EDAA40

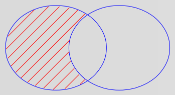
Discrete Structures in Computer Science

4: To infinity and beyond

Jörn W. Janneck, Dept. of Computer Science, Lund University

$R = \{x : x \notin x\}$

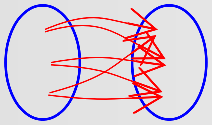
sets



$\supset$

$\heartsuit \subseteq P \times Q$

relations



$\supset$

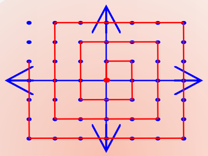
$f : A \longrightarrow B$

functions

$A \hookrightarrow B$

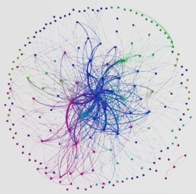
investigate

infinity



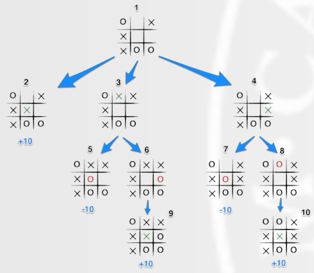
$\cup$

graphs



$\supset$

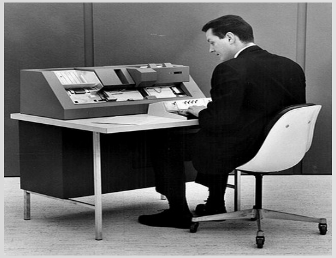
trees



working with infinite
(or arbitrarily large) stuff

definition, construction,
recursion, induction
(also: proofs, logic)

why infinity matters to computing



Everything is finite.

So are computers.

Then why do we care about infinity in maths for computing?

Infinity is often used as a *model*, an abstraction, of two kinds of phenomena:

- an awful lot, i.e. very many

  e.g. the size of a computer's main memory

- the absence of a finite bound

  e.g. the length of a video stream

why infinity matters to this course

It is a property of the math toolbox we use.
A professional knows his or her tools.

This is also an *application* of the tools we have looked at so far.
Using them, we can investigate a part of math, and maybe uncover a few non-trivial, maybe even surprising, truths about it.

intuition

> “Je le vois, mais je ne le crois pas!”
>
> Cantor to Dedekind, about having shown that $[0,1]_{\mathbb{R}} \sim [0,1]_{\mathbb{R}}^2$

Some of the following may seem to run against intuition.

This is a good thing.

Intuitions are extremely useful, but they summarize and stereotype past experiences. So when they are applied to new stuff, they sometimes break.

Then they need updating. That’s one of the goals of this lecture: update your intuitions about amounts of things in sets, for infinite sets.

a simple problem

You all know the natural numbers: 0, 1, 2, 3, 4, … and so on.

A boolean function on the natural numbers is one that yields for each natural number either *true* or *false*: f(177) = true, f(100234) = false, ...

