|F_1| = |F_2|$ then there exist bijections $f : E_1 \to E_2$ and $g : F_1 \to F_2$.
We define $h : E_1 \times F_1 \to E_2 \times F_2$ by $h(x, y) = (f(x), g(y))$. Let's check that h is a bijection.

- *h is injective.*
  Let $(x, y), (x', y') \in E_1 \times F_1$ be such that $h(x, y) = h(x', y')$.
  Then $f(x) = f(x')$ and $g(y) = g(y')$, thus $x = x'$ and $y = y'$ since f and g are injectives.
  We proved that $(x, y) = (x', y')$.
- *h is surjective.*
  Let $(z, w) \in E_2 \times F_2$. Since f is surjective, there exists $x \in E_1$ such that $z = f(x)$.
  Since g is surjective, there exists $y \in F_1$ such that $w = g(y)$.
  Then $h(x, y) = (f(x), g(y)) = (z, w)$. ■

Theorem 38. *Given two sets E and F, $|E| \leq |F|$ if and only if there exists a surjective function $g : F \to E$.*

Proof.
⇒ Assume that there exists an injective function $f : E \to F$, then $\tilde{f} : E \to f(E)$ is bijective.
If $E = \varnothing$, then there is nothing to prove. So we may assume that there exists $u \in E$.
Define $g : F \to E$ by $g(y) = \begin{cases} \tilde{f}^{-1}(y) & \text{if } y \in f(E) \\ u & \text{otherwise} \end{cases}$
Let $x \in E$, then $g(f(x)) = \tilde{f}^{-1}(f(x)) = x$. Thus g is surjective.

⇐ Assume that there exists a surjective function $g : F \to E$, then[5] $\forall x \in E$, $\exists y_x \in g^{-1}(x)$.
Define $f : E \to F$ by $f(x) = y_x$. Then f is injective, so $|E| \leq |F|$.
Indeed, assume that $f(x) = f(x')$ then $g(f(x)) = g(f(x'))$.
But $g(f(x)) = g(y_x) = x$ and similarly $g(f(x')) = x'$. Thus $x = x'$. ■

Theorem 39. *Given two sets E and F, if $|E| = |F|$ then $|\mathcal{P}(E)| = |\mathcal{P}(F)|$.*

Proof. Let E and F be such that $|E| = |F|$. Then there exists a bijection $f : E \to F$.
Note that $\tilde{f} : \mathcal{P}(E) \to \mathcal{P}(F)$ defined by $\tilde{f}(A) = f(A)$ is bijective too (*prove it!*).
Therefore $|\mathcal{P}(E)| = |\mathcal{P}(F)|$. ■

4 Countable sets

In what follows, we set $\aleph_0 \coloneqq |\mathbb{N}|$ (pronounced *aleph nought*).

Definition 40. A set E is countable if either E is finite or $|E| = \aleph_0$.

Proposition 41. *If $S \subset \mathbb{N}$ is infinite then $|S| = \aleph_0$.*

Proof. Let's define the function $f : \mathbb{N} \to S$ by induction as follows.
Set $f(0) = \min S$ (which is well-defined by the well-ordering principle since $S \neq \varnothing$ as it is infinite).
And then, assuming that $f(n)$ is already defined, we set $f(n+1) = \min\{k \in S : k > f(n)\}$ (which is well-defined by the well-ordering principle: the involved set is non-empty since otherwise S would be finite).
It is easy to check that f is injective (note that $\forall n \in \mathbb{N}$, $f(n+1) > f(n)$), therefore $\aleph_0 \leq |S|$.
But since $S \subset \mathbb{N}$, we also have $|S| \leq \aleph_0$.
Thus, by Cantor–Schröder–Bernstein theorem, $|S| = \aleph_0$. ■

[5] (AC) See Remark 57.

