GTI

Diagonalization

A. Ada, K. Sutner

Carnegie Mellon University

Fall 2017

1. Comments
    - Cardinality
    - Infinite Cardinality
    - Diagonalization

Personal Quirk 1

“Theoretical Computer Science (TCS)” sounds distracting–computers are just a small part of the story.

I prefer Theory of Computation (ToC) and will refer to that a lot.

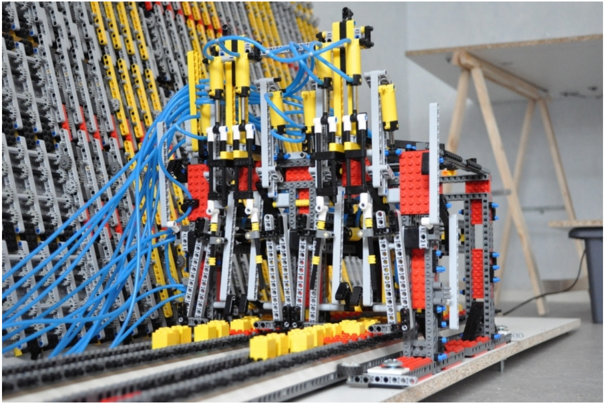
ToC:

- computability theory
- complexity theory
- proof theory
- type theory/set theory
- physical realizability

Personal Quirk 2

To my mind, the exact relationship between physics and computation is an absolutely fascinating open problem.

It is obvious that the standard laws of physics support computation (ignoring resource bounds).



There even are people (Landauer 1996) who claim

> . . . this amounts to an assertion that mathematics and computer science are a part of physics.

I think that is total nonsense, but note that Landauer was no chump: in fact, he was an excellent physicists who determined the thermodynamical cost of computation and realized that reversible computation carries no cost.

At any rate . . .

Note the caveat: “ignoring resource bounds.”

Just to be clear: it is not hard to set up computations that quickly overpower the whole (observable) physical universe. Even a simple recursion like this one will do.

$$\begin{aligned}
A(0,y) &= y^{+} \\
A(x^{+},0) &= A(x,1) \\
A(x^{+},y^{+}) &= A(x,A(x^{+},y))
\end{aligned}$$

This is the famous Ackermann function, and I don’t believe its study is part of physics.

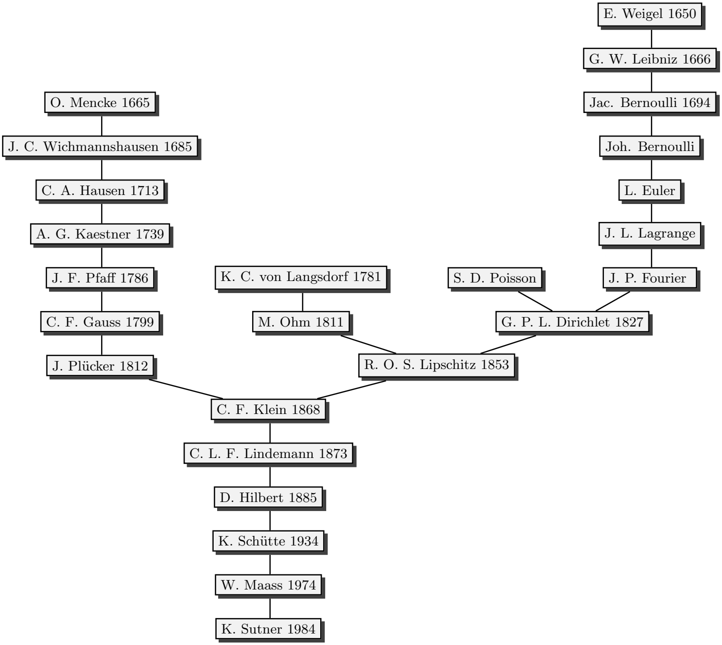
And there are much worse examples.

But the really hard problem is going in the opposite direction: no one knows how to axiomatize physics in its entirety, so one cannot **prove** that all physical processes are computable.

