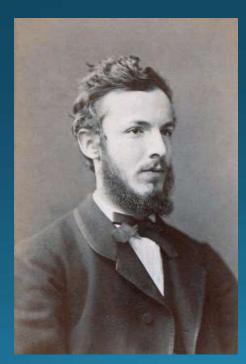
15-251: Great Theoretical Ideas in Computer Science Fall 2016 Lecture 5 Sept. 13, 2016

Cantor's legacy: Countability & Diagonalization



Our heroes for this week

father of set theory

father of computer science

1845-1918

1912-1954

a bein

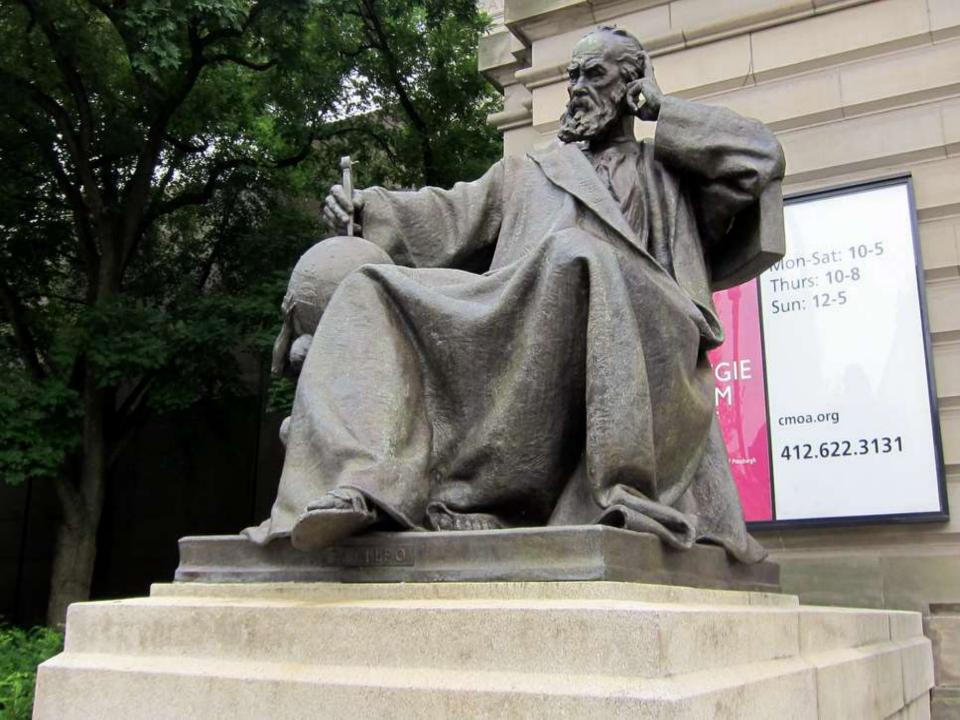
Uncountability

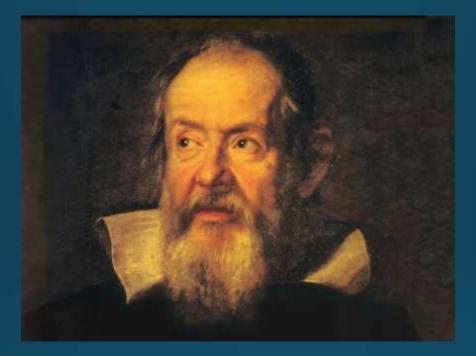
Uncomputability

Poll

Select the ones that apply to you:

- I know what an uncountable set means.
- I know Cantor's diagonalization argument.
- I used to know what uncountable meant, I forgot.
- I used to know the diagonalization argument, I forgot.
- I've never learned about uncountable sets.
- I've never learned about the diagonalization argument.





Galileo (1564–1642)

Best known publication:

Dialogue Concerning the Two Chief World Systems

The three characters

Salviati:

Argues for the Copernican system. The "smart one". (Obvious Galileo stand-in.) Named after one of Galileo's friends.

Sagredo:

"Intelligent layperson". He's neutral. Named after one of Galileo's friends.

Simplicio:

Argues for the Ptolemaic system. The "idiot". Modeled after two of Galilelo's enemies.







Salviati

Simplicio



I take it for granted that you know which of the numbers are squares and which are not.

> I am quite aware that a squared number is one which results from the multiplication of another number by itself; thus 4, 9, etc., are squared numbers which come from multiplying 2, 3, etc., by themselves.

Very well. If I assert that all numbers, including both squares and non-squares, are more than the squares alone, I shall speak the truth, shall I not?

Most certainly.



Salviati

Simplicio



If I should ask further how many squares there are one might reply truly that there are as many as the corresponding number of square-roots, since every square has its own square-root and every square-root its own square...