Proposition 42. *A set E is countable if and only if $|E| \leq \aleph_0$ (i.e. there exists an injection $f : E \to \mathbb{N}$), otherwise stated E is countable if and only if there exists a bijection between E and a subset of $\mathbb{N}$.*

Proof.
$\Rightarrow$ Assume that E is countable.
- Either E is finite and then there exists $n \in \mathbb{N}$ together with a bijection $g : \{k \in \mathbb{N} \,:\, k < n\} \to E$.
  We define $f : E \to \mathbb{N}$ by $f(x) = g^{-1}(x)$ (which is well-defined since $\{k \in \mathbb{N} \,:\, k < n\} \subset \mathbb{N}$).
  And f is an injection since g^{-1} is.
- Or $|E| = \aleph_0$, i.e. there exists a bijection $f : E \to \mathbb{N}$.

$\Leftarrow$ Assume there exists an injection $f : E \to \mathbb{N}$.
Assume that E is infinite. Then $|E| = |f(E)| = \aleph_0$ by Proposition 41.
Thus either E is finite or $|E| = \aleph_0$. In both cases E is countable. ■

Proposition 43. *The set of finite subsets of $\mathbb{N}$ is countably infinite, i.e. $|\{S \in \mathcal{P}(\mathbb{N}) \,:\, \exists n \in \mathbb{N},\, |S| = n\}| = \aleph_0$.*

Proof. Define $f : \{S \in \mathcal{P}(\mathbb{N}) \,:\, \exists n \in \mathbb{N},\, |S| = n\} \to \mathbb{N}$ by $f(S) = \sum_{k \in S} 2^k$.
Then f is bijective by existence and uniqueness of the binary positional numeral system. ■

Proposition 44. $|\mathbb{N} \times \mathbb{N}| = \aleph_0$

Proof. Define $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $f(a, b) = 2^a 3^b$. Then f is injective by uniqueness of the prime decomposition.
Thus $|\mathbb{N} \times \mathbb{N}| \leq \aleph_0$.
Besides $\{0\} \times \mathbb{N} \subset \mathbb{N} \times \mathbb{N}$, thus $\aleph_0 = |\{0\} \times \mathbb{N}| \leq |\mathbb{N} \times \mathbb{N}|$.
Hence $|\mathbb{N} \times \mathbb{N}| = \aleph_0$ by Cantor–Schröder–Bernstein theorem. ■

Theorem 45. *A countable union of countable sets is countable,*
i.e. if I is countable and if for every $i \in I$, E_i is countable then $\bigcup_{i \in I} E_i$ is countable.

Proof. WLOG we may now assume that $I \subset \mathbb{N}$.
Let $i \in I$. Since E_i is countable, there exists an injection $f_i : E_i \to \mathbb{N}$[6].
We define $\varphi : \bigcup_{i \in I} E_i \to \mathbb{N} \times \mathbb{N}$ by $\varphi(x) = (n, f_n(x))$ where $n = \min\{i \in I \,:\, x \in E_i\}$ (*well-ordering principle*).
It is not difficult to check that φ is injective ***TODO:*** *do it?* .
Therefore $\bigcup_{i \in I} E_i$ is countable. ■

Theorem 46. *If E is an infinite set then there exists $T \subset E$ such that $|T| = \aleph_0$, i.e. $\aleph_0$ is the least infinite cardinal.*

Proof. For $n \in \mathbb{N}$, set $E_n = \{S \in \mathcal{P}(E) \,:\, |S| = n\}$. By Theorem 29, $E_n \neq \varnothing$.
So for every $n \in \mathbb{N}$, we can pick $S_n \in E_n$[7].
Then $T \coloneqq \bigcup_{n \in \mathbb{N}} S_n$ is countable by Theorem 45.
Besides, $\forall n \in \mathbb{N}$, $S_n \subset T$ and $|S_n| = n$. Therefore T is infinite by Theorem 29.
Thus $|T| = \aleph_0$ as an infinite countable set. ■