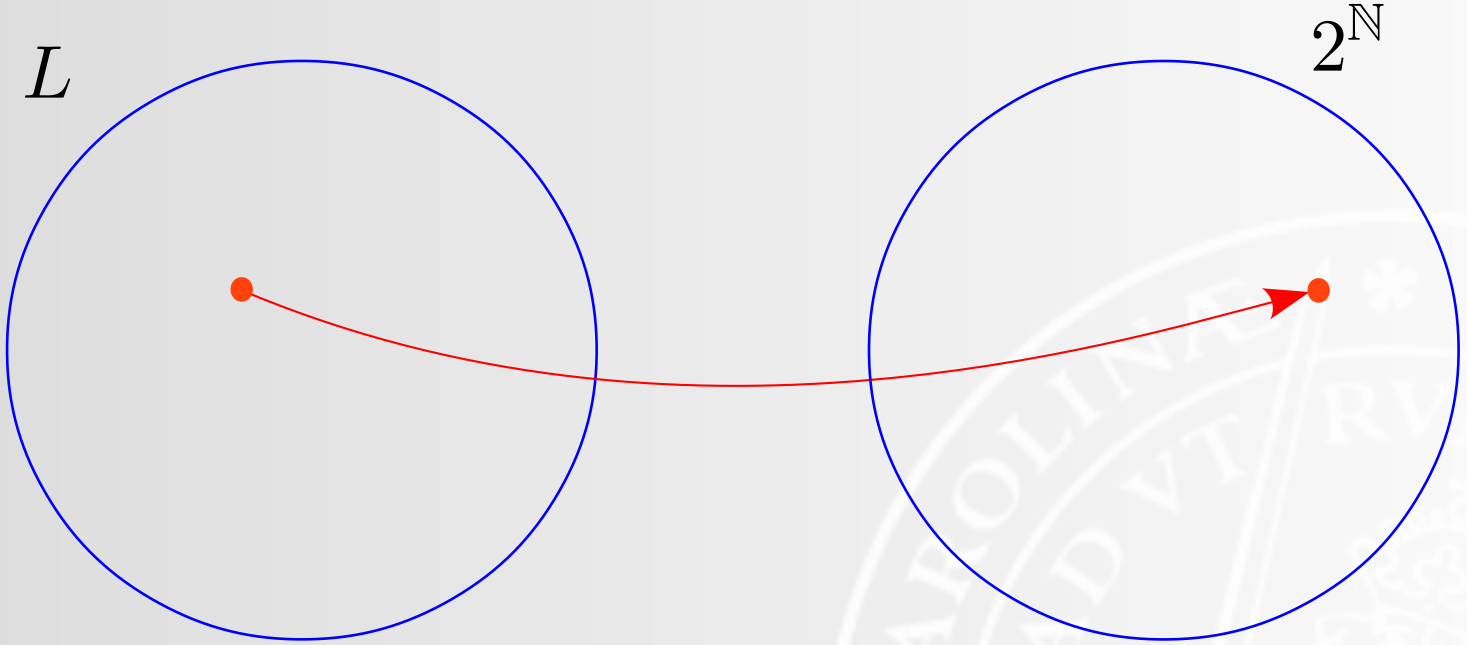
Let's suppose we have an infinite computer, i.e. we ignore any physical constraints of the computer itself, such as address space, memory size, word size, speed, …

A program for that computer is an arbitrarily long (but finite) string of characters in some programming language, arbitrarily "powerful", let's call it L.

The question:

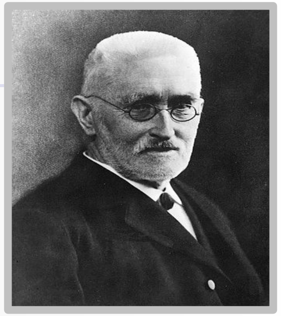
> Is it possible to create a programming language L, such that
> every boolean function on the natural numbers can be written
> as a program in L?

a simple problem - approach



Could there be a surjective function
from L (the set of programs in language L)
onto the set of all functions from the natural numbers to a set of two values?

infinite sets



Richard Dedekind
1831-1916

There are various ways of defining infinite sets.
This is one by Dedekind, 1888:

A is *infinite* if it is equinumerous to a <u>proper</u> subset of itself.
That is, there is some S such that

$$S \subset A \text{ and } S \sim A$$

Show that the natural numbers are an infinite set.
1. Find a proper subset.
2. Construct a bijection between it and the natural numbers.

denumerable (countable) sets

A is *denumerable* (*countable*) if it is equinumerous to the natural numbers, i.e.

$$A \sim \mathbb{N}$$

Denumerable sets are very important to math and CS.
They are the smallest infinite sets. (We won't prove that.)