Precisely so.

But if I inquire how many square-roots there are, it cannot be denied that there are as many as the numbers because every number is the square-root of some square.

This being granted, we must say that there are as many squares as there are numbers because they are just as numerous as their square-roots, and all the numbers are square-roots.

Yet at the outset we said that there are many more numbers to squares.



Sagredo:

What then must one conclude under these circumstances?



Salviati

Cantor (1845–1918)

... Neither is the number of squares less than the totality of all the numbers, ...

... nor the latter greater than the former, ...

... and finally, the attributes "equal," "greater," and "less," are not applicable to infinite, but only to finite, quantities. Good, good...

Good, good...

OOOHHHH! So close! You were almost there, Galileo! Why not say that they are indeed equal?

Let's review Salviati's arguments

 $\mathbb{N} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, \dots\}$ $\bigwedge = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, \dots\}$

"All numbers include[s] both squares and non-squares."

 $S \subsetneq \mathbb{N}$

"Every square has its own square-root and every square-root its own square..." There is a bijection between ℕ and S.

Cantor's Definition

Sets A and B have the same 'cardinality' (size), written |A| = |B|, if there exists a bijection between them.

Note: This is **not** a definition of "|A|". This is a definition of the phrase "|A| = |B|".

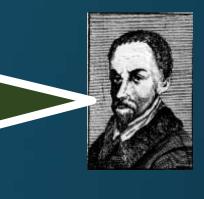
Reminder: what's a bijection?

- It's a perfect matching between A and B.
- It's a mapping f : A → B which is: an injection
 - (i.e., 'one-to-one': $f(a) \neq f(a')$ if $a \neq a'$) & a surjection
 - (i.e., 'onto': $\forall b \in B, \exists a \in A \text{ s.t. } f(a)=b$).
- It's a function f : A → B which has an inverse function, f⁻¹ : B → A (also a bijection).

Cantor's Definition

Sets A and B have the same 'cardinality' (size), written |A| = |B|, if there exists a bijection between them.

E.g.: $|\mathbb{N}| = |\text{Squares}|$ because the function f : $\mathbb{N} \to \text{Squares}$ defined by f(a)=a² is a bijection.

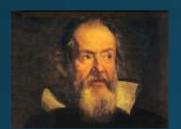


Hold on a sec. We just overloaded notation. Can we at least double-check this all makes sense for **finite** sets?

Sure, that's easy

Hold on a sec. We just overloaded notation. Can we at least double-check this all makes sense for **finite** sets?





Let A = {red, green, blue}. Let B = $\{.03, -2, 18\}$. Let C = $\{1, 2, 3, 4\}$

There is a bijection between A and B, so |A| = |B|. There is no bijection between B and C, so $|B| \neq |C|$.

There is no bijection between C and N, so $|C| \neq |N|$.

Perhaps this definition just captures the difference between finite and infinite?

> Good question. If A and B are infinite sets do we always have |A| = |B|?

That's exactly what I was wondering in 1873...

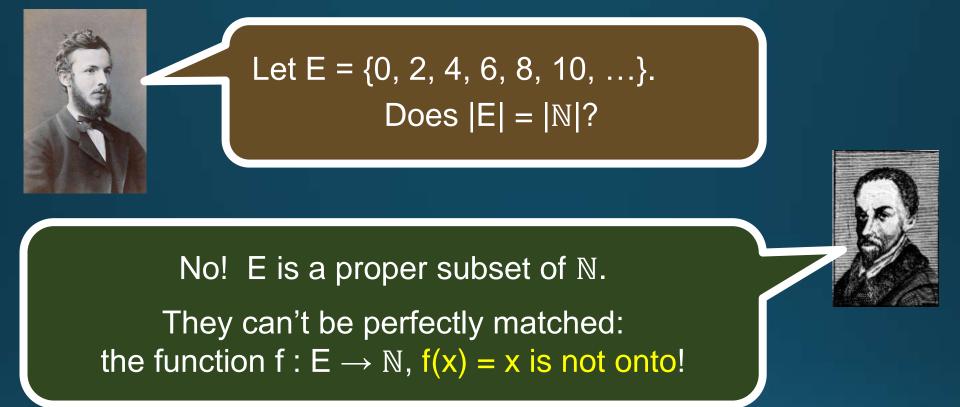
Let's try some examples!