Theorem 47. $|\mathbb{Z}| = \aleph_0$

Proof 1. Since $\mathbb{N} \subset \mathbb{Z}$, we have $|\mathbb{N}| \leq |\mathbb{Z}|$.
Define $f : \mathbb{Z} \to \mathbb{N}$ by $f(n) = \begin{cases} 2^n & \text{if } n \geq 0 \\ 3^{-n} & \text{if } n < 0 \end{cases}$.
Then f is injective by uniqueness of the prime factorization. Therefore $|\mathbb{Z}| \leq |\mathbb{N}|$.
Hence $|\mathbb{Z}| = |\mathbb{N}|$ by Cantor–Schröder–Bernstein theorem. ■

Proof 2.
Define $f : \mathbb{Z} \to \mathbb{N}$ by $f(n) = \begin{cases} 2n & \text{if } n \geq 0 \\ -(2n+1) & \text{if } n < 0 \end{cases}$.
Then f is bijective with inverse $f^{-1}(m) = \begin{cases} k & \text{if } \exists k \in \mathbb{N},\, m = 2k \\ -k-1 & \text{if } \exists k \in \mathbb{N},\, m = 2k+1 \end{cases}$.
Therefore $|\mathbb{Z}| = |\mathbb{N}|$. ■

[6] (ACC) See Remark 58.
[7] (ACC) See Remark 59.

Theorem 48. $|\mathbb{Q}| = \aleph_0$

Remark 49. This theorem asserts that there are as many rational numbers than natural numbers. Which seems counter-intuitive. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, we could expect that $\mathbb{R}$ is also countable. That's not the case as we will see in the next section.

Proof 1. Note that $\mathbb{N} \subset \mathbb{Q}$, therefore $\aleph_0 \leq |\mathbb{Q}|$.
Define $f : \mathbb{Q} \to \mathbb{Z} \times \mathbb{Z}$ by $f\left(\frac{a}{b}\right) = (a, b)$ where $\frac{a}{b}$ is in lowest form.
By uniqueness of the lowest form expression of a rational number, f is well-defined and injective.
Thus $|\mathbb{Q}| \leq |\mathbb{Z} \times \mathbb{Z}|$. Since $|\mathbb{Z}| = |\mathbb{N}|$, we get $|\mathbb{Z} \times \mathbb{Z}| = |\mathbb{N} \times \mathbb{N}| = \aleph_0$.
We conclude using Cantor–Schröder–Bernstein theorem. ■

Proof 2. Note that $\mathbb{N} \subset \mathbb{Q}$, therefore $\aleph_0 \leq |\mathbb{Q}|$.
The function $f : \mathbb{Z} \times \mathbb{N} \setminus \{0\} \to \mathbb{Q}$ defined by $f(a, b) = \frac{a}{b}$ is surjective.
Thus, by Proposition 38[8], $|\mathbb{Q}| \leq |\mathbb{Z} \times \mathbb{N} \setminus \{0\}|$.
Since $|\mathbb{Z}| = |\mathbb{N}|$ and $|\mathbb{N} \setminus \{0\}| = |\mathbb{N}|$, we get $|\mathbb{Z} \times \mathbb{N} \setminus \{0\}| = |\mathbb{N} \times \mathbb{N}| = \aleph_0$.
We conclude using Cantor–Schröder–Bernstein theorem. ■

Proof 3.
Note that $\mathbb{N} \subset \mathbb{Q}$, therefore $\aleph_0 \leq |\mathbb{Q}|$.
Besides $\mathbb{Q} = \bigcup_{(a,b)\in\mathbb{Z}\times\mathbb{N}\setminus\{0\}} \left\{\frac{a}{b}\right\}$, so that $\mathbb{Q}$ is countable by Theorem 45[9], i.e. $|\mathbb{Q}| \leq \aleph_0$.
We conclude using Cantor–Schröder–Bernstein theorem. ■

5 Cantor's diagonal argument

Theorem 50. $\aleph_0 < |\mathbb{R}|$

The following proof relies on Cantor's diagonal argument[10]. That's a very general method that we will use later to prove Cantor's theorem[11].