Hilbert was the first to realize this and posed the following problem (#6 on his list) in 1900:

> Mathematical Treatment of the Axioms of Physics.
>
> The investigations on the foundations of geometry suggest the problem: To treat in the same manner, by means of axioms, those physical sciences in which already today mathematics plays an important part; in the first rank are the theory of probabilities and mechanics.

Personal Quirk 3



E. Weigel 1650

G. W. Leibniz 1666

Jac. Bernoulli 1694

Joh. Bernoulli

L. Euler

J. L. Lagrange

J. P. Fourier

S. D. Poisson

G. P. L. Dirichlet 1827

O. Mencke 1665

J. C. Wichmannshausen 1685

C. A. Hausen 1713

A. G. Kaestner 1739

J. F. Pfaff 1786

C. F. Gauss 1799

J. Plücker 1812

K. C. von Langsdorf 1781

M. Ohm 1811

R. O. S. Lipschitz 1853

C. F. Klein 1868

C. L. F. Lindemann 1873

D. Hilbert 1885

K. Schütte 1934

W. Maass 1974

K. Sutner 1984

- Comments

2. Cardinality

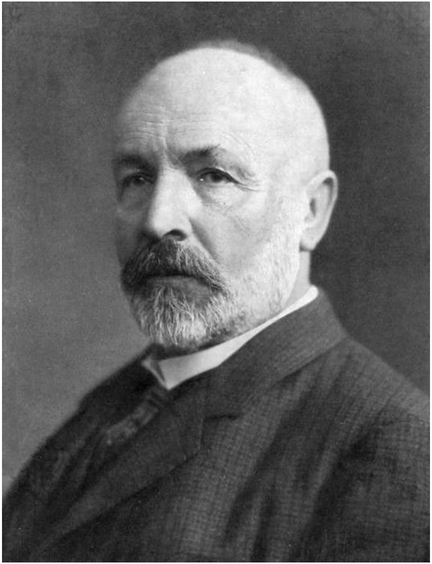
- Infinite Cardinality

- Diagonalization

Today: Diagonalization

set theory	uncountability
computability	unsolvability
complexity	hardness, separation
proof theory	Gödel incompleteness

Georg Cantor



Cantor single-handedly invented modern set theory in the late 19th century.

Incidentally, while studying Fourier transforms.

You see, not all applications are useless.

A Definition of Set

Here is a feeble first attempt at a “definition” of a set.

Definition

A set is an arbitrary collection of objects.

In the words of Cantor:

> By an “aggregate” we are to understand any collection into a whole M of definite and separate objects m of our intuition or our thought. The objects are called “elements” of M. In signs we express this thus: $M = \{m\}$.

Modern Notation

Cantor's symbolic notation is rather old-fashioned. Nowadays, and following G. Peano, one would usually write

$$M = \{ m \mid P(m) \}$$

indicating that we wish to collect all objects m that have property P into a set M. Upper/lower case letters are a feeble attempt at typing.

This set formation principle is the core of set theory. Unfortunately, in its full unrestricted form it also causes major problems.

Frege's Axioms

As part of his foundational work in logic, G. Frege developed a set theory that essentially boils down to just two axioms.

- **Extensionality**

$$x = y \;\text{ if }\; \forall\, z\, (z \in x \iff z \in y)$$

- **Formation**
  For any property $P(z)$:

$$\exists\, x\, \forall\, z\, (z \in x \iff P(z))$$

The quantifiers here all range over the collection of all sets. Note that by Extensionality the set x in Formation is unique.

Inconsistency

Russell realized that Frege's system has internal contradictions. Fixes:

- Frege changed his axioms; unfortunately causing his universe to have only one element.
- Russell invented type theory; horrible system (reducibility axiom) that no one uses.
- Zermelo-Fraenkel, von Neumann-Gödel-Bernays, Kelly-Morse: reasonable axiom systems, not terribly complicated.

Practical Advice: Simply ignore all these proplems.

"A foolish consistency is the hobgoblin of little minds" R. W. Emerson

A Century's Worth of Experience

. . . shows that these two axioms are enough to construct all of math and computer science.

This is a white lie, but more than good enough for our purposes.

Set theory provides an extremely powerful and even elegant way to organize and structure any discourse in math and computer science.