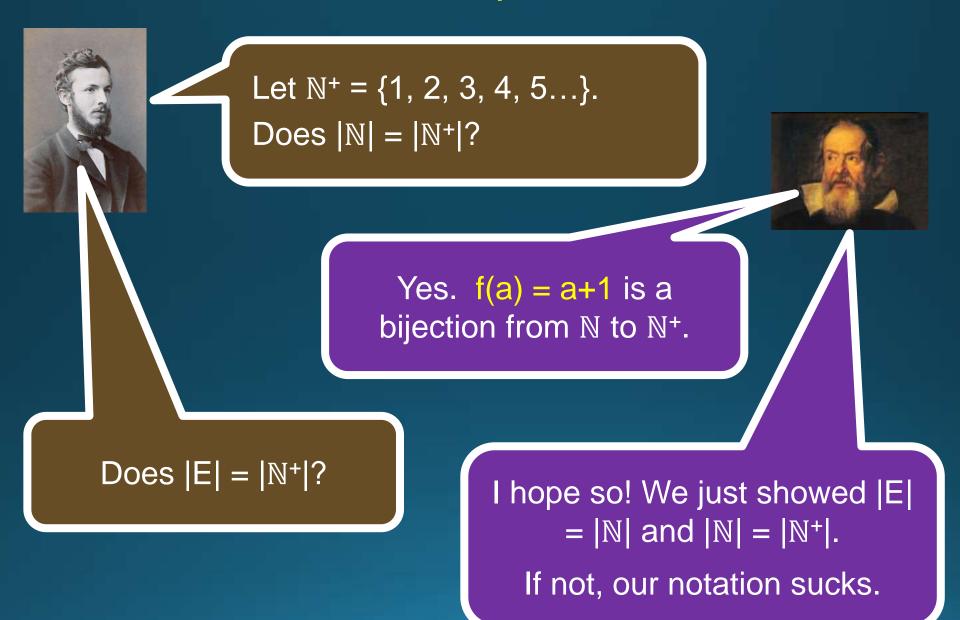
Examples





Wrong Simplicio, that doesn't matter. There does exist a bijection $f : E \rightarrow \mathbb{N}$, namely f(x) = x/2. So $|E| = |\mathbb{N}|$.

Examples



Transitivity

Theorem:

If there is a bijection from A to B (say, f), and there is a bijection from B to C (say, g), then there is a bijection from A to C.

I.e., if |A| = |B| and |B| = |C| then |A| = |C|.

Proof: g • f is a bijection from A to C. (Why?)





$\mathbb{N} = \{0, 1, 2, 3, 4, 5, 6, 7, \ldots\}$ $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$



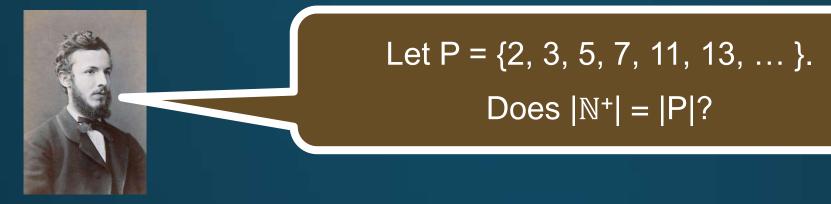


$\mathbb{N} = \{ 0, 1, 2, 3, 4, 5, 6, 7, \ldots \}$ $\mathbb{Z} = \{ 0, -1, +1, -2, +2, -3, +3, -4, \ldots \}$

It's looking good...

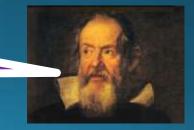
 $f(a) = (-1)^a [a/2]$ is a bijection from N to Z.

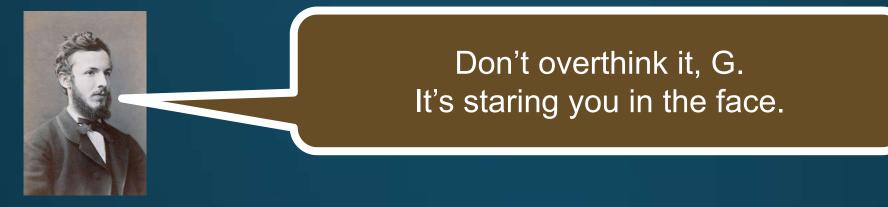


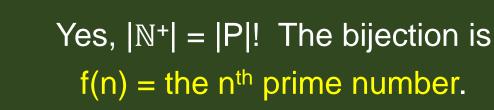


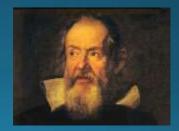
$\mathbb{N}^{+} = \{ 1, 2, 3, 4, 5, 6, 7, 8, \ldots \}$ $\mathsf{P} = \{ 2, 3, 5, 7, 11, 13, 17, 19, \ldots \}$

Hmm... It's looking good... And yet...











He's right, Galileo. Totally legit.

That can't be legit.

Why not? It's a well-defined function, isn't it? It's a bijection, isn't it?



