

## More on comparing sizes of sets

Recall: Sets  $A$  and  $B$  are the same size if there is a bijection from  $A$  to  $B$ .

A set is countable if it is finite or is the same size as  $\mathcal{N}$ .

Some methods for showing that a set is countable.

- ▶ Every subset of a countable set is countable.
- ▶ Set  $S$  is countable iff there is an injection from  $S$  to some countable set.
- ▶ Set  $S$  is countable iff there is a surjection from some countable set to  $S$ .
- ▶ The image of a countable set is countable.
- ▶ A countable union of countable sets is countable.

Some countable sets:

- ▶  $\mathcal{N} \times \mathcal{N}$ ,
- ▶ rational numbers,
- ▶  $A^*$  for any alphabet  $A$ ,
- ▶  $\text{lists}(\mathcal{N})$ ,
- ▶ the set of finite subsets of  $\mathcal{N}$ .



Since the sequence below is arbitrary, if we can show that not all real numbers from 0 to 1 are represented in this sequence, then we can conclude that **no enumeration of all real numbers exists**.

$d_{0,0}$	$d_{0,1}$	$d_{0,2}$	$d_{0,3}$	$d_{0,4}$	$d_{0,5}$	$d_{0,6}$	$d_{0,7}$	...
$d_{1,0}$	$d_{1,1}$	$d_{1,2}$	$d_{1,3}$	$d_{1,4}$	$d_{1,5}$	$d_{1,6}$	$d_{1,7}$	...
$d_{2,0}$	$d_{2,1}$	$d_{2,2}$	$d_{2,3}$	$d_{2,4}$	$d_{2,5}$	$d_{2,6}$	$d_{2,7}$	...
$d_{3,0}$	$d_{3,1}$	$d_{3,2}$	$d_{3,3}$	$d_{3,4}$	$d_{3,5}$	$d_{3,6}$	$d_{3,7}$	...
$d_{4,0}$	$d_{4,1}$	$d_{4,2}$	$d_{4,3}$	$d_{4,4}$	$d_{4,5}$	$d_{4,6}$	$d_{4,7}$	...
$d_{5,0}$	$d_{5,1}$	$d_{5,2}$	$d_{5,3}$	$d_{5,4}$	$d_{5,5}$	$d_{5,6}$	$d_{5,7}$	...
$d_{6,0}$	$d_{6,1}$	$d_{6,2}$	$d_{6,3}$	$d_{6,4}$	$d_{6,5}$	$d_{6,6}$	$d_{6,7}$	...
$d_{7,0}$	$d_{7,1}$	$d_{7,2}$	$d_{7,3}$	$d_{7,4}$	$d_{7,5}$	$d_{7,6}$	$d_{7,7}$	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

OK, so how can we show that this is not an enumeration of *all* real numbers from 0 to 1? We should produce an infinite sequence of digits 0–9 that does not occur here!

And this is where the name “diagonalization” comes from. It is easy to produce such a sequence — it need only differ from each of the sequences above at one digit, right?

We produce this new sequence by taking each of the digits along the diagonal and altering it! Thus, the new sequence will differ from  $n$ th sequence in the enumeration at the  $n$ th digit...

$d_{0,0}$	$d_{0,1}$	$d_{0,2}$	$d_{0,3}$	$d_{0,4}$	$d_{0,5}$	$d_{0,6}$	$d_{0,7}$	$\dots$
$d_{1,0}$	$d_{1,1}$	$d_{1,2}$	$d_{1,3}$	$d_{1,4}$	$d_{1,5}$	$d_{1,6}$	$d_{1,7}$	$\dots$
$d_{2,0}$	$d_{2,1}$	$d_{2,2}$	$d_{2,3}$	$d_{2,4}$	$d_{2,5}$	$d_{2,6}$	$d_{2,7}$	$\dots$
$d_{3,0}$	$d_{3,1}$	$d_{3,2}$	$d_{3,3}$	$d_{3,4}$	$d_{3,5}$	$d_{3,6}$	$d_{3,7}$	$\dots$
$d_{4,0}$	$d_{4,1}$	$d_{4,2}$	$d_{4,3}$	$d_{4,4}$	$d_{4,5}$	$d_{4,6}$	$d_{4,7}$	$\dots$
$d_{5,0}$	$d_{5,1}$	$d_{5,2}$	$d_{5,3}$	$d_{5,4}$	$d_{5,5}$	$d_{5,6}$	$d_{5,7}$	$\dots$
$d_{6,0}$	$d_{6,1}$	$d_{6,2}$	$d_{6,3}$	$d_{6,4}$	$d_{6,5}$	$d_{6,6}$	$d_{6,7}$	$\dots$
$d_{7,0}$	$d_{7,1}$	$d_{7,2}$	$d_{7,3}$	$d_{7,4}$	$d_{7,5}$	$d_{7,6}$	$d_{7,7}$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