Proof. We are going to prove that there is no surjection $\mathbb{N} \to \mathbb{R}$ (and hence no such bijection).
Let $f : \mathbb{N} \to \mathbb{R}$ be a function. Given $n \in \mathbb{N}$, we know[12] that $f(n)$ has a unique proper decimal expansion

$$f(n) = \sum_{k=0}^{+\infty} a_{nk} 10^{-k}$$

where $a_{n0} \in \mathbb{Z}$ and $a_{nk} \in \{0, 1, \ldots, 9\}$ for $k \geq 1$, i.e.

$$\begin{aligned}
f(0) &= a_{00} \,.\, a_{01}\, a_{02}\, a_{03}\, a_{04}\, a_{05} \ldots \\
f(1) &= a_{10} \,.\, a_{11}\, a_{12}\, a_{13}\, a_{14}\, a_{15} \ldots \\
f(2) &= a_{20} \,.\, a_{21}\, a_{22}\, a_{23}\, a_{24}\, a_{25} \ldots \\
f(3) &= a_{30} \,.\, a_{31}\, a_{32}\, a_{33}\, a_{34}\, a_{35} \ldots \\
f(4) &= a_{40} \,.\, a_{41}\, a_{42}\, a_{43}\, a_{44}\, a_{45} \ldots \\
&\vdots \qquad\qquad\qquad\quad \vdots
\end{aligned}$$

[8] Actually we only need a weak version of Proposition 38 which doesn't involve the axiom of choice: using the well-ordering principle, we can prove that a surjective function whose domain is $\mathbb{N}$ admits a right inverse.
[9] The axiom of countable choice is not necessary here: the sets are singletons, so there is no choice.
[10] This elegant argument was published by Cantor in 1891, but he gave a previous proof of the uncountability of $\mathbb{R}$ in 1874 (with a modified version in 1879).
[11] It can also be used to prove that the box topology on $\mathbb{R}^{\mathbb{N}}$ is not first-countable, or to derive from Erdös–Kaplansky's theorem (if E is an infinite dimensional vector space then $\dim E^* = |E^*|$) that if E is an infinite dimensional vector space then E is not isomorphic to its dual E^*.
[12] Chapter 6, Theorem 56.

Given $k \in \mathbb{N}$, we set $b_k = \begin{cases} 1 & \text{if } a_{kk} = 0 \\ 0 & \text{otherwise} \end{cases}$.
Then $b = \sum_{k=0}^{+\infty} b_k 10^{-k}$ is a real number written with its unique proper decimal expansion.
Note that for every $n \in \mathbb{N}$, $b \neq f(n)$ since $b_n \neq a_{nn}$ (we use the uniqueness of the proper decimal expansion).
Therefore $b \notin \operatorname{Im}(f)$ and f is not surjective. ■

Theorem 51 (Cantor's theorem)**.** *Given a set E, $|E| < |\mathcal{P}(E)|$.*

Remark 52. As a consequence, we get that there is no greatest cardinal.

Proof of Cantor's theorem.
First, note that $g : E \to \mathcal{P}(E)$ defined by $g(x) = \{x\}$ is injective, therefore $|E| \leq |\mathcal{P}(E)|$.
We are going to prove that there is no surjection $E \to \mathcal{P}(E)$ (and hence no such bijection).
Let $f : E \to \mathcal{P}(E)$ be a function. Define $S = \{x \in E : x \notin f(x)\}$.
Let $x \in E$. If $x \in f(x)$ then $x \notin S$. Otherwise, if $x \notin f(x)$ then $x \in S$. Therefore $f(x) \neq S$ (since one contains x but not the other one).
Thus $S \notin \operatorname{Im}(f)$ and f is not surjective. ■

We already know that $|\mathbb{N}| < |\mathbb{R}|$ and that $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})|$.
The following theorem asserts that actually $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$.

Theorem 53. $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$

Proof.
Define $f : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ by $f(S) = \sum_{n \in S} 10^{-n}$.
Then f is injective by uniqueness of the proper decimal expansion. Thus $|\mathcal{P}(\mathbb{N})| \leq |\mathbb{R}|$.