 $\mathbb{N} = \{0, 1, 2, 3, 4, 5, \overline{6}, 7, \ldots\}$ $\mathbb{E} = \{0, 2, 4, 6, 8, 10, 12, 14, \ldots\}$ $\mathbb{Z} = \{0, -1, +1, -2, +2, -3, +3, -4, \ldots\}$ $\mathbb{P} = \{2, 3, 5, 7, 11, 13, 17, 19, \ldots\}$

If S is an infinite set and you can list off its elements as s_0 , s_1 , s_2 , s_3 , ... uniquely, in a well-defined way, then |S| = |N|.

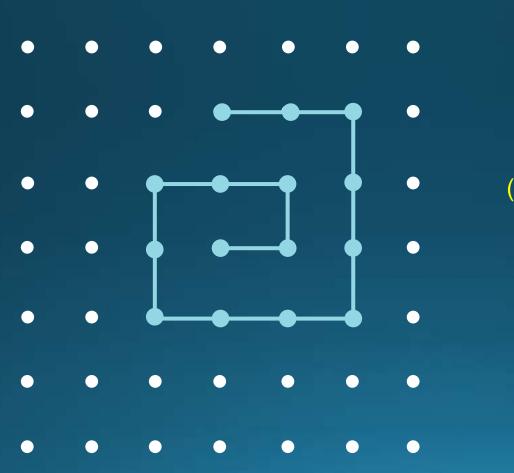


Any set S with $|S| = |\mathbb{N}|$ is called **countably infinite**.

A set is called **countable** if it is either finite or countably infinite.

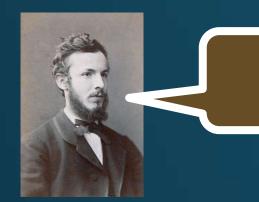


So \mathbb{Z} is countable. Is \mathbb{Z}^2 countable?



(0,0) (1,0) (1,1) (0,1) (-1,1)(-1,0) (-1,-1) (0,-1) (1,-1) (2,-1) (2,0) (2,1) (2,2) (1,2) (0,2)

Spiral. Etc. So yes.



What about **Q**, the rationals? Countable?

Come on, no way! How could you list them in a sequence? Between any two rationals there are infinitely many more.

Not so fast...



Take our listing of \mathbb{Z}^2 :

 $(0,0), (1,0), (1,1), (0,1), (-1,1), \dots, (2,-1), \dots, (2,1), \dots, (1,2), \dots, (-1,2), \dots$

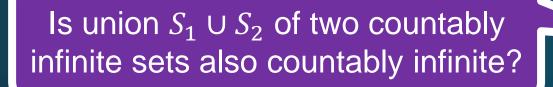
1, 0, -1, -2, 2, 1/2, -1/2, ...

To get a listing of Q, go through the above list in order.

If you are at (p,q), output p/q...

... if $q \neq 0$ and you haven't output this rational yet. If q = 0 or you've seen p/q before, just go on to next one.

> This indeed lists all of the rationals exactly once. So \mathbb{Q} is countable.





Yes, just list elements of S_1 first and then those of S_2



Sure, Simplicio? Will you ever get to elements in S_2 ?

Oops, sorry. Alternate elements of S_1 and S_2 f(0), g(0), f(1), g(1), f(2), g(2), ...if $f : \mathbb{N} \to S_1$ and $g : \mathbb{N} \to S_1$ are bijections.



Right. Similarly any finite union of countably infinite sets is also countably infinite



What about a countable union of countably infinite sets, G. ?

Hmm, seems tricky, yet familiar...



Good practice problem...

If there is an injection (one-to-one map) from A to B, we say $|A| \le |B|$.

E.g.: f(a) = a is an injection from Squares $\rightarrow \mathbb{N}$; f(x) = (p,q) when x = p/q in lowest terms is an injection from $\mathbb{Q} \rightarrow \mathbb{Z}^2$.

Suppose there is an injection $A \rightarrow B$, so $|A| \leq |B|$. Suppose there's also an injection $B \rightarrow A$, so $|B| \leq |A|$. If our notation doesn't suck, it should mean that |A|=|B|. So must there be a **bijection** between A and B?

> Yes! This is the "Cantor– Bernstein–Schröder Theorem". Shall I prove it for you? It'll take 3-4 slides.

Nah, I'll check it on Wikipedia

If there is a surjection (onto map) from A to B, we say $|A| \ge |B|$.

Here's a clearer way to show \mathbb{Q} is countable: \mathbb{Z}^2 is countable so it suffices to show $|\mathbb{Z}^2| \ge |\mathbb{Q}|$. Define $f: \mathbb{Z}^2 \to \mathbb{Q}$ by $f(p, q) = \begin{cases} p/q & \text{if } q \neq 0, \\ 0 & \text{if } q = 0. \end{cases}$ This is clearly a surjection, so $|\mathbb{Z}^2| \ge |\mathbb{Q}|$.