The cardinality of the natural numbers (and thus all denumerable sets) has a name:

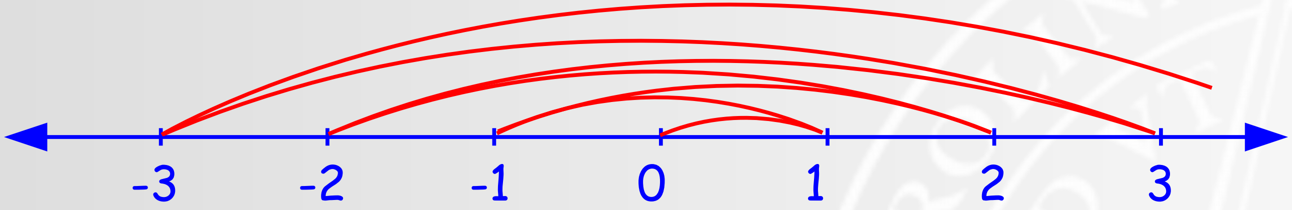
$$\#(\mathbb{N}) = \aleph_0$$

$\aleph_0$ is therefore the smallest *transfinite cardinal number*.

$\mathbb{Z}$: the integers

$\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$ is the set of integers.

$\#(\mathbb{Z})$?



And this is the bijection:

$$z : \mathbb{Z} \longleftrightarrow \mathbb{N}$$

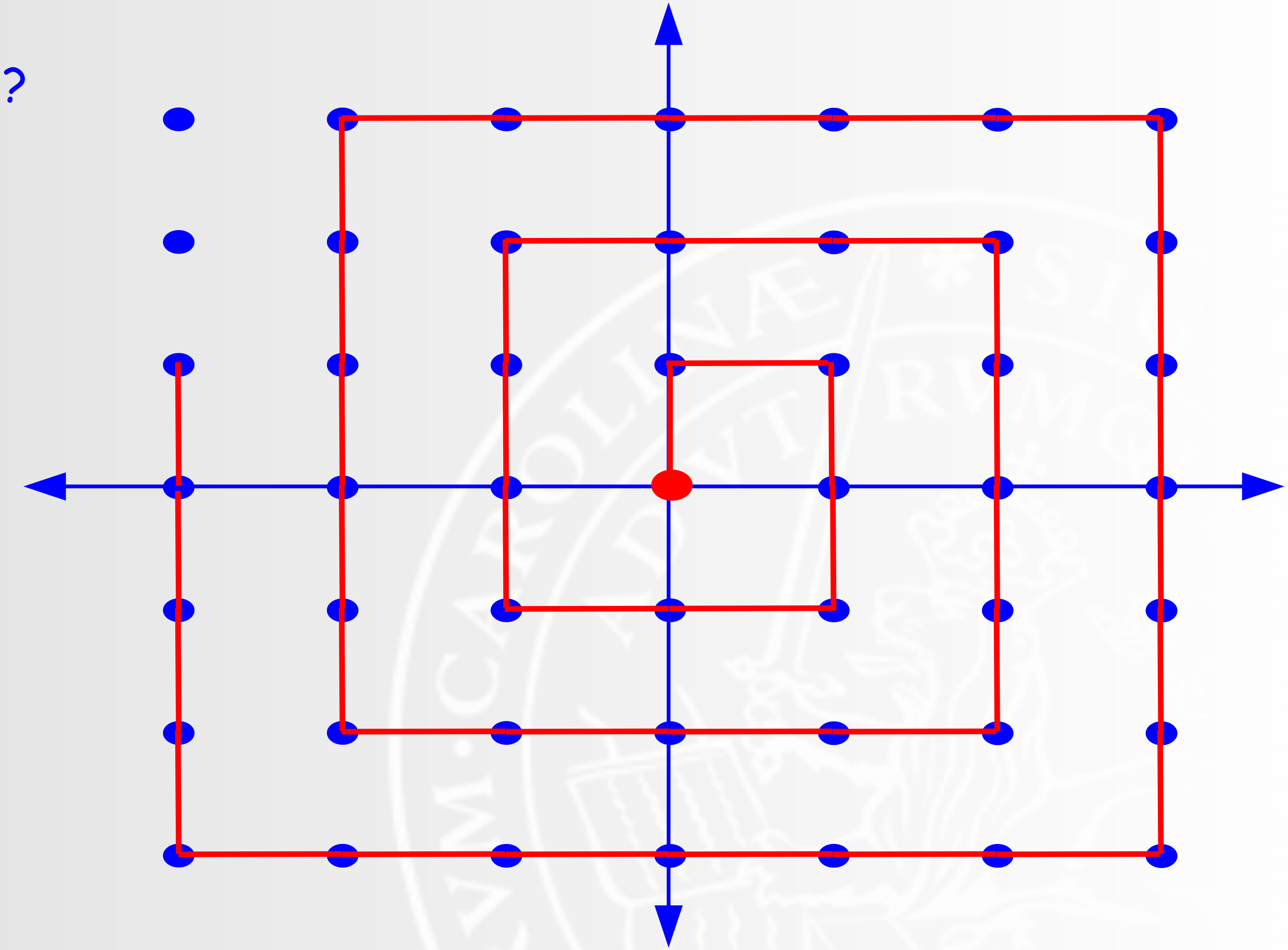
$$z : i \longmapsto \begin{cases} 2i - 1 & \text{for } i > 0 \\ -2i & \text{otherwise} \end{cases}$$