It has become the de facto gold standard: a rigorous argument is one that can be reconstructed in terms of set theory (at least in principle). Bourbaki's whole oeuvre is built on this idea, and has conquered the world of math.

Bourbaki is cilantro.

Assembly Language

As a ToC person, you can think of set theory as a universal assembly language: any mathematical concept such as integer, rational, real, series, function, integral, vector field, finite field, . . . can be interpreted as a set. With a little more work we get machines, languages, problems, complexity classes, . . .

This interpretation may be overly technical, and wreak havoc on our cherished intuitions, but it provides a rock-solid foundation: all ambiguity evaporates, all proofs are perfectly reliable (and they can be carried out by machines). But at a cost . . .

Things tend to get very formal, abstract and technical. Sometimes overly.

Charles Hermite

The impression that Cantor's memoirs produce on us is disastrous. Reading them seems to us a complete torture . . . Even acknowledging that he has opened up a new field of research, none of us is tempted to follow him. It has been impossible for us to find, among the results that can be understood, just one that possesses a *real and present interest*.

(Proved that e is transcendental.)

And Yet . . .

Bourbaki-style arguments have been the gold standard for more than half a century.

This will not change any time soon. If you want to work in ToC, you have no choice.

Cilantro.

Aside: The Problem With Definitions

Beware: When you try to come to grips with a new concept (like DFA or TM), it is entirely pointless to simply stare at the formal, set-theoretic definition. Instead, create a little table in your mind:

- intuitive meaning
- formal definition
- examples
- counterexamples
- basic results

Bad things happen to people who cling solely to formal definitions.

Rant

Some people tell you that definitions are incapable of being wrong.

That is complete nonsense.

In formal logic, definitions are indeed defined as arbitrary abbreviations.

In the real world, definitions have a clear cognitive purpose: they must help to organize your thought and your arguments.

If they don't, they are wrong.

Ask Cauchy about continuity.

Today's Challenge: The Size of a Set

We want to make sense out of the concept of the size of a set.

This seems quite straightforward for simple sets like

$$A = \{\emptyset, 42, \triangle, \square\}$$

Clearly, A has size 4.

More generally, if a set looks like

$$A = \{a_0, a_1, a_2, \ldots, a_{n-1}\}$$

then it has size n (here we assume tacitly that the enumeration does not contain any repetitions).

The Real Challenge: Infinity

So we can handle finite sets (more later).

But what should we do with infinite sets like

$$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{N} \to \mathbb{N}, \mathbb{R} \to \mathbb{R}, \Sigma^\star, \text{trees}, \text{ugraphs}, \text{TMs}, \ldots$$

Simply calling them all “infinite” is not good enough.

Cardinality

If you don't understand something, enshrine it in a preliminary definition.

Definition

The size of a set A is called its cardinality or cardinal number and written symbolically as $|A|$.

This is really a figure of speech more than a definition, it says nothing about the nature of cardinal numbers. In the words of G. Cantor:

> Every aggregate M has a definite "power," which we also call its "cardinal number."
> ... the general concept which, by means of our active faculty of thought, arises from the aggregate M when we make abstraction of the nature of its various elements m and of the order in which they are given.

The “Easy” Case

So how do we define a cardinal number?

For finite sets we clearly want to use natural numbers as cardinals. Everybody understands the naturals, right?

Awkward Question: What is a natural number?

This sounds like an inane question, but remember: we would like to be able to trace everything back to a set, at least in principle.

How should we model natural numbers as sets?

von Neumann Ordinals

John von Neumann was one of the leading mathematicians of the 20th century, and did groundbreaking work in CS.

Here is his answer. Recursively define

$$\underline{n} = \{\underline{0}, \underline{1}, \ldots, \underline{n-1}\}$$

So, we use $\emptyset$ to model zero, and, to increase a finite ordinal by one, we apply the successor function $S(x) = x \cup \{x\}$

$$\underline{n+1} = S(\underline{n}) = \underline{n} \cup \{\underline{n}\}$$

Awful Notation

n	$\underline{n}$
0	$\emptyset$
1	$\{\emptyset\}$
2	$\{\emptyset, \{\emptyset\}\}$
3	$\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$
4	$\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}\}$

Positively awful to look at, we'll stick to our standard notation and write things like 5 and 42. This is similar to the difference between assembly and Python.

