

## 3.3 Diagonal Arguments: Uncountable Sets and Cantor's Theorem

Every infinite set met so far has turned out to be countable:  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$ , finite products of countable sets, countable unions of them (Theorem 3.2.3), and even the algebraic numbers (Theorem 3.2.4). Is every infinite set countable? This section shows that the answer is no, by one argument, which starts from the game of Section 1.5 and is then used three times. Given any list, it builds an object that the list misses. The argument proves that the set of 0–1 sequences is uncountable, that no set can be mapped onto the set of its subsets (Cantor's theorem), and, with decimals, that  $\mathbb{R}$  is uncountable.

### 3.3.1 Back to the five strings

Section 1.5 left a game open: one player lists five strings of five symbols 0 or 1, and the other must write a string that is not on the list. Searching all 32 strings works; a rule is better.

**Example. Back to the five strings**

Find a rule that, from any list  $s_1, \dots, s_5$  of strings of five symbols 0 or 1, builds a string not on the list. Run it on 00000, 11111, 01010, 10101, 00111, and prove that it always works.

*Solution.* Let the new string  $g$  have, as its  $k$ -th symbol, the flip of the  $k$ -th symbol of the  $k$ -th string:

$$g(k) := 1 - s_k(k) \quad (k = 1, \dots, 5)$$

(strings are functions  $\{1, \dots, 5\} \rightarrow \{0, 1\}$ , Section 1.5). On the given list, the symbols  $s_1(1), \dots, s_5(5)$  are 0, 1, 0, 0, 1, so  $g = 10110$ . It differs from 00000 in place 1, from 11111 in place 2, from 01010 in place 3, from 10101 in place 4 and from 00111 in place 5.

For any list: for each  $k$ ,  $g(k) \neq s_k(k)$ , since  $1 - b \neq b$  for  $b \in \{0, 1\}$ . So  $g \neq s_k$  (Proposition 1.5.1), for  $k = 1, \dots, 5$ . The  $k$ -th step looks only at the  $k$ -th string, in its own place  $k$ .

The rule never used the number 32. It used one place for each string on the list. Give the strings infinitely many places, indexed by  $\mathbb{N}$ , and let the list be infinite too, and the same rule still applies.

*Quick check.* Can the rule handle a list of six strings of five symbols? Does that contradict anything?

*Solution.* No. The sixth string has no sixth place to be flipped in. Nothing is contradicted: a list of 32 strings may contain every string. The rule needs as many places as strings, and in the next subsection both are indexed by  $\mathbb{N}$ .

### 3.3.2 A sequence that is not on the list

Recall from Section 1.5 that  $\{0, 1\}^{\mathbb{N}}$  is the set of all functions  $\mathbb{N} \rightarrow \{0, 1\}$ : the sequences of zeros and ones, indexed from 0. These are the strings of the game with places indexed by  $\mathbb{N}$ . It is infinite, and it looks no bigger than  $\mathbb{N} \times \mathbb{N}$ , which is countable. To show that it is **not** countable we must show, by Definition 3.1.1, that it is non-empty and that **no** function  $f: \mathbb{N} \rightarrow \{0, 1\}^{\mathbb{N}}$  is onto. So we take an arbitrary list  $f(0), f(1), f(2), \dots$  of sequences and look for a sequence that is not on it.

Write the list as a table: row  $n$  is the sequence  $f(n)$ , and the entry in row  $n$  and column  $k$  is the number  $f(n)(k)$ . Here is the start of one such table, with the entries in places  $(0, 0), (1, 1), (2, 2), (3, 3)$  boxed:

|          |                                                 |                                                 |                                                 |                                                 |   |     |
|----------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|---|-----|
| place    | 0                                               | 1                                               | 2                                               | 3                                               | 4 | ... |
| $f(0)$   | <span style="border: 1px solid black;">0</span> | 1                                               | 0                                               | 1                                               | 0 | ... |
| $f(1)$   | 1                                               | <span style="border: 1px solid black;">1</span> | 1                                               | 1                                               | 1 | ... |
| $f(2)$   | 0                                               | 0                                               | <span style="border: 1px solid black;">0</span> | 0                                               | 0 | ... |
| $f(3)$   | 1                                               | 0                                               | 0                                               | <span style="border: 1px solid black;">1</span> | 1 | ... |
| $\vdots$ |                                                 |                                                 |                                                 |                                                 |   |     |
| $g$      | 1                                               | 0                                               | 1                                               | 0                                               |   | ... |

The boxed entries run down the **diagonal** of the table. The bottom row  $g$  is the diagonal with every entry flipped:  $g(n) = 1 - f(n)(n)$ . It cannot be row 0, because it differs from  $f(0)$  in place 0; it cannot be row 1, because it differs from  $f(1)$  in place 1; and so on down the table. Nothing in this reasoning used the particular entries.

#### Theorem 3.3.1. The 0–1 Sequences Are Uncountable

The set  $\{0, 1\}^{\mathbb{N}}$  of all sequences of zeros and ones is uncountable.