Suppose there's a surjection $f : A \rightarrow B$, so $|A| \ge |B|$. If our notation doesn't suck, then presumably $|B| \le |A|$, meaning there should be an injection g : B \rightarrow A. Is there?

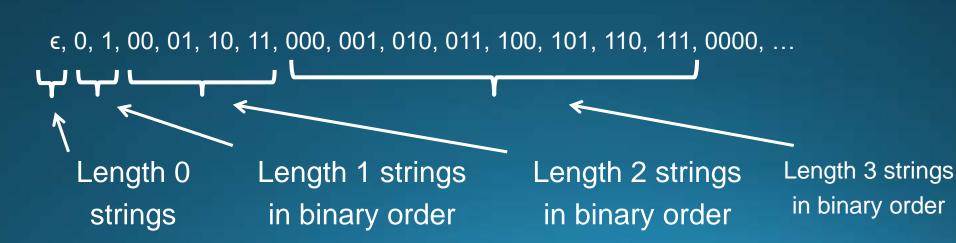
Sure. For any b∈B, define g(b) to be any element a such that f(a)=b.
(Such an a must exist ∵ f is a surjection.) This g is an injection (why?).



This requires the Axiom of Choice, which we will always assume in 251. Let's do one more example. Let {0,1}* denote the set of all binary strings of any finite length. Is {0,1}* countable?

Yes, this is easy. Here is my listing:





15 slides ago, Simplicio and I asked if every infinite set has the same cardinality. Now we've seen the squares, evens, primes, integers, rationals, {0,1}*, Z²
etc. are all countably infinite: they have the same cardinality as N. So are all infinite sets countable?

Yeah, I was thinking about all this in 1873. In particular, about the next obvious question: Is ℝ (set of real numbers) countable? *My motivation was a simpler proof of Liouville's thorem that transcadental numbers exist*

Anyway, I proved **ℝ** is uncountable in December 1873.

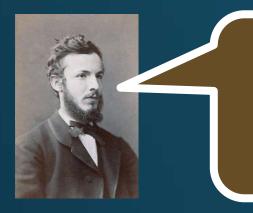
But when I wrote the paper, I kind of focused on countability of Z^d, the number theory application, etc.

'Cause, you know, I could tell there was going to be a lot of controversy over my radical new ideas on "different sizes of infinity". l feel you, man. The 1873 proof was specifically tailored to \mathbb{R} .

In 1891, I described a much slicker proof of uncountability.

People call it...

The Diagonal Argument



I'll use the diagonal argument to prove the set of all **infinite** binary strings, denoted {0,1}[∞], is uncountable.

Examples of infinite binary strings:

 $\begin{array}{l} x = 0000000000000000000000000... \\ y = 010101010101010101010101010... \\ z = 10110111011110111110111110... \\ w = 001101010001010001010001000... \\ (Here w_n = 1 \mbox{ if and only if n is a prime.}) \end{array}$



Yep.

I'll use the diagonal argument to prove the set of all **infinite** binary strings, denoted {0,1}[∞], is uncountable.

Interesting! I remember we showed that {0,1}*, the set of all finite binary strings, is countable.

What about \mathbb{R} ?

We'll come back to it. Anyway, strings are more interesting than real numbers, don't you think?

Suppose for the sake of contradiction that you can make a list of all the infinite binary strings.

For illustration, perhaps the list starts like this:

- **2**: 1011011101111011111011...
- **3**: 0011010100010100010100...
- **4**: 0101001111111111111111...

Consider the string formed by the 'diagonal':

0: 00000000000000000000000000000... 1: 010101010101010101010101... 2: 101101110111101111011... 3: 001101010000101000101000... 4: 010100111111111111111111... 5: 1100010000000000000000000...

Consider the string formed by the 'diagonal':

0: 000000000000000000000000000000...
1: 010101010101010101010101...
2: 1011011101111011111011...
3: 0011010100010101000101000...
4: 0101001111111111111111111...
5: 110001000000000000000000000...

Actually, take the negation of the string on the diagonal: 100010...

It can't be anywhere on the list, since it differs from every string on the list! **Contradiction.**

- 1: 010101010101010101010101...
- **2**: 101101110111101111011...
- **3**: 0011010100010100010100...
- **4**: 0101001111111111111111...

Here is the same proof, using different words:

Suppose for contradiction's sake that $\{0,1\}^{\infty}$ is countable. Thus $|\mathbb{N}| \ge |\{0,1\}^{\infty}|$;

i.e., there's a surjection f : N → {0,1}[∞].
Define an infinite binary string w∈{0,1}[∞] by w_n = ¬ f(n)_n.
We claim that w ≠ f(m) for every m∈N. This is because, by definition, they disagree in the mth position.
Therefore f is not a surjection onto {0,1}[∞], contradiction.