But the point is that a von Neumann ordinal $\underline{n}$ is the prototype of a finite set of size n; in a sense it is the most basic set with n elements.

WTF?

Isn't this overly complicated? Why not simply represent 4 as

$$\{\{\{\{\{\}\}\}\}\}$$

Absolutely, this coding is easier to define. But it's much harder to use.

But here is the killer: von Neumann's definition carries over beautifully to infinite numbers, the simpleton approach collapses miserably.

Lining Things Up

For a set A to have finite size n means we can write a two-column table: $0, \ldots, n-1$ in column one, the elements of A in column two.

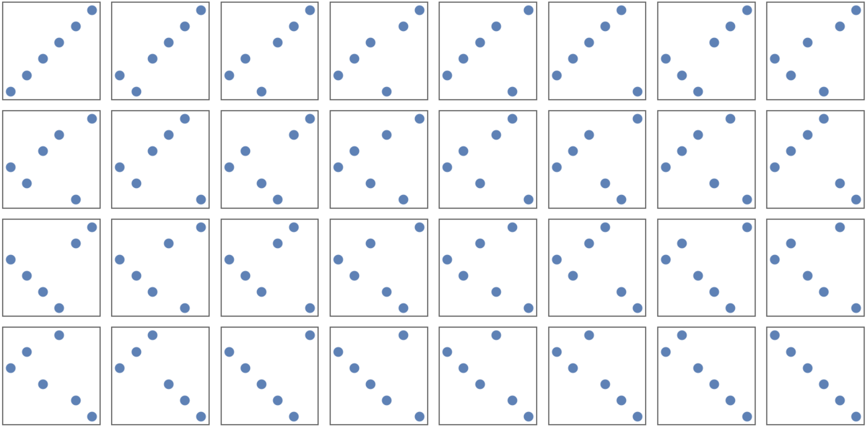
0	♡
1	♢
2	♣
3	♠

More formally, we need a bijection

$$f : \underline{n} \longleftrightarrow A$$

for some n.

Remember “Adjacent” Permutations?



How many are there? What is the cardinality of P_n?

Counting

- Show that $|P_n| \geq 2|P_{n-1}|$ and $|P_n| \leq 2|P_{n-1}|$.
- Find an (easily computable) bijection $f : \mathbf{2}^{n-1} \to P_n$.

Both require a little analysis of adjacent permutations.

Comparing Cardinality

G. Cantor suggests the following method to determine whether two sets have the same size:

> We say that two aggregates M and N are "equivalent" if it is possible to put them, by some law, in such a relation to one another that to every element of each one of them corresponds one and only one element of the other.

This idea was not new, Galileo Galilei already used it to show that there are as many squares as there are naturals.

- Comments
- Cardinality

3. Infinite Cardinality

- Diagonalization

And Infinite Sets?

Time to get serious: how do we make sense out of $|A|$ when A fails to be finite?

$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{N} \to \mathbb{N}, \mathbb{R} \to \mathbb{R}, \Sigma^\star, \text{trees}, \text{ugraphs}, \text{TMs}, \ldots$

As already mentioned, one can generalize von Neumann ordinals to the infinite case. Alas, this leads deep into to the swamp of set theory, so we try to weasel around it. Take a look at ordinals if you want to know more.

A closer look at the finite case shows that we can deal with cardinal numbers without knowing what they are . . .

A Dirty Trick

We can compare cardinalities without having to worry about the ontological status of a cardinal number.

Definition

Let A and B be two arbitrary sets.

$$|A| = |B| \iff \exists\, f \text{ bijective}\,(\, f\colon\ A \longleftrightarrow B\,)$$

$$|A| \leq |B| \iff \exists\, f \text{ injective}\,(\, f\colon\ A \longrightarrow B\,)$$

$$|A| \geq |B| \iff \exists\, f \text{ surjective}\,(\, f\colon\ A \longrightarrow B\,)$$

Sets with the same cardinality are called equipotent or equinumerous.