*Idea* (Scratch work). By Definition 3.1.1, “uncountable” has two parts: the set is non-empty, and no map from  $\mathbb{N}$  onto it exists. The second is a statement about **every**  $f$ , so take  $f$  arbitrary and look for a sequence it misses. As givens and goal:

| Givens                                                                          | Goal                                                                |
|---------------------------------------------------------------------------------|---------------------------------------------------------------------|
| $f: \mathbb{N} \rightarrow \{0, 1\}^{\mathbb{N}}$ arbitrary                     | $f$ is not onto; and $\{0, 1\}^{\mathbb{N}} \neq \emptyset$         |
| —                                                                               | some $g \in \{0, 1\}^{\mathbb{N}}$ with $g \neq f(n)$ for every $n$ |
| $g \neq f(n)$ as soon as $g$ and $f(n)$ differ at one place (Proposition 1.5.1) | pick the place for each $n$ : place $n$ , as in the game            |
| gap: $g(n) \neq f(n)(n)$ for every $n$                                          | closed by $g(n) := 1 - f(n)(n)$                                     |

The proof writes this forwards.

*Proof.* Let  $f: \mathbb{N} \rightarrow \{0, 1\}^{\mathbb{N}}$  be any function. Define  $g: \mathbb{N} \rightarrow \{0, 1\}$  by

$$g(n) := 1 - f(n)(n) \quad (n \in \mathbb{N});$$

its values lie in  $\{0, 1\}$ , since  $1 - 0 = 1$  and  $1 - 1 = 0$ . So  $g \in \{0, 1\}^{\mathbb{N}}$ . For every  $n \in \mathbb{N}$ ,  $g(n) \neq f(n)(n)$ , since no  $b \in \{0, 1\}$  has  $1 - b = b$ . Hence  $g \neq f(n)$ : if  $g = f(n)$ , then  $g(n) = f(n)(n)$ , which is false; the two functions differ at place  $n$  [=E,  $\neg$ I; Proposition 1.5.1]. So  $g$  is not a value of  $f$ , and  $f$  is not surjective. Since  $f$  was arbitrary, no function  $\mathbb{N} \rightarrow \{0, 1\}^{\mathbb{N}}$  is surjective. The set  $\{0, 1\}^{\mathbb{N}}$  is non-empty, since it contains the zero sequence. By Definition 3.1.1,  $\{0, 1\}^{\mathbb{N}}$  is not countable. ■

This is the move of the section.

### Example. Method: the diagonal

**Template.** To show that no map from  $I$  onto  $X$  exists (for “uncountable”:  $I = \mathbb{N}$ , and check  $X \neq \emptyset$ ), let  $f: I \rightarrow X$  be arbitrary. Build  $d \in X$  whose entry at  $i$  differs from the entry of  $f(i)$  at  $i$ , for every  $i \in I$ . Then  $d \neq f(i)$  for each  $i$ , so  $f$  is not onto. Since  $f$  was arbitrary, no such map exists.

**When to use it.** The elements of  $X$  are themselves functions or subsets on the index set of the list: sequences, digit strings, subsets of  $I$ .

**The usual failure.** A  $d$  that falls outside  $X$  (a decimal ending in 9s, met below; a sequence of the wrong kind). A difference at place  $i$  that is not a real difference (two expansions of one number). Running the construction on one list only (the warning below).

*Quick check.* Would the unflipped diagonal,  $g(n) := f(n)(n)$ , do? Test it on the list whose every entry is the zero sequence.

*Solution.* No. There  $g$  is the zero sequence, which is  $f(0)$ . The new object must differ from the  $n$ -th entry at place  $n$ , for every  $n$ .

*Quick check.* Run the proof on the list  $f(n) := \mathbf{1}_{\{n\}}$ , the sequence with a 1 in place  $n$  and 0 elsewhere. What is  $g$ ? Is it on the list?

*Solution.* Here  $f(n)(n) = \mathbf{1}_{\{n\}}(n) = 1$  for every  $n$ , so  $g(n) = 1 - 1 = 0$  for every  $n$ :  $g$  is the zero sequence. It is indeed not on the list, since every  $f(n)$  has a 1 in place  $n$ .

**Warning.** The diagonal object depends on the list. “The diagonal sequence” is not one sequence: each list has its own, and a different list may well contain the  $g$  built from

this one (the list  $g, f(0), f(1), \dots$  does). And the conclusion “uncountable” needs the list to be **arbitrary**. Running the construction on one list, or on a few, shows only that those lists miss something; concluding “so no list works” from them is the error (N1) of Logic and Proof, Example (What is not allowed). The proof above starts from “let  $f$  be **any** function”, and that is what makes its conclusion universal.

### Example. A function that outgrows a list

Let  $f_0, f_1, f_2, \dots$  be any list of functions  $\mathbb{N} \rightarrow \mathbb{N}$ . Find  $g: \mathbb{N} \rightarrow \mathbb{N}$  such that, for every  $k \in \mathbb{N}$ ,  $g(n) > f_k(n)$  for every  $n \geq k$ .

*Solution. Idea.* The flipped diagonal beats  $f_n$  at place  $n$  only. To beat  $f_k$  from place  $k$  on, let place  $n$  beat  $f_0, \dots, f_n$  at once: finitely many, so a maximum exists.