Awesome. So not every infinite set is countable. {0,1}[∞] has larger cardinality than the set ℕ.

So what about \mathbb{R} ?

 \mathbb{R} is uncountable. Even the set [0,1] of all reals between 0 and 1 is uncountable.

This is because there is a bijection between [0,1] and $\{0,1\}^{\infty}$. Hence $|\mathbb{R}| \ge |[0,1]| = |\{0,1\}^{\infty}| > |\mathbb{N}|$.

What's the bijection between [0,1] and {0,1}[∞]?

 $\pi - 3$



It's just the function f which maps each real number between 0 and 1 to its binary expansion!

- $1/3 = 1/4 + 1/16 + 1/64 + \dots$
 - ↔ .0101010101...
 - .14159265358979323...₁₀ .00100100001111110...₂

Um, technically that's not a surjection. It misses, e.g., .0111111111111111...



It's just the function f which maps each real number between 0 and 1 to its binary expansion.

 $\pi - 3$

- $1/3 = 1/4 + 1/16 + 1/64 + \dots$
 - ↔ .0101010101...
 - .14159265358979323...₁₀ .00100100001111110...₂

Um, technically that's not a surjection. It misses, e.g., .0111111111111111...

> You're saying because this also equals 1/2? In the same way that, in base 10, .499999... is the same as .500000...?

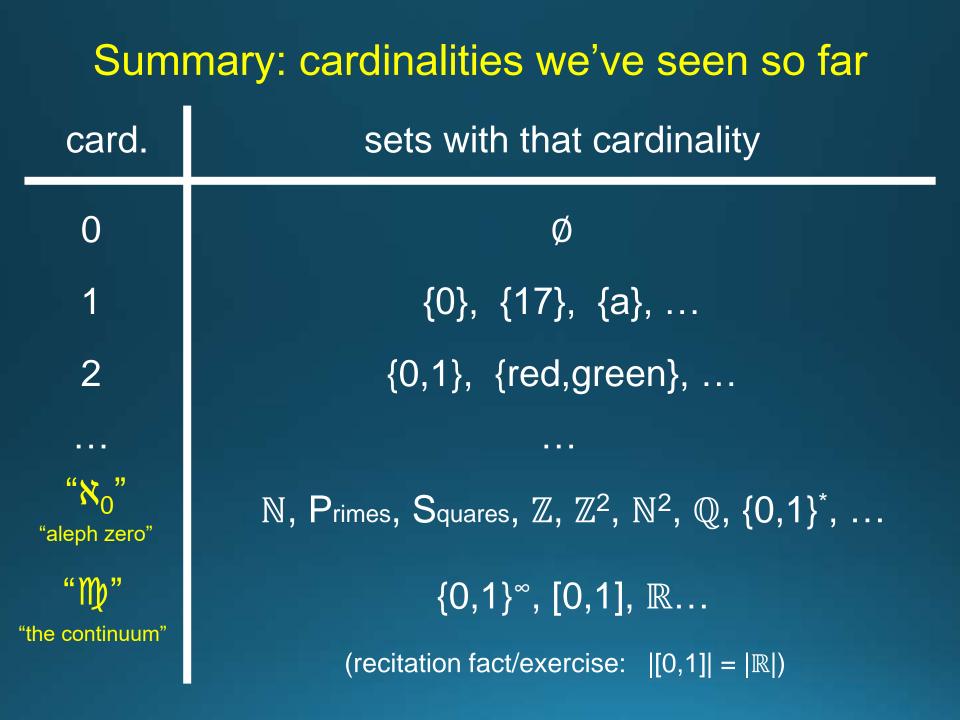
Yeah.

Sorry.

Ugh. I was hoping you wouldn't notice that. This was all so elegant – and you had to go and bring that up!

There are a variety of hacks you can use to get around this issue.

I'll make the TAs go over one or two such hacks in recitation.



Cantor's Theorem

Theorem: For any non-empty set A, $|A| < |\mathcal{P}(A)|.$

 $S = \{1, 2, 3\}$

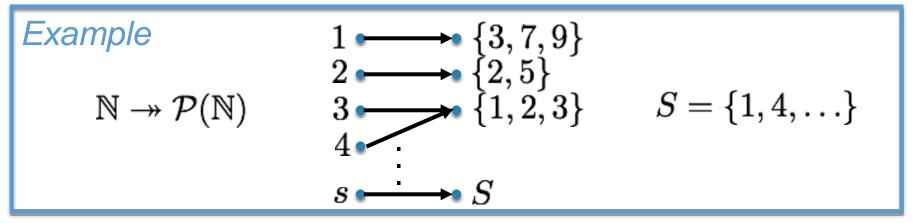
$$\begin{split} \mathcal{P}(S) &= \{ \emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{2,3\}, \{1,3\}, \{1,2,3\} \} \\ & S = \{1,2,3\} \\ |\mathcal{P}(S)| = 2^{|S|} & 1 & 0 & 1 & \longleftrightarrow \{1,3\} \\ \mathcal{P}(S) &\leftrightarrow \{0,1\}^{|S|} & 0 & 0 & \longleftrightarrow \emptyset \end{split}$$