Define $g : \mathbb{R} \to \mathcal{P}(\mathbb{Q})$ by $g(x) = \{q \in \mathbb{Q} : q < x\}$.
Then g is injective. Indeed, let $x, y \in \mathbb{R}$ be such that $x < y$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists $q \in \mathbb{Q}$ such that $x < q < y$. So $q \notin g(x)$ but $q \in g(y)$. Therefore $g(x) \neq g(y)$.
Hence $|\mathbb{R}| \leq |\mathcal{P}(\mathbb{Q})| = |\mathcal{P}(\mathbb{N})|$ using Theorem 39 since $|\mathbb{Q}| = |\mathbb{N}|$.

We conclude thanks to Cantor–Schröder–Bernstein theorem. ■

TODO: *parler de CH et GCH (par ailleurs dire que ZF+GCH implies AC), peut-être dans une annexe?*
TODO: *Ajouter une annexe sur les nombres algébriques/transcendants avec un commentaire historique (Cantor)*

A What is a set?

The notion of *set* turned out to be necessary in order to handle rigorous definitions of $\mathbb{R}$ as the ones provided by Dedekind and later by Cantor. It is worth noting that set theory first observed a great resistance[13], probably because of the influence of Gauss and Kronecker who shared the *horror of the infinite* from ancient Greek philosophers.

Originally Cantor defined a set as *"a gathering together into a whole of distinguishable objects (which are called the elements of the set)"*[14]. According to current standards, it is a very informal definition. This *naive set theory* was governed by two principles: the *comprehension principle* from which any predicate (i.e. statement) defines a set (i.e. we can define the set of all elements satisfying a given property) and the *extension principle* asserting that two sets are equal if and only if they contain the same elements.

Such an intuitive approach is enough to manipulate sets in everyday mathematics (that's what you did in your courses about calculus, multivariable calculus, linear algebra...). Nonetheless it is not satisfactory since it leads to several paradoxes such as Russell's paradox[15] that we can state as follows using modern notations (and particularly Peano's notation for set membership $\in$).

Since any statement defines a set (*comprehension principle*), $S = \{x : x \notin x\}$ must be a set. Therefore either $S \in S$ but then $S \notin S$ by definition of S, or $S \notin S$ but then $S \in S$ by definition of S. Which leads to a contradiction.

Zermelo (1908) was the first to suggest a more careful axiomatic set theory. Particularly the comprehension principle is weakened to the *separation principle*: given a set, we can define its subset of elements satisfying a given predicate (that's the set-builder notation $\{x \in E : P(x)\}$).

This theory has been subsequently refined by Fraenkel, Solem, von Neumann, and others, giving rise to Zermelo–Fraenkel (ZF) set theory. It is a first order theory[16] with equality and the set membership binary predicate symbol $\in$.

In such a theory, we don't define what is a set: they are the atomic objects over which we use quantifiers[17]. Instead, we have a list of axioms ensuring the existence of some sets and how to define new sets from already defined ones.
There are several equivalent formulations of ZF, for instance this one:

- **Axiom of extensionality:**

$$\forall x \forall y (\forall z (z \in x \Leftrightarrow z \in y) \Leftrightarrow x = y)$$

  Intuitively, this axiom states that two sets are equal if and only if they contain the same elements. Particularly, order doesn't matter and $\{a, a\} = \{a\}$.

- **Axiom of pairing:**

$$\forall x \forall y \exists z \forall w (w \in z \Leftrightarrow (w = x \vee w = y))$$

  This axiom asserts that given two sets x *and* y*, the set* $z = \{x, y\}$ *containing* x *and* y *is well-defined.*
  It is often given as an axiom although it is a consequence of the axiom schema of replacement.
  Note that for a given set x *the axiom of pairing and the axiom of extensionality allow us to define the singleton* $\{x\} = \{x, x\}$.

[13] To which Hilbert later replied with the following wonderful and well-known sentence: *Aus dem Paradies, das Cantor uns geschaffen, soll uns niemand vertreiben können* [*No one should be able to expel us out of the paradise that Cantor has created for us.*].