Let

$$g(n) := \max\{f_0(n), \dots, f_n(n)\} + 1 \quad (n \in \mathbb{N}).$$

The set is  $\{x_1, \dots, x_{n+1}\}$  with  $x_{k+1} := f_k(n)$ , so the maximum exists by Proposition 2.3.8; it is one of the  $x_j$ , so  $g(n) \in \mathbb{N}$ . For  $k \in \mathbb{N}$  and  $n \geq k$ ,  $f_k(n)$  is one of the numbers in the set, so  $f_k(n) \leq \max\{f_0(n), \dots, f_n(n)\} < g(n)$ . The example asks more of  $g$  than a diagonal does, and each step still looks at finitely many entries at one place.

## 3.3.3 Back to Russell: Cantor's theorem

Section 1.2 proved Russell's theorem (Theorem 1.2.4): the set  $R_A$  disagrees with each candidate  $x \in A$  about  $x$  itself, so it is not in  $A$ . Here the same disagreement becomes a theorem about sizes.

The table had a row for each  $n \in \mathbb{N}$  and a column for each place  $k \in \mathbb{N}$ . Nothing about the diagonal needs the rows and the columns to be indexed by  $\mathbb{N}$ : it needs only that they are indexed by the **same** set. A subset  $S$  of a set  $X$  is a row with one entry for each  $x \in X$ , “in” or “out”, the values of  $\mathbf{1}_S$ . So the argument applies to any set and the set of its subsets.

### Theorem 3.3.2. Cantor's Theorem

Let  $X$  be a set. Then there is **no** surjection  $X \rightarrow \mathcal{P}(X)$ .

*Try it first (ten minutes): given a function  $f: X \rightarrow \mathcal{P}(X)$ , which subset of  $X$  could be missed by  $f$ ? Try  $X = \{1, 2, 3\}$  with a function of your own first.*

*Idea.* The first attempts look for a subset that is “too big” or “too small” to be a value, such as  $X$  itself or  $\emptyset$ , and fail:  $f$  may well take both values. Size is the wrong idea. A set  $D$  differs from  $f(x)$  as soon as the two disagree about **one** element, and the only element we know anything about in connection with  $f(x)$  is  $x$  itself. So we make  $D$  disagree with  $f(x)$  at  $x$ : put  $x$  into  $D$  exactly when  $x \notin f(x)$ . This is the diagonal of the table, flipped. It is the method of [Example \(Method: the diagonal\)](#), with the rows and the places both indexed by  $X$ .

*Before reading on: where is surjectivity used?*

*Proof.* Let  $f: X \rightarrow \mathcal{P}(X)$  be a function, and let

$$D := \{x \in X : x \notin f(x)\},$$

a subset of  $X$  by Definition 1.2.3, so  $D \in \mathcal{P}(X)$  (Definition 1.3.2). Suppose  $f$  is surjective. Then there is  $d \in X$  with  $f(d) = D$  (Definition 1.5.4). If  $d \in D$ , then  $d \notin f(d)$  by the definition of  $D$ , that is,  $d \notin D$ , a contradiction. Hence  $d \notin D$ . Then  $d \notin f(d)$ , and since  $d \in X$ , the definition of  $D$  gives  $d \in D$ , a contradiction again. So  $f$  is not surjective. This proves the theorem. ■

Surjectivity was used once, to obtain  $d$ ; everything else holds for every function  $f$ . Read without the supposition, the proof shows that  $D \neq f(x)$  for **every**  $x \in X$ : the two sets disagree about  $x$ . Three consequences follow.

- (i) *Russell's theorem is the same argument.* In Theorem 1.2.4, the set  $R_A = \{x \in A : x \notin x\}$  is the set  $D$  built from the rule  $f(x) := x$ : each  $x \in A$  is tested against itself. The proof that  $D$  differs from every  $f(x)$  becomes the proof that  $R_A$  differs from every element of  $A$ , that is,  $R_A \notin A$ . (The rule  $x \mapsto x$  need not map into  $\mathcal{P}(A)$ , but the argument never used that  $f(x) \subseteq X$ , only the test " $x \in f(x)$ ".)
- (ii) *The power set is strictly bigger.* The map  $x \mapsto \{x\}$  is an injection  $X \rightarrow \mathcal{P}(X)$ , since  $\{x\} = \{y\}$  gives  $x = y$ . So  $\mathcal{P}(X)$  has at least as many elements as  $X$ , in the sense that  $X$  is in bijection with a subset of it, and by the theorem it is not equinumerous with  $X$ , since a bijection would be a surjection. Section 3.4 names this relation:  $X$  is strictly smaller than  $\mathcal{P}(X)$ .
- (iii) *There is no largest size.*  $\mathbb{N}, \mathcal{P}(\mathbb{N}), \mathcal{P}(\mathcal{P}(\mathbb{N})), \dots$ : each set in this sequence is strictly bigger than the one before.

### Corollary 3.3.3. The Subsets of the Naturals

The power set  $\mathcal{P}(\mathbb{N})$  is uncountable.