products

What is the cardinality of $\mathbb{Z}^2$?

$\#(\mathbb{Z}^3)$?

$\#(\mathbb{Z}^n)$?

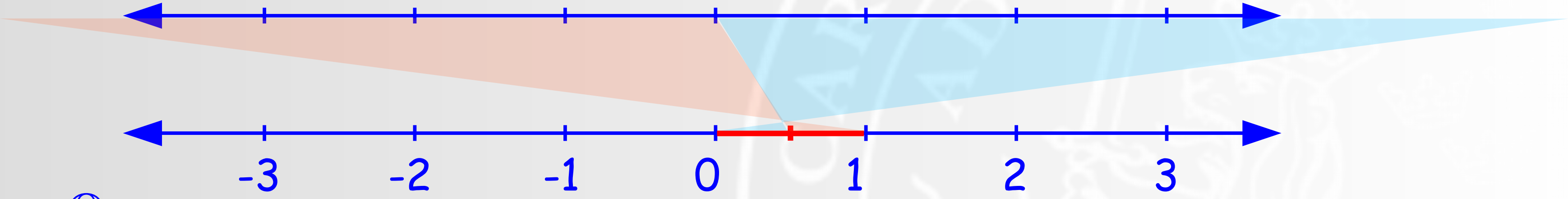


$\mathbb{Q}$: the rational numbers

Some properties:

1. $\mathbb{Q}$ is *dense*: between any two distinct $r, s \in \mathbb{Q}$ there is $\frac{r+s}{2} \in \mathbb{Q}$

2. Any non-empty open interval $]r, s[\subset \mathbb{Q}$ is equinumerous to $\mathbb{Q}$



For example:

$$p_{\mathbb{Q}} : \;]0,1[_{\mathbb{Q}} \longleftrightarrow \mathbb{Q}$$

$$x \longmapsto \begin{cases} \frac{1}{x} - 2 \text{ for } x < \frac{1}{2} \\ \frac{1}{x-1} + 2 \text{ for } x \geq \frac{1}{2} \end{cases}$$

$\mathbb{Q}$: the rational numbers

$\#(\mathbb{Q})$?

Let's start with the simple stuff, i.e. $\aleph_0 = \#(\mathbb{N}) \leq \#(\mathbb{Q})$

Then we define an injection from $\mathbb{Q}$ into $\mathbb{Z}^2$:

$$\begin{array}{rcl} f : \mathbb{Q} & \hookrightarrow & \mathbb{Z}^2 \\ \dfrac{p}{q} & \longmapsto & (p, q) \end{array}$$

(We assume a fully reduced fraction.)

Therefore, we know that $\#(\mathbb{Q}) \leq \#(\mathbb{Z}^2) = \aleph_0$

Put this together: $\aleph_0 = \#(\mathbb{N}) \leq \#(\mathbb{Q}) \leq \#(\mathbb{Z}^2) = \aleph_0$

$$\#(\mathbb{Q}) = \aleph_0$$

finite sequences/strings

Let A be a finite set of n symbols $A = \{a_1, ..., a_n\}$.

The set of all finite sequences (strings) of these symbols is A^*

The empty sequence is $\varepsilon \in A^*$.