For all  $n \in \mathcal{N}$ ,

$$d_n = \begin{cases} 1 & , \text{ if } d_{n,n} > 1 \\ 2 & , \text{ otherwise} \end{cases}$$

It is easy to see that every one of the included sequences

$$d_{n,0}, d_{n,1}, d_{n,2}, \dots, d_{n,n-1}, d_{n,n}, d_{n,n+1}, \dots$$

is different from the defined sequence

$$d_0, d_1, d_2, \dots, d_{n-1}, d_n, d_{n+1}, \dots$$

Just notice that  $d_{n,n} \neq d_n$ , for all  $n \in \mathcal{N}$ .

It is a little harder to see that every one of the included sequences represents a different real number than that represented by the defined sequence.

(The problem:  $.099\dots = .1000\dots$ )

The key: we don't replace 0 with 9 or replace 9 with 0.

OK, that's the basic idea of diagonalization, but it seems to apply only to the problem of showing that a set is uncountably infinite. (By showing that the set is not enumerable.)

We can simplify things, now that we have seen the idea, and in doing so make the method more generally applicable. . .

## Simplifying the prior argument

Notice: A countably infinite sequence of elements from a set  $S$  can be understood as a function from  $\mathcal{N}$  to  $S$ .

So, for all  $n \in \mathcal{N}$ , let  $f_n$  be an arbitrary function from  $\mathcal{N}$  to  $\{0, \dots, 9\}$ .

Now, each of these countably many, arbitrarily chosen functions  $f_n$  represents a countable sequence of digits 0 through 9.

We will show that at least one function from  $\mathcal{N}$  to  $\{0, \dots, 9\}$  is not among

$$f_0, f_1 \dots$$

Take  $f : \mathcal{N} \rightarrow \{0, \dots, 9\}$  s.t. for all  $n \in \mathcal{N}$ ,

$$f(n) = \begin{cases} 1 & , \text{ if } f_n(n) \neq 1 \\ 0 & , \text{ otherwise} \end{cases}$$

Then, for all  $n \in \mathcal{N}$ ,

$$f \neq f_n$$

since

$$f(n) \neq f_n(n).$$

## Further simplifying the prior argument

Notice that in the prior argument, we represent each countably infinite sequence of digits as a function from  $\mathcal{N}$ , but we did not do the same for the arbitrary countably infinite sequence

$$f_0, f_1 \dots$$

of such functions. Let's do so this time...

Take an arbitrary function

$$g : \mathcal{N} \rightarrow (\mathcal{N} \rightarrow \{0, \dots, 9\}).$$

(So  $g$  maps each natural number  $n$  to a function from  $\mathcal{N}$  to  $\{0, \dots, 9\}$ . That is,  $g$  represents an arbitrary countably infinite sequence of countably infinite sequences of digits 0 through 9.)

To show that  $g$  is not surjective, take  $f : \mathcal{N} \rightarrow \{0, \dots, 9\}$  s.t. for all  $n \in \mathcal{N}$ ,

$$f(n) = \begin{cases} 1 & , \text{ if } g(n)(n) \neq 1 \\ 0 & , \text{ otherwise} \end{cases}$$

Notice that for all  $n \in \mathcal{N}$ ,  $g(n)(n) \neq f(n)$ , and so  $g(n) \neq f$ .