*Proof.* The set  $\mathcal{P}(\mathbb{N})$  is non-empty, since  $\emptyset \in \mathcal{P}(\mathbb{N})$ , and by Theorem 3.3.2 with  $X = \mathbb{N}$  there is no surjection  $\mathbb{N} \rightarrow \mathcal{P}(\mathbb{N})$ . By Definition 3.1.1,  $\mathcal{P}(\mathbb{N})$  is not countable. ■

The two diagonal proofs are one. By the correspondence of Section 1.5,  $S \mapsto \mathbf{1}_S$  is a bijection  $\mathcal{P}(\mathbb{N}) \rightarrow \{0, 1\}^{\mathbb{N}}$ , with inverse  $b \mapsto \{n \in \mathbb{N} : b(n) = 1\}$  (Theorem 1.6.4). Given  $f: \mathbb{N} \rightarrow \mathcal{P}(\mathbb{N})$ , the sequence  $g(n) = 1 - \mathbf{1}_{f(n)}(n)$  of Theorem 3.3.1 is 1 exactly when  $n \notin f(n)$ , so  $g = \mathbf{1}_D$  for the set  $D$  of Cantor's proof.

### 3.3.4 Decimal expansions

The real numbers should be next. A real number in  $[0, 1)$  is written as a decimal  $0.d_1d_2d_3\dots$ , a sequence of digits, so the diagonal suggests itself: list the numbers, write their decimals as the rows of a table, and change the diagonal digits. Two questions must be answered first. What does it **mean** that a sequence of digits is the decimal of  $x$ ? And what is  $0.999\dots$ ? School says it equals 1, and a diagonal that produces  $0.4999\dots$  when  $0.5000\dots$  is on the list has produced a number that is on the list.

An infinite sum would answer the first question, but this book has not yet defined one. Finite sums are enough. The first  $n$  digits of  $x$  should pin  $x$  down to within  $10^{-n}$ .

*A digit sequence is a decimal of  $x$  when each of its finite truncations is within one unit of its last place below  $x$ .*

#### Definition 3.3.1. Decimal Expansion

A **digit sequence** is a function  $k \mapsto d_k$  from  $\mathbb{Z}_{>0}$  to  $\{0, 1, \dots, 9\}$ . For  $n \geq 1$  let

$$t_n := \sum_{k=1}^n d_k 10^{-k}$$

(Definition 2.3.1), where  $10^{-k} := (1/10)^k$  (Definition 2.3.2). The digit sequence is a **decimal expansion** of  $x \in [0, 1]$  if **for every**  $n \geq 1$ ,

$$t_n \leq x \leq t_n + 10^{-n}.$$

In words:  $t_n$  is the number written with the first  $n$  digits,  $0.d_1\dots d_n$ , and the condition says that  $x$  lies in the closed interval of length  $10^{-n}$  starting at  $t_n$ , for every  $n$ . The digits are indexed from 1, so that  $d_k$  is the digit in the  $k$ -th place after the point. No infinite sum appears: every  $t_n$  is a finite sum.

#### Examples.

1. *Zero.* The all-zero sequence has  $t_n = 0$ , and  $0 \leq 0 \leq 10^{-n}$ ; it is an expansion of 0.
2. *One half, twice.* The sequence  $5, 0, 0, \dots$  has  $t_n = \frac{1}{2}$  for every  $n \geq 1$ , and  $\frac{1}{2} \leq \frac{1}{2} \leq \frac{1}{2} + 10^{-n}$ . The sequence  $4, 9, 9, \dots$  has  $t_n = \frac{4}{10} + \sum_{k=2}^n 9 \cdot 10^{-k} = \frac{1}{2} - 10^{-n}$ , by Proposition 2.3.6, so  $t_n \leq \frac{1}{2} \leq t_n + 10^{-n} = \frac{1}{2}$ . Both are expansions of  $\frac{1}{2}$ .
3. *One.* The all-nine sequence has  $t_n = \sum_{k=1}^n 9 \cdot 10^{-k} = 1 - 10^{-n}$  (Proposition 2.3.6), so  $t_n \leq 1 \leq t_n + 10^{-n} = 1$ : it is an expansion of 1.

**A non-example by minimal change.** Keep the sequence  $5, 0, 0, \dots$  and move  $x$  from  $\frac{1}{2}$  to  $0.51$ . The inequalities still hold for  $n = 1$  ( $0.5 \leq 0.51 \leq 0.6$ ) and for  $n = 2$  ( $0.50 \leq 0.51 \leq 0.51$ ), and they fail at  $n = 3$ :  $0.51 > 0.500 + 0.001$ . So  $5, 0, 0, \dots$  is not an expansion of  $0.51$ ; the clause “for every  $n \geq 1$ ” is what fails.

**Why this definition.** Both inequalities are non-strict. With a strict right inequality,  $x < t_n + 10^{-n}$ , the all-nine sequence would be an expansion of no number at all: it would need  $x < 1$  together with  $1 - 10^{-n} \leq x$  for every  $n$ , and no real number satisfies both, by the Archimedean property (a school fact; *Analysis on the Real Line*, Chapter 1 proves it). The non-strict form is what makes  $0.999\dots = 1$  true, and true **by the definition**: the inequalities for  $x = 1$  are checked in Example 3, and no other number satisfies them, since  $1 - 10^{-n} \leq x \leq 1$  for every  $n$  forces  $x = 1$ , again by the Archimedean property (this is the uniqueness half of (D2) below).

Example 2 shows that the price is two expansions for some numbers.

What this book does not prove are the facts that make expansions behave as school says. We state them once and use them in this section and the next, always by name.