What is A^* and $\#(A^*)$? if A is...

$$A = \{\bullet\}$$
$$A = \{0, 1\}$$
$$A = \{a, b, c, ..., v, w, x, y, z\}$$
$$A = \text{UTF-16}$$

This is one injection:

$$\begin{aligned} val : A^* &\hookrightarrow \mathbb{N} \\ c_1 ... c_L &\longmapsto \sum_{i=1..L} v(c_i) n^{i-1} \\ \varepsilon &\longmapsto 0 \end{aligned}$$

with

$$\begin{aligned} v : A &\longrightarrow \mathbb{N} \\ a_j &\longmapsto j \end{aligned}$$

infinite sequences

Let A be a finite set of n symbols $A = \{a_1, ..., a_n\}$.

An *infinite sequence in A* is a function $s : \mathbb{N} \longrightarrow A$

The set of all infinite sequences in A: $A^{\mathbb{N}}$

As always: $\#(A^{\mathbb{N}})$?

$$A = \{\bullet\}$$

$$A = \{0, 1\} = 2$$

Note: You can think of $\{0,1\}^{\mathbb{N}}$ as the powerset $\mathcal{P}(\mathbb{N})$ of $\mathbb{N}$.

Cantor's diagonal construction

Georg Cantor
1845-1918

1. Let's start by <u>assuming</u> that $\mathbb{N} \sim 2^{\mathbb{N}}$, i.e. there must be a bijection $f : \mathbb{N} \longleftrightarrow 2^{\mathbb{N}}$.
   Recall that a bijection is also surjective, i.e. $f(\mathbb{N}) = 2^{\mathbb{N}}$

2. Assuming an f, we can construct the diagonal sequence D:
   $D = 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, \ldots$

3. Invert D:
   $\overline{D} = 0, 1, 0, 1, 1, 0, 1, 1, 0, 0, \ldots$

4. Note that
   $$\overline{D} \notin f(\mathbb{N})$$

5. This contradicts the assumption that f is a bijection.

Conclusion:
There is no bijection $\mathbb{N} \longleftrightarrow 2^{\mathbb{N}}$

$s_i : \mathbb{N} \longrightarrow 2$

$f : \mathbb{N} \longrightarrow 2^{\mathbb{N}}$

	0	1	2	3	4	5	6	7	8	9	…
0	1	0	0	1	1	1	0	0	0	1	
1	1	0	1	1	0	0	1	1	0	1	
2	0	0	1	1	0	1	1	0	1	0	
3	1	1	1	0	1	0	0	1	0	0	
4	0	1	0	1	0	1	0	1	0	1	…
5	1	1	0	0	1	1	0	0	1	1	
6	0	1	0	1	0	1	0	0	1	0	
7	0	0	0	0	1	1	1	0	0	0	
8	1	1	0	0	0	1	1	1	1	0	
9	0	1	1	0	0	0	1	1	1	1	
⋮						⋮					⋱

but wait...

We can list all finite strings of 0s and 1s systematically, like so:

0	ε
1	0
2	1
3	00
4	01
5	10
6	11
7	000
8	001
9	010
10	011
11	100
...	...

Why could we not do the same for infinite sequences of 0s and 1s?

0	00000000000000000000000000...
1	10000000000000000000000000...
2	01000000000000000000000000...
3	11000000000000000000000000...
4	00100000000000000000000000...
5	10100000000000000000000000...
6	01100000000000000000000000...
7	11100000000000000000000000...
8	00010000000000000000000000...
9	10010000000000000000000000...
10	01010000000000000000000000...
11	11010000000000000000000000...
...	...

Is there an infinite sequence not in that list?