binary strings of length |S|

We just proved $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})|.$

Proof of Cantor's Theorem

Assume $|\mathcal{P}(A)| \leq |A|$ for some set ASo $A \twoheadrightarrow \mathcal{P}(A)$ Let f be such a surjection.



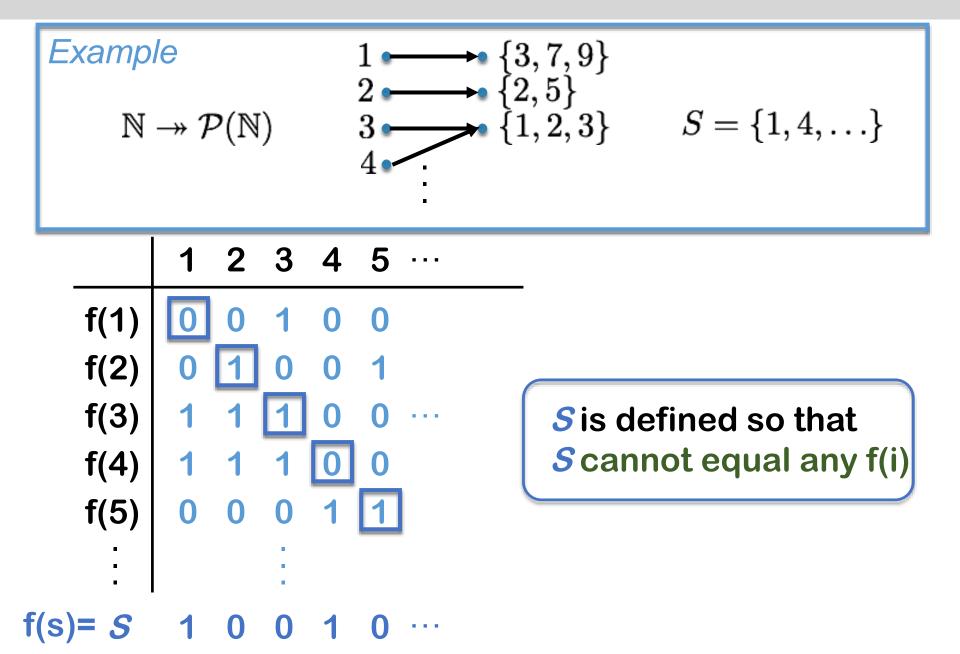
Define $S = \{a \in A : a \notin f(a)\} \in \mathcal{P}(A).$

Since f is a surjection, $\exists s \in A$ s.t. f(s) = S

But this leads to a contradiction: Is $s \in S$?