- (D1) Every  $x \in [0, 1)$  has a decimal expansion that does **not** end in all 9s.
- (D2) Every digit sequence is a decimal expansion of **exactly one**  $x \in [0, 1]$ .
- (D3) If two **different** digit sequences are expansions of the same  $x$ , then one of them ends in all 9s and the other ends in all 0s.

Here a digit sequence **ends in all 9s** if there is  $N \geq 1$  with  $d_k = 9$  for every  $k \geq N$ , and likewise for 0s. By the uniqueness in (D2), we may write  $x = 0.d_1d_2d_3\ldots$  for “ $x$  is the number of which  $d_1, d_2, d_3, \ldots$  is an expansion”. So  $0.4999\ldots = 0.5000\ldots = \frac{1}{2}$ , as in (D3).

Neither (D1) nor (D2) follows from what we have so far. (D3) does, by a finite computation at the first place where the two sequences differ (Theorem 2.1.3, Proposition 2.3.6); we state it with the others to quote it by name. The **existence** half of (D2), that some  $x$  satisfies all the inequalities at once, is where the completeness of  $\mathbb{R}$  hides, and *Analysis on the Real Line*, Chapter 1 takes completeness as an axiom. The **uniqueness** half of (D2) and the “not ending in all 9s” of (D1) rest on the Archimedean property: two numbers  $x, x'$  with  $|x - x'| \leq 10^{-n}$  for every  $n$  must be equal, which *Analysis on the Real Line*, Chapter 1 also proves. Until then we use them as school facts, and each use names the fact it uses.

One consequence is used below. *A digit sequence with no digit 0 and no digit 9 is the **only** expansion of its number.* Indeed, by (D3) a second expansion of the same number would force one of the two to end in all 9s and the other in all 0s; the given sequence ends in neither, since it has no 9s and no 0s at all.

### 3.3.5 The real numbers are uncountable

With the facts in hand, the diagonal runs on digits. Two precautions are built in. Each number on the list gets its expansion by a **rule**, so that no choosing is needed. And the new digits avoid 0 and 9, so that the new number has only one expansion.

#### Theorem 3.3.4. The Reals Are Uncountable, by Decimals

The interval  $[0, 1)$  is uncountable. Hence  $\mathbb{R}$  is uncountable.

*Idea.* Take an arbitrary list of numbers in  $[0, 1)$ , write each as its expansion with no tail of 9s, and change the  $n$ -th digit of the  $n$ -th number to 4 or 5. The resulting digit sequence is an expansion of some  $y$ , by (D2). If  $y$  were the  $n$ -th number, it would have two different expansions, one with no 0s or 9s, against (D3).

*Proof.* Let  $f: \mathbb{N} \rightarrow [0, 1)$  be any function. For  $n \geq 1$  let  $x_n := f(n - 1)$ , so that the list  $x_1, x_2, x_3, \ldots$  has the same values as  $f$ .

For each  $n \geq 1$ ,  $x_n$  has an expansion not ending in all 9s, by (D1), and only one: by (D3), two different such expansions would include one ending in all 9s. Let  $(d_{n,k})_{k \geq 1}$  be this expansion. It is named by a rule, “the expansion of  $x_n$  not ending in all 9s”, so no choice is involved.

Define a digit sequence by

$$e_n := \begin{cases} 5 & \text{if } d_{n,n} \neq 5, \\ 4 & \text{if } d_{n,n} = 5, \end{cases} \quad (n \geq 1).$$

Then  $e_n \neq d_{n,n}$  for every  $n \geq 1$ . By (D2), there is  $y \in [0, 1]$  of which  $(e_k)_{k \geq 1}$  is an expansion. Its first truncation is  $t_1 = e_1/10 \in \{0.4, 0.5\}$ , so  $0.4 \leq t_1 \leq y \leq t_1 + 0.1 \leq 0.6$ , and in particular  $y \in [0, 1]$ .

Let  $n \geq 1$ , and suppose  $y = x_n$ . Then  $(e_k)$  and  $(d_{n,k})$  are both expansions of  $y$ , and they are different, since they differ in place  $n$ . But  $(e_k)$  has no digit 0 and no digit 9, so by the consequence of (D3) stated above it is the **only** expansion of  $y$ , a contradiction. So  $y \neq x_n$  for every  $n \geq 1$ , and  $y$  is not a value of  $f$ . Since  $f$  was arbitrary, no function  $\mathbb{N} \rightarrow [0, 1]$  is surjective; and  $[0, 1]$  is non-empty, since it contains 0. By Definition 3.1.1,  $[0, 1]$  is not countable.

Finally, if  $\mathbb{R}$  were countable, its subset  $[0, 1]$  would be countable by Proposition 3.1.3. So  $\mathbb{R}$  is uncountable. ■

**Why the digits 4 and 5.** The first rule that comes to mind is  $e_n := 9 - d_{n,n}$ , which also changes every diagonal digit. Run it on the list  $x_1 = \frac{1}{2}$ ,  $x_n = 0$  for  $n \geq 2$  (that is,  $f(0) = \frac{1}{2}$  and  $f(n) = 0$  for  $n \geq 1$ ). The expansions are  $0.5000\dots$  and  $0.000\dots$ , so  $d_{1,1} = 5$  and  $d_{n,n} = 0$  for  $n \geq 2$ . The naive rule produces  $e_1 = 4$  and  $e_n = 9$  for  $n \geq 2$ : the sequence  $0.4999\dots$ , which is an expansion of  $\frac{1}{2} = x_1$  by Example 2. The new digit sequence differs from every row, and its number is on the list. The rule with 4 and 5, on the same list, gives  $0.4555\dots = \frac{41}{90}$ , which is on no row. Avoiding 0 and 9 is exactly what the last paragraph of the proof uses.