[14] *"Unter einer 'Menge' verstehen wir jede Zusammenfassung M von bestimmten wohlunterscheidbaren Objekten M unserer Anschauung oder unseres Denkens (welche die 'Elemente' von M genannt werden) zu einem Ganzen"*, in *Beiträge zur Begründung der transfiniten Mengenlehre* by Cantor (1895).

[15] It was discovered by Zermelo in 1899, but he did not published it, and then rediscovered by Russell in 1901.

[16] A first order theory generalizes propositional calculus by introducing quantified variables.

[17] Actually, it is possible to work with a theory about more general objects: that's for example the case in von Neumann–Bernays–Gödel theory where atomic objects are classes and where a set is defined as a class which is contained in another class. It is known that the statements about sets that can be proved within vNBG coincides with the statement that can be proved within ZFC. Additionally, it doesn't involve axiom schema, i.e. it is described using only finitely many axioms.

- **Axiom of union:**

$$\forall x \exists y \forall u (u \in y \Leftrightarrow \exists w \in x (u \in w))$$

This axiom ensures that given a set x (of sets, everything is a set here), the set $\bigcup_{w \in x} w$ is well-defined. We will use the abbrevation $\cup$ in what follows.
By the way, note that we have to be more careful about intersections: what would be $\bigcap_{w \in \varnothing} w$?

- **Axiom of power set:**

$$\forall x \exists y \forall z [z \subset x \Leftrightarrow z \in y]$$

Here $z \subset x$ is an abbreviation for $\forall u (u \in z \implies u \in x)$. This axiom asserts that given a set x, the set $y = \mathcal{P}(x)$ of its subsets is well-defined.

- **Axiom of empty set:**

$$\exists x \forall y \neg (y \in x)$$

This axiom ensures that the empty set exists (and it is unique by extensionality), therefore, in what follows, we introduce the term $\varnothing$ to denote the empty set.

- **Axiom of infinity:**

$$\exists x (\varnothing \in x \wedge \forall y \in x (y \cup \{y\} \in x))$$

This axiom states that there exists a set containing a copy of $\mathbb{N}$.
We are going to use it to give a construction of $\mathbb{N}$. Set $0 := \varnothing$ and $s(y) := y \cup \{y\}$ (that's the successor function), therefore $1 = \{\varnothing\}$, $2 = \{\varnothing, \{\varnothing\}\}$, $3 = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$...
By the axiom of infinity, there exists a set E containing 0 which is closed by s. Then it is not too difficult to prove that the intersection of all the subsets of E containing 0 and closed by s satisfies the induction principle. It is a nice construction of $\mathbb{N}$ where the order is given by inclusion.

- **Axiom schema of replacement:**

$$(\forall x \in a \exists! y P(x, y)) \implies (\exists b \forall y (y \in b \Leftrightarrow \exists x \in a P(x, y)))$$

This one is an axiom schema and not an axiom (i.e. we need it for all formulae $P(x, y)$).
It asserts that if a formula defines a "function" then its "range" is a set.
Together with the axiom of empty set, the axiom schema of replacement implies that given a set, we may define a subset of elements satisfying a given property (set-builder notation): that's a weak/restricted version of the comprehension principle called the separation principle.

- **Axiom of foundation:**

$$\forall x (x \neq \varnothing \Rightarrow \exists y \in x (x \cap y = \varnothing))$$

This axiom is a little bit special: it doesn't define new sets but it is here to avoid paradoxes by removing circular arguments. Particularly, as a consequence of it, a set can't be an element of itself: if x satisfies $x \in x$ then the singleton $\{x\}$ doesn't satisfy the axiom of foundation since $x \cap \{x\} = \{x\}$.

Note that during the interwar period, the French group Bourbaki started to formalize most of known mathematics within set theory.

It is commonly believed that ZF is very likely to be consistent, i.e. there is no contradition (like Russell's paradox). In what follows, I assume that ZF is consistent (otherwise every statement would be true). Nonetheless, a consequence of Gödel's second incompleteness theorem is that if ZF is consistent then we can't prove within ZF that it is.