What is the property of the sequences in the list?

more than $\aleph_0$

To summarize:

1. We have $\aleph_0 = \#(\mathbb{N}) \leq \#(2^{\mathbb{N}}) = 2^{\aleph_0}$
2. … but we cannot construct a bijection $\mathbb{N} \longleftrightarrow 2^{\mathbb{N}}$
3. Conclusion:

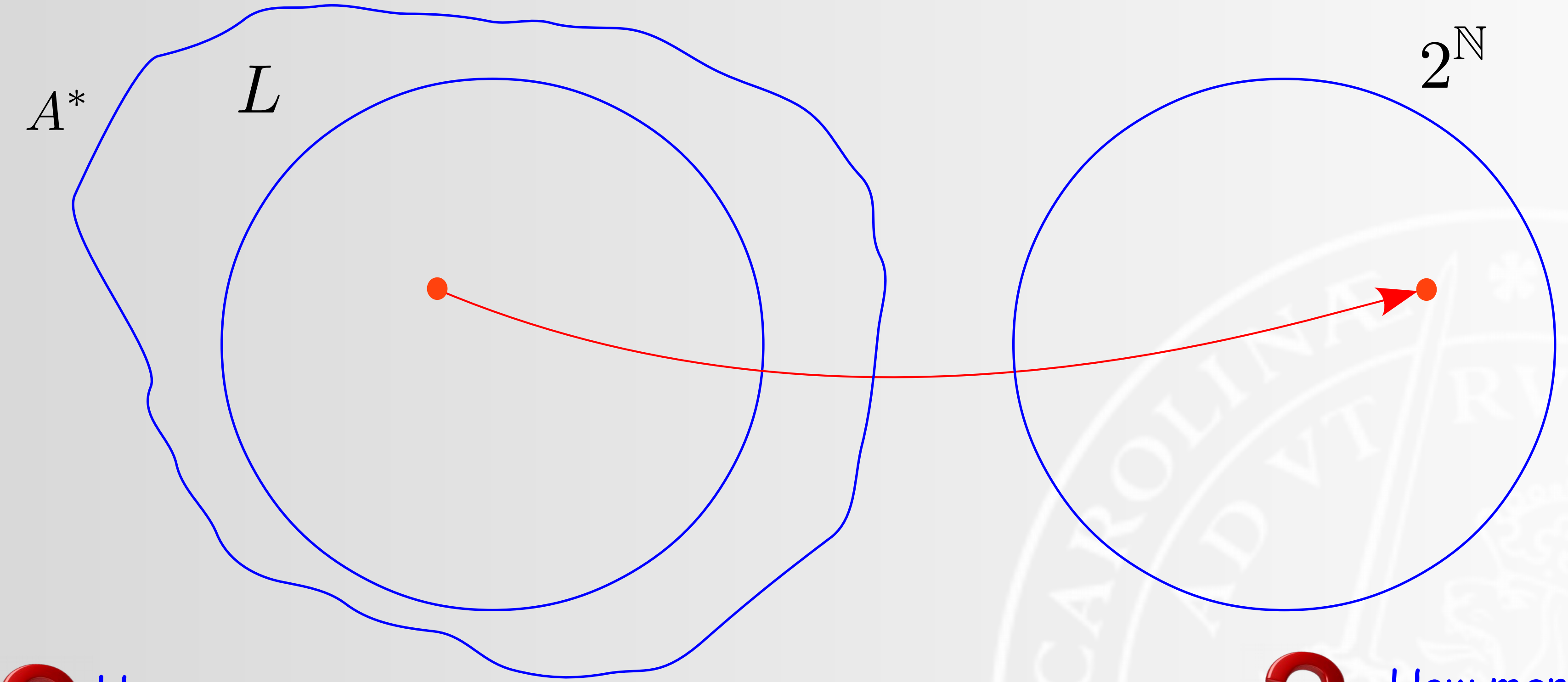
$$\aleph_0 = \#(\mathbb{N}) < \#(2^{\mathbb{N}}) = 2^{\aleph_0}$$

$$\aleph_0 < 2^{\aleph_0}$$

We discovered a new transfinite cardinal number: $\mathfrak{c} = 2^{\aleph_0}$

Proposition: It is the case that $n^{\aleph_0} = 2^{\aleph_0} = \mathfrak{c}$ for all finite $n \geq 2$

coming back to the simple problem...



How many programs at most in L?