*Analysis on the Real Line*, Chapter 1 proves that  $\mathbb{R}$  is uncountable again, with no decimals at all, directly from its least upper bound axiom; that proof needs none of (D1)–(D3).

### 3.3.6 Transcendental numbers

A real number that is not algebraic is called **transcendental**. Section 3.2 showed that the set Alg of algebraic numbers is countable (Theorem 3.2.4).

#### Corollary 3.3.5. Transcendental Numbers Exist

The set  $\mathbb{R} \setminus \text{Alg}$  of transcendental numbers is uncountable. In particular, there is a real number that is not algebraic.

*Proof.* By excluded middle (Logic and Proof, Proposition 2.5.3), every real number is in Alg or not, so  $\mathbb{R} = \text{Alg} \cup (\mathbb{R} \setminus \text{Alg})$ . Suppose  $\mathbb{R} \setminus \text{Alg}$  is countable. Since Alg is countable by Theorem 3.2.4,  $\mathbb{R}$  is countable by Proposition 3.1.5, contradicting Theorem 3.3.4. So  $\mathbb{R} \setminus \text{Alg}$  is uncountable. In particular  $\mathbb{R} \setminus \text{Alg} \neq \emptyset$ , since  $\emptyset$  is countable (Definition 3.1.1). So not every  $x$  satisfies  $x \notin \mathbb{R} \setminus \text{Alg}$ : otherwise  $\mathbb{R} \setminus \text{Alg}$  would be an empty set (Definition 1.2.4), hence equal to  $\emptyset$  (Proposition 1.2.3). By the classical law Logic and Proof, Proposition 6.2.1 (a) and  $\neg \neg E$ , some  $x$  lies in  $\mathbb{R} \setminus \text{Alg}$ ; this  $x$  is a real number that is not algebraic. This last passage, from “not empty” to “has an element”, is a classical step (*Logic and Proof*, Section 8.7).

*Remark.* The proof names no transcendental number: it shows that the set of them is too large to be countable, hence not empty. Its classical steps are the split of  $\mathbb{R}$  into Alg and its complement (excluded middle) and the last one, from “not empty” to “has an element”. Compare the classical proof of Logic and Proof, Theorem 8.7.1 in *Logic and Proof*, Section 8.7, which also names no witness (unlike Logic and Proof, Proposition 8.7.2, which does). Proving that a particular number, such as  $\pi$ , is transcendental is a different and much harder matter.

### 3.3.7 Exercises

#### 3.3.7.1 A. Check your understanding

##### Exercise. A1

These parts are about the proof of [Theorem 3.3.2](#).

- Where in the proof is the surjectivity of  $f$  used, and what does it supply?
- Run the proof on  $X = \{1, 2, 3\}$  and  $f(1) = \{1, 2\}$ ,  $f(2) = \emptyset$ ,  $f(3) = \{1, 3\}$ . What is  $D$ ? Is  $D$  a value of  $f$ ?
- State the idea of the proof in two sentences.
- Why does the same argument not show that there is no surjection  $X \rightarrow X$ ?

*Solution.*

- It is used once, right after  $D$  is defined: surjectivity supplies an element  $d \in X$  with  $f(d) = D$ . The two cases  $d \in D$  and  $d \notin D$  are then both refuted. Everything before that line holds for every function  $f$ .
- Test each  $x$  against  $f(x)$ :  $1 \in f(1)$ , so  $1 \notin D$ ;  $2 \notin f(2) = \emptyset$ , so  $2 \in D$ ;  $3 \in f(3)$ , so  $3 \notin D$ . Hence  $D = \{2\}$ . It is not a value of  $f$ : the values are  $\{1, 2\}$ ,  $\emptyset$ ,  $\{1, 3\}$ , and  $D$  differs from  $f(x)$  about  $x$  in each case ( $1 \in f(1) \setminus D$ ,  $2 \in D \setminus f(2)$ ,  $3 \in f(3) \setminus D$ ).
- Build the set  $D$  of those  $x$  that are not in their own image  $f(x)$ . Then  $D$  disagrees with every  $f(x)$  about the element  $x$ , so it is not a value of  $f$ .
- The set  $D$  is a **subset** of  $X$ , so it is a candidate value only for a function into  $\mathcal{P}(X)$ . For  $f: X \rightarrow X$ , the argument (where the test  $x \in f(x)$  makes sense at all) shows that  $D$  is not a value of  $f$ , but  $D$  need not be an element of  $X$ , so nothing contradicts surjectivity. Indeed  $\text{id}_X$  is a surjection  $X \rightarrow X$ ; with  $f = \text{id}_X$  the argument is Russell's, and shows only that  $D = R_X$  is not an element of  $X$  (Theorem 1.2.4).