You should keep in mind that such a theory was given in response to the foundational crisis of mathematics in the late 19^{th} century in order to free mathematics from contradictions and to add more rigor in it. But most mathematicians work at a higher level, not directly from these axioms, and don't care too much about them (we did mathematics before set theory). So, should a contradiction be found, it probably won't

impact that much other fields of mathematics: maybe it would be possible to simply fix the axioms in a way to remove the contradiction, or otherwise to work on new foundations for mathematics... Anyway, some choices were made and they can be changed (and even without finding a contradiction, some mathematicians have objections about using set theory as foundations for mathematics, especially since some fields of mathematics involve proper classes which are too big to be sets, so they are *de facto* excluded from set theory).

Theorem 54. *There is no set containing all sets.*

Proof. Assume that such a set V exists, then the powerset $\mathcal{P}(V)$ exists too and $\mathcal{P}(V) \subset V$ by definition of V. Therefore $|\mathcal{P}(V)| \leq |V|$, but $|V| < |\mathcal{P}(V)|$ by Cantor's theorem. Hence a contradiction. ■

We may similarly prove that there is no set containing all finite sets, or even containing all singletons.

Theorem 55. *There is no set containing all singletons.*

Proof. Assume that the set S of all singletons exists.
Define $f : \mathcal{P}(S) \to S$ by $f(x) = \{x\}$ (which is well-defined). Since f is one-to-one, we get that $|\mathcal{P}(S)| \leq |S|$.
Which contradicts $|S| < |\mathcal{P}(S)|$ (Cantor's theorem). ■

B *Un morceau de choix*

The following statement is (a formulation of) the *axiom of choice*

$$\text{(AC)} \qquad \forall x((\varnothing \notin x \wedge \forall u, v \in x(u = v \vee u \cap v = \varnothing)) \implies \exists y \forall u \in x \exists w(u \cap y) = \{w\})$$

It asserts that given a set x of non-empty pairwise disjoint sets, there exists a set y which contains exactly one element for each set in x. Informally, it means that given infinitely many non-empty sets, we can simultanuously pick an element in each set.

We can also state it in the following way. For I a set together with $(X_i)_{i\in I}$ a family of sets indexed by I, we have

$$\left(\forall i \in I,\ X_i \neq \varnothing\right) \implies \prod_{i\in I} X_i \neq \varnothing$$

i.e. there exists $(x_i)_{i\in I}$ where $x_i \in X_i$ (we can simultaneously pick $x_i \in X_i$ for each $i \in I$).

Gödel and Cohen respectively showed that the axiom of choice is not disprovable in ZF and that it is not provable in ZF (assuming that ZF is consistent)[18]. Therefore the axiom of choice can be added to ZF as an axiom without changing its consistency, in this case the theory is denoted ZFC.

Acceptance of the axiom of choice is a little bit controversial: on the one hand it seems very natural and useful in some areas of mathematics[19] but some consequences are counter intuitive (for instance the well-known Banach–Tarski paradox). For this reason, some mathematicians try to avoid it or to use weaker versions (such as the axiom of countable choice, i.e. only when I is countable).
Here is a (funny) quote summarizing the situation[20]:

"The axiom of choice is obviously true, the well-ordering principle obviously false, and who can tell about Zorn's lemma?"
– Jerry L. Bona[21].

A statement equivalent to the axiom of choice and which is related to the content of this chapter is the following one (which generalizes Remark 23 to infinite sets):

[18] According to Gödel's first incompleteness theorem, ZF contains at least one statement which is undecidable, the axiom of choice is such a statement.
[19] For instance, the axiom of choice is equivalent to the fact that every vector space has a basis.
[20] These three statements are equivalent.
[21] Steven G. Krantz. *Handbook of Logic and Proof Techniques for Computer Science*, p121. Birkhäuser (2002).

Theorem 56 (Trichotomy principle for cardinality)**.**
Given two sets A and B, exactly one of the following occurs:

- $|A| < |B|$
- $|A| = |B|$
- $|A| > |B|$

When Tarski submitted to the *Comptes Rendus de l'Académie des Sciences* his proof that the trichotomy principle is equivalent to the axiom of choice, both Fréchet and Lebesgue refused it: Fréchet because *"an implication between two well known propositions is not a new result"*, and Lebesgue because *"an implication between two false propositions is of no interest"*[22].