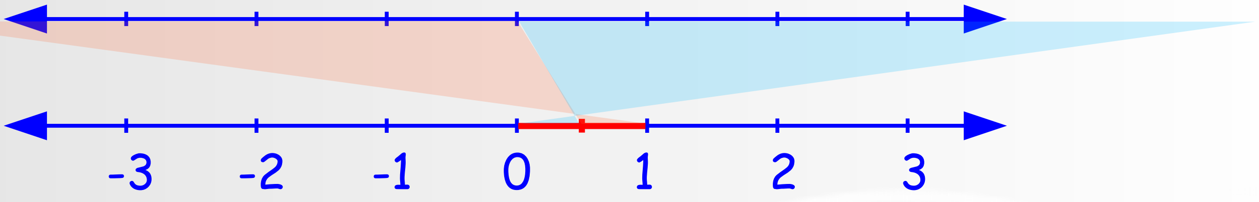
How many functions $\mathbb{N} \longrightarrow \{0,1\}$?

Conclusion?

$\mathbb{R}$: the real numbers

$$p_{\mathbb{R}} : \ ]0,1[_{\mathbb{R}} \longleftrightarrow \mathbb{R}$$

$$x \longmapsto \begin{cases} \frac{1}{x} - 2 \text{ for } x < \frac{1}{2} \\ 2 - \frac{1}{x-1} \text{ for } x \geq \frac{1}{2} \end{cases}$$



$$\mathbb{R} \sim \ ]0,1[_{\mathbb{R}}$$

So for the purposes of determining the cardinality of the real numbers, we can focus on *non-terminating* decimal sequences:

$$0.d_1 d_2 \ldots d_n \ldots$$

Non-terminating means there is no n such that all digits after d_n are 0. Otherwise, this would be the set of *all* sequences of ten symbols $\{0,1,2,3,4,5,6,7,8,9\}^{\mathbb{N}}$, with cardinality

$$10^{\aleph_0} = 2^{\aleph_0} = \mathfrak{c}$$

Even so, the cardinality of the real numbers still comes out to

$$\#(\mathbb{R}) = \mathfrak{c}$$

(proof omitted)

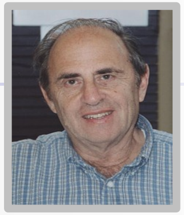
power sets

Note that for $\aleph_0$, it is the case that $\aleph_0 < 2^{\aleph_0}$, i.e. the set of natural numbers is strictly smaller than its powerset. This holds more generally:

$$\begin{array}{ll} \text{For any set } A, & \left.\begin{array}{l} \#(A) < \#(\mathcal{P}(A)) \\ C < 2^C \end{array}\right\} \text{same thing} \\ \text{For any cardinal number } C, & \end{array}$$

What does that mean for transfinite cardinal numbers?

more transfinite cardinals


Paul Cohen
1934-2007

sidebar

So far, we have encountered two transfinite cardinals:

$$\aleph_0 = \#(\mathbb{N}) \text{ and } \mathfrak{c} = 2^{\aleph_0} = \#(\mathbb{R}).$$

As we have seen, there are infinitely many transfinite cardinals.
Starting from $\aleph_0$, they are called in order

$$\aleph_0 < \aleph_1 < \aleph_2 < \ldots$$

Such that between any two $\aleph_n, \aleph_{n+1}$ there is no other cardinal number.

Where does $\mathfrak{c}$ fit in? All we know is that $\mathfrak{c} > \aleph_0$, so it's at least $\aleph_1$.
So, is $\mathfrak{c} = \aleph_1$? This is the *continuum hypothesis* (CH).

CH was shown to be *independent* of ZFC (Cohen, 1963).
Since ZFC doesn't tell us how big those alephs are, we get beths:

$$\beth_0 < \beth_1 < \beth_2 < \ldots$$

Such that $\beth_0 = \aleph_0$ and $\beth_{n+1} = 2^{\beth_n}$. At least we know that $\mathfrak{c} = \beth_1$

Note: We assume ZFC for this discussion, i.e. Zermelo-Fraenkel set theory with the axiom of choice. Do not worry about it.