#### 3.3.7.2 B. Practice

##### Exercise. B1: A faded diagonal

Let  $\mathbb{N}^{\mathbb{N}}$  be the set of all functions  $\mathbb{N} \rightarrow \mathbb{N}$ . Here is a proof that  $\mathbb{N}^{\mathbb{N}}$  is uncountable, with its last step left blank. *Name the move, then complete it.*

Let  $f: \mathbb{N} \rightarrow \mathbb{N}^{\mathbb{N}}$  be any function, and write  $f_n := f(n)$ , so that  $f_0, f_1, f_2, \dots$  is a list of elements of  $\mathbb{N}^{\mathbb{N}}$ . Define  $g: \mathbb{N} \rightarrow \mathbb{N}$  by  $g(n) := f_n(n) + 1$ . (Last step: \_\_\_\_\_.) Since  $f$  was arbitrary, no function  $\mathbb{N} \rightarrow \mathbb{N}^{\mathbb{N}}$  is surjective, and as  $\mathbb{N}^{\mathbb{N}}$  is non-empty (it contains the zero function),  $\mathbb{N}^{\mathbb{N}}$  is not countable.

*Solution.* The move is **the diagonal**:  $g$  is built to disagree with the list's  $n$ -th entry in place  $n$ . The last step: for every  $n \in \mathbb{N}$ ,  $g(n) = f_n(n) + 1 \neq f_n(n)$ , so  $g \neq f_n$  by Proposition 1.5.1 (two equal functions agree at  $n$ ). Hence  $g$  is not on the list:  $g$  is an element of  $\mathbb{N}^{\mathbb{N}}$  (its values  $f_n(n) + 1$  are natural numbers) that is not a value of  $f$ , and  $f$  is not surjective.

### Exercise. B2

Prove that the set of **infinite** subsets of  $\mathbb{N}$  is uncountable.

*Solution.* Let  $\mathcal{F}$  be the set of finite subsets of  $\mathbb{N}$  and  $\mathcal{I}$  the set of infinite ones, both subsets of  $\mathcal{P}(\mathbb{N})$  defined by specification. By excluded middle (Logic and Proof, Proposition 2.5.3), every subset of  $\mathbb{N}$  is finite or not finite, that is, finite or infinite (Definition 2.5.3), so  $\mathcal{P}(\mathbb{N}) = \mathcal{F} \cup \mathcal{I}$ . The set  $\mathcal{F}$  is countable by Exercise (C1: Finite subsets of the natural numbers). If  $\mathcal{I}$  were countable,  $\mathcal{P}(\mathbb{N})$  would be countable by Proposition 3.1.5, contradicting [Corollary 3.3.3](#). So  $\mathcal{I}$  is uncountable.

### Exercise. B3: Expansions of three numbers

For each of  $x = \frac{1}{3}, \frac{1}{2}, 1$ , find **all** decimal expansions of  $x$ , checking the inequalities of [Definition 3.3.1](#) for each one you give, and say which of (D1)–(D3) each case illustrates.

*Solution.* We use Proposition 2.3.6 with  $r = \frac{1}{10}$ , after taking the constant factor  $\frac{1}{10}$  out of the sum:  $\sum_{k=1}^n 10^{-k} = \frac{1}{10} \sum_{k=1}^n (\frac{1}{10})^{k-1} = \frac{1}{10} \cdot \frac{1-10^{-n}}{9/10} = \frac{1}{9}(1-10^{-n})$ .

$x = \frac{1}{3}$ . The all-three sequence has  $t_n = \frac{3}{9}(1-10^{-n}) = \frac{1}{3} - \frac{1}{3}10^{-n}$ , so  $t_n \leq \frac{1}{3} \leq t_n + 10^{-n}$ . It is an expansion, and it does not end in all 9s, as (D1) promises. It ends in neither 9s nor 0s, so by (D3) no other expansion of  $\frac{1}{3}$  exists. Answer: 0.333... only.

$x = \frac{1}{2}$ . Example 2 after [Definition 3.3.1](#) checks that 0.5000... and 0.4999... are expansions. There are no others: if three different expansions existed, two of them would both end in all 9s or both not; but by (D3), of two different expansions of one number, one ends in all 9s and the other in all 0s, and a sequence ending in all 0s does not end in all 9s (from some place on, its digits would be both 0 and 9). Answer: exactly these two, which illustrate (D3); the first is the one (D1) promises.

$x = 1$ . The all-nine sequence is an expansion (Example 3). Conversely, let  $(d_k)$  be an expansion of 1, and suppose  $d_m \leq 8$  for some  $m$ . Since every digit is at most 9,  $t_m \leq \sum_{k=1}^m 9 \cdot 10^{-k} - 10^{-m} = 1 - 2 \cdot 10^{-m}$ , so  $t_m + 10^{-m} \leq 1 - 10^{-m} < 1$ , against the definition. So every digit is 9. Answer: 0.999... only. The all-nine sequence is an expansion of 1, an

instance of the existence half of (D2). Here (D3) has no second expansion to compare. The case also shows why (D1) is stated for  $[0, 1)$  only: the one expansion of 1 ends in all 9s.

### Exercise. B4: Prove or disprove

Prove each true statement, and show each false one false.