if
$$s \in S$$
 then $s \notin f(s) = S$
if $s \notin S$ then $s \in f(s) = S$

Cantor's Theorem – Why is this diagonalization



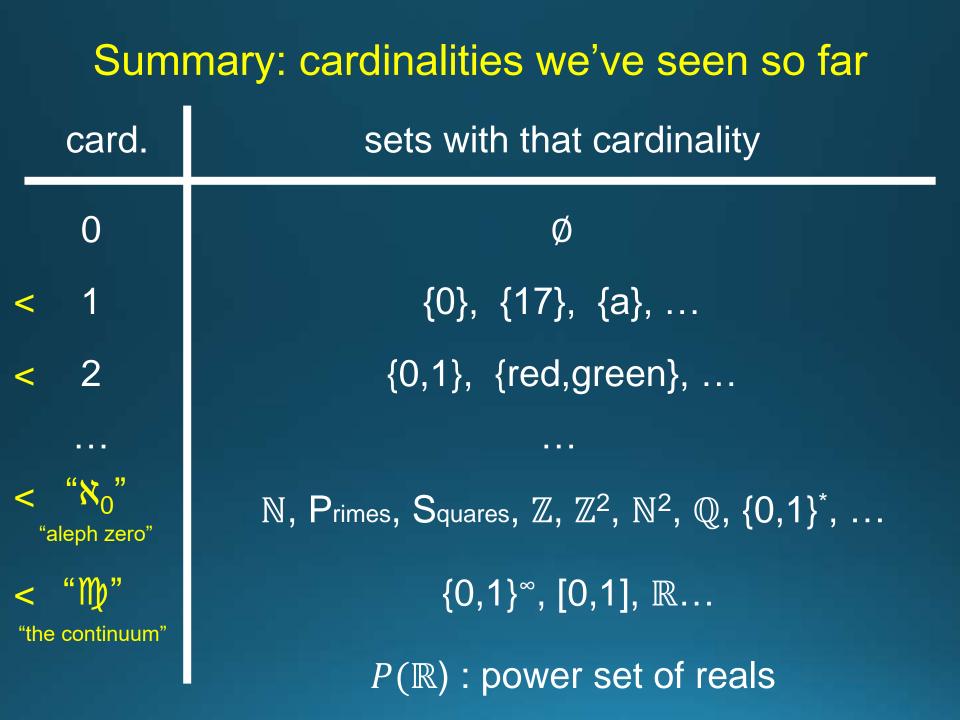
Cantor's Theorem

Theorem: For any non-empty set A, $|A| < |\mathcal{P}(A)|$.

So:

$|\mathbb{N}| < |\mathcal{P}(\mathbb{N})|$. I.e. $\mathcal{P}(\mathbb{N})$ is uncountable.

$|\mathbb{N}| < |\mathcal{P}(\mathbb{N})| < |\mathcal{P}(\mathcal{P}(\mathbb{N}))| < |\mathcal{P}(\mathcal{P}(\mathcal{P}(\mathbb{N})))| < \cdots$ (an infinity of infinities)



Fact: There are no infinite sets with cardinality less than $|\mathbb{N}|$.

Question: Is there any set S with $|\mathbb{N}| < S < |\mathbb{R}|$?

I didn't think so, and called this the Continuum Hypothesis. I spent a really long time trying to prove it, with no success. 🛞 There's a reason you failed... And it's not because the Continuum Hypothesis is false...

Question: Is there any set S with $|\mathbb{N}| < S < |\mathbb{R}|$?





I didn't think so, and called this the Continuum Hypothesis. I spent a really long time trying to prove it, with no success. 🛞 Proving sets countable: the computer scientist's method

We showed $|\{0,1\}^*| = |\mathbb{N}|$.

Actually, if Σ is any finite "alphabet" (set) then $\Sigma^* = \{ all \text{ finite strings over alphabet } \Sigma \}$ is also countably infinite.

E.g., if $\Sigma = \{0, 1, ..., 9, a, b, ..., z, +, -, *, /, ^\}$:

ε, 0, 1, ..., a, ..., /, ^, 00, 01, ..., 0a, 0/, 0^, 10, ..., ^/, ^^, 000, 001, ...

Proving sets countable: the computer scientist's method

Suppose we want to show set S is countable.

Since $|\Sigma^*|$ is countably infinite, it suffices to find a surjection $\Sigma^* \to S$. This implies $|\mathbb{N}| = |\Sigma^*| \ge S$.

To give such a surjection, just need to describe a well-defined rule which maps each string to an element of S, and which covers all elements of S. Proving sets countable: the computer scientist's method

Ex. problem: Prove that Q[x] is countable.

Valid solution:

Any polynomial in $\mathbb{Q}[x]$ can be described by a finite string over the alphabet $\Sigma = \{0, 1, ..., 9, x, +, -, *, /, ^\}.$

(For example: $x^3-1/4x^2+6x-22/7$.)

Proving sets countable using computation

Remember Galileo was a little uncomfortable with the bijection $f : \mathbb{N} \rightarrow Primes$, defined by f(n) = the nth prime' ?

We said it was okay as long as f is a 'well-defined rule'.

A particular kind of well-defined rule: anything "computable by a computer program" (in your favorite language). Proving sets countable using computation

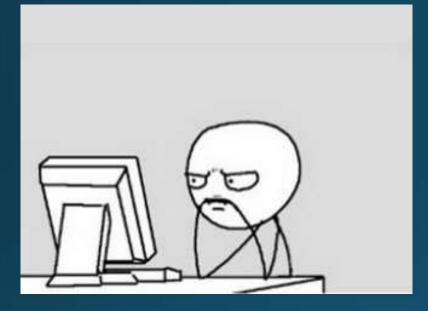
For example, f(n) = 'the nth prime'. You could write a program (Turing machine) to compute f. So this is a well-defined rule.

Or: $f(n) = the n^{th}$ rational in our listing of \mathbb{Q} . (List \mathbb{Z}^2 via the spiral, omit the terms p/0, omit rationals seen before...) You could write a program to compute this f.

A caveat (and spoiler)

There are well-defined rules which cannot be computed by a computer program.





Definitions: Cardinality Countable

Theorem/proof: Countability of various sets.

Study Guide

The diagonal method: uncountability of {0,1}[∞] and [0,1]