Basic Sanity Check

At the very least, “same-cardinality” should be an equivalence relation.

- reflexive: $I_A \colon A \longleftrightarrow A$
- symmetric: $f \colon A \longleftrightarrow B$ yields $f^{-1} \colon B \longleftrightarrow A$
- transitive: $f \colon A \longleftrightarrow B$ and $g \colon B \longleftrightarrow C$ yields $f \circ g \colon A \longleftrightarrow C$

So far, so good.

Likewise, “at-most-same-cardinality” is a pre-order (reflexive and transitive).

But it’s not a partial order: the whole point of cardinality is to compare different sets.

Comparability holds, given sufficiently strong axioms of set theory (AC).

More Sanity

Lemma

There is an injection $f : A \to B$ *if, and only if, there is a surjection* $g : B \to A$*.*

Proof.

Assume $A \neq \emptyset$ and let f be the injection. Pick $a_0 \in A$ and set

$$g(b) = \begin{cases} a & \text{if } f(a) = b, \\ a_0 & \text{if } b \notin \operatorname{rng} f. \end{cases}$$

Assume g is a surjection. Hence, for each $a \in A$, there exists an $b \in B$ such that $g(b) = a$. Pick one such b, say, b_0, and set $f(a) = b_0$.

□

Yet More Sanity

Theorem (Cantor-Schröder-Bernstein)

Suppose $f : A \to B$ and $g : B \to A$ are injective.
Then A and B have the same cardinality.

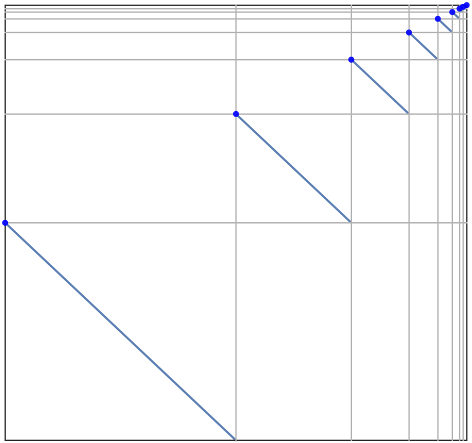
Incidentally, Cantor had no proof, Schröder had a wrong proof, Bernstein had correct proof, but Dedekind had a better proof.

At any rate, the result makes sure that, for all cardinals κ and λ,

$$\kappa \leq \lambda \text{ and } \lambda \leq \kappa \quad \text{implies} \quad \kappa = \lambda$$

We'll skip the proof of CSB, see cardinality for two proofs.

A Bijection



Constructing a bijection $[0,1] \leftrightarrow (0,1]$ directly. Try.

Countably Infinite

The naturals $\mathbb{N}$ are the most basic example of an infinite set. Surprisingly, there are lots of other sets that appear to be larger but are all equinumerous with $\mathbb{N}$;

$$\mathbb{Z},\ \mathbb{N} \times \mathbb{N},\ \mathbb{Q},\ \Sigma^\star,\ \text{finite trees, finite graphs, TMs, } \dots$$

Definition

A set A countably infinite if there is a bijection $\mathbb{N} \leftrightarrow A$.

Exercise

Show that all the sets above are indeed countably infinite.

A Basic Bijection

A very useful tool is a pairing function, an injective map $\pi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$.

There are many possibilities, the following is my favorite choice:

$$\pi(x, y) = 2^x(2y + 1)$$

For example

$$\pi(5, 27) = 32 \cdot 55 = 1760 = 110111\,00000_2$$

Using π and recursive extension to $\mathbb{N}^\star$ it is not hard to show that lots of infinite sets are countable.

Terminology

It is standard to say

set A is countable

if A is either finite or countably infinite.

Thus, a countable set can be enumerated as

$$a_0, a_1, a_2, \ldots, a_{n-1}$$

or

$$a_0, a_1, a_2, \ldots, a_{n-1}, a_n, \ldots$$

Countable sets are critical in ToC; in a sense, anything larger is automatically off limits (at least in the classical theory).