Since  $g$  was an arbitrary function from  $\mathcal{N}$  to  $\mathcal{N} \rightarrow \{0, \dots, 9\}$ , we can conclude that no function of this type is surjective, and so there is no bijection from  $\mathcal{N}$  to  $\mathcal{N} \rightarrow \{0, \dots, 9\}$ . That is, the sets are not the same size.

## Example: no surjection from $\mathcal{N}$ to $\text{power}(\mathcal{N})$

Let's use this approach to show there is no surjection from  $\mathcal{N}$  to  $\text{power}(\mathcal{N})$ ...

Take an arbitrary function  $g : \mathcal{N} \rightarrow \text{power}(\mathcal{N})$ .

We want to use diagonalization to show that  $g$  is not surjective. So what is the idea? We will *use*  $g$  to construct an element of  $\text{power}(\mathcal{N})$  that does not belong to  $\text{range}(g)$ .

We need a set  $D$  of natural numbers s.t. for all  $n \in \mathcal{N}$ ,

$$D \neq g(n).$$

That is, for each  $n \in \mathcal{N}$ , there must be some  $m \in \mathcal{N}$  s.t.

$$m \in D \quad \text{iff} \quad m \notin g(n).$$

That may seem hard, but it turns out not to be.

Take  $D$  s.t. for all  $n \in \mathcal{N}$ ,

$$n \in D \quad \text{iff} \quad n \notin g(n).$$

That is,

$$D = \{ n \mid n \in \mathcal{N}, n \notin g(n) \}.$$

Notice that, for all  $n \in \mathcal{N}$ ,  $g(n) \neq D$ . (Do you see why?)

## No surjection from $A$ to $\text{power}(A)$ !

The prior example still involved showing uncountability. But by talking about surjections instead of enumerations, we have shown how to generalize the diagonalization method.

For example, essentially the same argument as before can show that  
**no** set is the same size as its powerset!

We'll show there is no surjection from  $A$  to  $\text{power}(A)$ . (So there is no bijection, so they are not the same size.)

Take an arbitrary  $g : A \rightarrow \text{power}(A)$ . (What is our goal?)

We want a subset  $D$  of  $A$  s.t. for all  $x \in A$ ,

$$D \neq g(x).$$

Take

$$D = \{x \mid x \in A, x \notin g(x)\}.$$

Notice that, for all  $x \in A$ ,  $g(x) \neq D$ . (Do you see why?)

Claim: This result implies that there are infinitely many sizes of infinite sets.  
(How does it imply this?)

## Not everything is computable

Here's the idea. . . Let's consider algorithms, or programs, that take as input a binary string and return a binary string as output. So each algorithm "encodes" a function from  $\{0, 1\}^*$  to  $\{0, 1\}^*$ .

The problem? There are too many such functions.

Every algorithm, or program, can be represented as a binary string. (Or it can be represented as a string over some alphabet, and then translated to binary, right?)

But  $\{0, 1\}^*$  is countable, whereas the set of functions from  $\{0, 1\}^*$  to  $\{0, 1\}^*$  is not countable.

Idea: Let  $F$  be the set of functions from  $\{0, 1\}^*$  to  $\{0, 1\}^*$ . Show there is no surjection from  $\mathcal{N}$  to  $F$ . To do this, take an arbitrary  $g : \mathcal{N} \rightarrow F$ , and construct a "diagonal" function  $f : \{0, 1\}^* \rightarrow \{0, 1\}^*$ , so that  $f \notin \text{range}(g)$ . That is, construct  $f$  so that, for all  $n \in \mathcal{N}$ ,  $g(n) \neq f$ . That is, construct  $f$  so that, for all  $n \in \mathcal{N}$ , there is an  $x \in \{0, 1\}^*$  s.t.  $g(n)(x) \neq f(x)$ . Here's a way to do it: Let  $s$  be a bijection from  $\mathcal{N}$  to  $\{0, 1\}^*$ . (Such an  $s$  exists, since  $\{0, 1\}^*$  is countably infinite.) For all  $n \in \mathcal{N}$ , let

$$f(s(n)) = \begin{cases} \Lambda & , \text{ if } g(n)(s(n)) \neq \Lambda \\ 0 & , \text{ otherwise.} \end{cases}$$