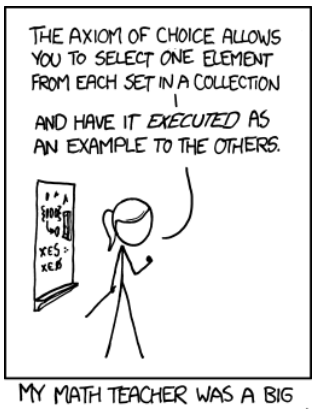
Below I highlight the places where I used either the axiom of choice or the axiom of countable choice in this chapter.

Remark 57. In Proposition 38, the part that $|E| \leq |F|$ implies the existence of a surjection $g : F \to E$ is true in ZF even without the axiom of choice.
Nonetheless, I used the axiom of choice to prove the converse when I pick $(y_x)_{x\in E} \in \prod_{x\in E} g^{-1}(x)$.
Actually the axiom of choice is equivalent to the fact a function is surjective if and only if it admits a right inverse (i.e. $g : F \to E$ is surjective if and only if there exists $f : E \to F$ such that $g \circ f = id_E$).

Remark 58. Within ZF, Theorem 45 is equivalent to the axiom of countable choice.
Nonetheless, using an induction, we can prove within ZF that a finite union of countable sets is countable (see the first part of the proof of Theorem 45).
In the proof of Theorem 45, I used the axiom of countable choice to pick simultaneously injective functions $f_i : E_i \to \mathbb{N}$ for every $i \in I$.

Remark 59. I used the axiom of countable choice in the proof Theorem 46 when applying Theorem 45.
A set is *Dedekind–infinite* if it contains an infinite countable subset[23]. It is true within ZF that a Dedekind–infinite set is infinite. The converse requires the axiom of choice: there exist models of ZF containing amorphous sets, i.e. which are infinite and Dedekind–finite.

[22] Jan Mycielski. *A System of Axioms of Set Theory for the Rationalists*. Notices of the AMS, Volume 53, Number 2.
[23] Another equivalent definition is: *a set E is Dedekind–complete if there exists $A \subsetneq E$ such that $|A| = |E|$.*



FROM: https://xkcd.com/982/

Below are a few other proof methods:

- **Proof by Example/Generalization.** *The statement holds for $n = 42$ so it holds for any $n \in \mathbb{N}$.*
- **Proof by Intimidation.** *Don't be silly, it is trivial.*
- **Proof by Terror.** *When proof by intimidation fails.*
- **Proof by Insignificance.** *Who really cares anyway?*
- **Proof by Homework.** *The proof is left as an exercise to the reader.*
- **Proof by Exhaustion.** *The result is an easy consequence of the following 271 pages.*
- **Proof by Obvious Induction.** *3 is prime, 5 is prime, 7 is prime... hence any odd number greater than 2 is a prime number.*
- **Proof by Omission.** *The reader may easily supply the remaining 314 cases in a similar way.*
- **Proof by the End of the Lecture.** *Since it is already the end of the lecture, I let you finish the proof at home.* (Sorry, I might have really used this one)
- **Proof by Lazyness.**
- **Proof by Postponement.** ***TODO:*** Finish the proof later.
- **Proof by General Agreement.** *All in Favor?*
- **Proof by My Agreement.** *Do you believe me? I believe me...*
(I have been told that I used this one often in MAT237... Can you believe that? I can't believe that!)
- **Proof by Intuition.** *I just have this gut feeling.* (Usually that's how we do research)
- **Proof by Supplication.** *Oh please, let it be true.* (Quite often, research looks like that)
- **Proof by Definition.** *We define it to be true.*
- **Proof by Design.** *We add it as an axiom.*
- **Proof by Authority.** *I've just met Gauss in the elevator, he told me that was true, so it must be!*
- **Proof by Stubbornness.** *The favorite method of a former student of mine.*
- **The Only Valid Proof.** *It is too beautiful to be false.*