Aleph Naught

OK, but if countable sets are so important, it would be nice to know what their cardinal number is.

To first order approximation, we can define the first infinite cardinal number to be

$$\aleph_0 = \{\, \underline{n} \mid n \geq 0 \,\}$$

As written, this definition is rather circular, we basically assume the naturals to define the naturals. This can be fixed, see Dedekind chains, but we won't go there.

Cardinality Zoo

So we now have the following infinite collection of cardinal numbers:

$$0, 1, 2, \ldots, 42, \ldots, 10^{10^{10}}, \ldots, \aleph_0$$

Note that we can even do arithmetic on these numbers:

$$\aleph_0 + n = \aleph_0 + \aleph_0 = \aleph_0 \cdot \aleph_0 = \aleph_0$$

Exercise

Figure out what the last comment means. Think about $\mathbb{Z}$, $\mathbb{Q}$ and the like.

And Beyond?

Natural Question: Are there sets that are not countable?

As we will see, the answer is emphatically YES: it makes perfect sense to talk about

$$\aleph_1, \aleph_2, \ldots, \aleph_{10^{10}}, \ldots, \aleph_{\aleph_0}, \ldots, \aleph_{\aleph_{\aleph_0}}, \ldots$$

The sequence of cardinals is itself wildly infinite and leads straight into the abyss (aka the math department).

Relax, though, all these higher cardinalities play hardly any role in the life of computer scientist. The reason we talk about them is the proof technique that is used to produce them: diagonalization.

- Comments

- Cardinality

- Infinite Cardinality

4. Diagonalization

OK, so here is a definition that just might be vacuous:

Definition

A set is uncountable if it is neither finite nor countably infinite.

We need to show that some uncountable set exists.

And it would be nice to come up with a well-known set that turns out to uncountable.

Cantor's Theorems

Theorem (Cantor 1874/1891)

The set of real numbers, $\mathbb{R}$, is not countable.

Theorem (Cantor 1891)

For any set A, the cardinality of $\mathfrak{P}(A)$ is greater than the cardinality of A.

Say What?

Cantor's first theorem shows that there are at least two levels of infinity, and that they play a central role in calculus: the continuum has to be larger than the naturals.

The second theorem is a bit harder to swallow: it shows that there are in fact infinitely many levels of infinity:

$$|\mathbb{N}| < |\mathfrak{P}(\mathbb{N})| < |\mathfrak{P}(\mathfrak{P}((\mathbb{N}))| < |\mathfrak{P}(\mathfrak{P}(\mathfrak{P}(\mathbb{N})))| < \dots$$

Even this infinite chain is misleading: cardinals form the backbone of the universe of sets, the collection of all cardinals is as large as the universe itself.

Sequences

Standard notation for sequences over some set X:

finite	List(X), $X^\star$, $X^{<\omega}$
infinite	X^ω, $X^\mathbb{N}$, $\mathbb{N} \to X$

Cantor: Diagonalization

Warm-up: The number of binary sequences of length n is larger than n.

Yes, yes, we can do this quite directly by counting, but ordinary counting does not work for infinite sets; we need a different approach.

We will prove something more constructive:

Given n binary sequences s_i, $i < n$, of length n, there is a binary sequence d of length n that differs from all of them.

Diagonalization

Here goes: given the s_i, define d by

$$d(i) = 1 - s_i(i)$$

for all $i < n$.

Then d differs from all the s_i in at least one bit, so $d \neq s_i$ for all $i < n$.

Note that d is obtained by mucking with the diagonal sequence $s_i(i)$.

Matrix View

We get d by flipping each bit along the diagonal of a matrix. Hence the resulting sequence cannot be a row in the matrix.

$$\begin{array}{ccccc}
\mathbf{s_0(0)} & s_0(1) & s_0(2) & \ldots & s_0(n-1) \\
s_1(0) & \mathbf{s_1(1)} & s_1(2) & \ldots & s_1(n-1) \\
s_2(0) & s_2(1) & \mathbf{s_2(2)} & \ldots & s_2(n-1) \\
\vdots & & & \ddots & \vdots \\
s_{n-1}(0) & s_{n-1}(1) & s_{n-1}(2) & \ldots & \mathbf{s_{n-1}(n-1)}
\end{array}$$

In general, it does not matter how we change the element $s_i(i)$ in d, it just has to be different. With bits there is only one choice, of course.

The Key Insight

This also works for infinite sequences.

Simply replace $i < n$ by $i < \aleph_0$ and everything works just fine for infinite sequences $s_i \in \mathbf{2}^{\mathbb{N}}$.

Claim

There are uncountably many binary sequences: $\mathbf{2}^{\mathbb{N}}$ is uncountable.

Corollary

$\mathfrak{P}(\mathbb{N})$ *is uncountable.*

The Real Thing

We want to show that $\mathbb{R}$ is uncountable. There are two main approaches:

- Modify the diagonalization argument to work for $\mathbb{R}$.
- Find an injection $\mathbf{2}^{\mathbb{N}} \to \mathbb{R}$.

The second approach is more elegant, but we want to get some more exercise in diagonalization, so let's do the first.

Exercise

Take the second approach to show that $\mathbb{R}$ is uncountable.

Diagonalization for $\mathbb{R}$

Let $D = \{0, 1, \ldots, 9\}$ be the decimal digits. Diagonalization easily shows that $D^{\mathbb{N}}$ is uncountable.

But there is a problem: we want to interpret $x \in D^{\mathbb{N}}$ as a decimal expansion

$$0 \,.\, x_0 x_1 x_2 \ldots x_n x_{n+1} \ldots$$

Thus, the numerical value of x is

$$\mathsf{val}(x) = \sum_{i \geq 0} x_i 10^{-i-1}$$

Clearly $0 \leq \mathsf{val}(x) \leq 1$ and the map is surjective.

Alas: it is not injective because of trailing 9's:

$$0.500000\ldots = 1/2 = 0.499999\ldots$$

Killing 9's

Let's agree to consider only sequences that have no trailing infinite block of 9's, so we are now looking at some set of sequences $D_0^{\mathbb{N}} \subsetneq D^{\mathbb{N}}$.

The value map is now injective, and has range $[0, 1)$.

Assume again that there is a sequence of digit sequences s_i, $i \geq 0$, that enumerates $D_0^{\mathbb{N}}$. Then the diagonal sequence

$$d(i) = \begin{cases} 3 & \text{if } s_i(i) = 2, \\ 2 & \text{otherwise.} \end{cases}$$

is in $D_0^{\mathbb{N}}$, but different from all the s_i. Contradiction.

Cantor II

The next task is to show that the cardinality of $\mathfrak{P}(A)$ is strictly greater than the cardinality of A, for any set A.

There is a trivial injection from A to $\mathfrak{P}(A)$: $a \mapsto \{a\}$.

So suppose there is a surjection $f : A \to \mathfrak{P}(A)$. Think of a as a "name" for the set $f(a)$ and define

$$B = \{\, a \in A \mid a \notin f(a) \,\} \subseteq A.$$

Since f is surjective, B must have a name: $B = f(b)$ for some $b \in A$.

But then $b \in B$ implies $b \notin B$, and conversely; contradiction.

Cantor vs. Russell

Note that Cantor's construction is very similar to Russell's paradox, surprisingly Cantor never made the transition.

$$S = \{\, x \mid x \notin x \,\}$$

$$B = \{\, a \in A \mid a \notin f(a) \,\}$$

The existence of S is contradictory, but can be proved from Frege's axioms (though presumably not in Zermelo-Fraenkel set theory).

But there is nothing wrong with B, it just shows that f cannot be surjective (and can be proved to exist in Zermelo-Fraenkel set theory).

More on Diagonalization

Diagonalization is a key technique in computability and complexity theory.

Next lecture we will see a proof of the undecidability of the Halting Problem, and it is based on exactly the same idea.

The only difference is that we will be dealing with computable maps, rather that set-theoretic ones.

So Cantor inadvertently also provided a fundamental technique for computability theory.

Definition: Cardinality

intuition

formal def (this is hard in the uncountable case, ignore)

examples

counterexpl

results