- (a) The set of periodic 0–1 sequences, those  $a \in \{0, 1\}^{\mathbb{N}}$  with some  $p \geq 1$  such that  $a(n + p) = a(n)$  for every  $n \in \mathbb{N}$ , is uncountable.
- (b) The set of 0–1 sequences with no two consecutive 1s is uncountable.
- (c) For every set  $X$  there is no injection  $\mathcal{P}(X) \rightarrow X$ .
- (d) If  $X$  is infinite, there is no surjection  $X \rightarrow X \times X$ .
- (e) The set of all functions  $\mathbb{Z} \rightarrow \{0, 1\}$  is uncountable.

*Solution.*

- (a) **False:** the set is countable. For  $p \in \mathbb{Z}_{>0}$  let  $P_p$  be the set of sequences  $a$  with  $a(n + p) = a(n)$  for every  $n \in \mathbb{N}$ . For  $a \in P_p$ ,  $a(qp + r) = a(r)$  for all  $q \in \mathbb{N}$  and  $0 \leq r < p$ , by induction on  $q$ . Every  $n \in \mathbb{N}$  is  $qp + r$  for such  $q$  and  $r$  (Theorem 2.2.3;  $q \geq 0$  since  $n \geq 0$  and  $r < p$ ). So  $a$  is determined by  $a(0), \dots, a(p - 1)$ , and  $a \mapsto a|_{\{0, \dots, p-1\}}$  is an injection of  $P_p$  into the set of functions  $\{0, \dots, p - 1\} \rightarrow \{0, 1\}$ , which is finite, with  $2^p$  elements (Proposition 2.5.8, through indicator functions). A finite set injects into  $\mathbb{N}$  (Theorem 3.1.2, (iii)  $\Rightarrow$  (ii)); composing,  $P_p$  injects into  $\mathbb{N}$ , so  $P_p$  is countable ((ii)  $\Rightarrow$  (i)). The set of periodic sequences is  $\bigcup_{p \in \mathbb{Z}_{>0}} P_p$ , countable by Theorem 3.2.3 (b), since  $\mathbb{Z}_{>0} \subseteq \mathbb{N}$  is countable (Proposition 3.1.3). Choice is not needed here, although the proof of the cited theorem uses Axiom 3.2.1 in general: the lists are given by a formula. Let  $w_{p,k}$  ( $k < 2^p$ ) be the  $k$ -th pattern  $\{0, \dots, p - 1\} \rightarrow \{0, 1\}$  in binary order, put  $w_{p,k} := w_{p,2^p-1}$  for  $k \geq 2^p$ , and let  $L(p, k)$  be the sequence  $n \mapsto w_{p,k}(r)$ ,  $r$  the remainder of  $n$  on division by  $p$ . Then  $L$  maps  $\mathbb{Z}_{>0} \times \mathbb{N}$  onto the set of periodic sequences, and composing it with a surjection  $\mathbb{N} \rightarrow \mathbb{Z}_{>0} \times \mathbb{N}$ , as in Example (Unions that come with their lists), lists them without Axiom 3.2.1. The diagonal fails here because the flipped diagonal of a list of periodic sequences need not be periodic: the first usual failure of [Example \(Method: the diagonal\)](#).
- (b) **True.** Send  $b \in \{0, 1\}^{\mathbb{N}}$  to the sequence  $a$  with  $a(2n) := b(n)$  and  $a(2n + 1) := 0$ . Every odd place of  $a$  is 0, and of two consecutive places one is odd, so  $a$  has no two consecutive 1s. The map is injective, since  $b$  is read back from the even places (Proposition 1.5.1). If the set were countable, composing with an injection of it into  $\mathbb{N}$  (Theorem 3.1.2, (i)  $\Rightarrow$  (ii)) would give an injection  $\{0, 1\}^{\mathbb{N}} \rightarrow \mathbb{N}$ , so  $\{0, 1\}^{\mathbb{N}}$  would be countable ((ii)  $\Rightarrow$  (i)), against [Theorem 3.3.1](#).
- (c) **True.** Let  $i: \mathcal{P}(X) \rightarrow X$  be injective. Since  $\emptyset \in \mathcal{P}(X)$ , the domain is non-empty, so  $i$  has a left inverse  $r: X \rightarrow \mathcal{P}(X)$  (Theorem 1.6.1 (b)). The equation  $r \circ i = \text{id}_{\mathcal{P}(X)}$  says that  $i$  is a right inverse of  $r$ , so  $r$  is surjective (Proposition 1.6.2), against [Theorem 3.3.2](#).
- (d) **False.** Take  $X = \mathbb{N}$ . The set  $\mathbb{N} \times \mathbb{N}$  is countable (Theorem 3.1.6) and non-empty, so Definition 3.1.1 gives a surjection  $\mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$ .
- (e) **True.** Restriction  $a \mapsto a|_{\mathbb{N}}$  maps  $\{0, 1\}^{\mathbb{Z}}$  onto  $\{0, 1\}^{\mathbb{N}}$ : a given  $b \in \{0, 1\}^{\mathbb{N}}$  is the restriction of its extension by 0 on the negative integers. If  $\{0, 1\}^{\mathbb{Z}}$  were countable, its image  $\{0, 1\}^{\mathbb{N}}$  would be countable (Proposition 3.1.4), against [Theorem 3.3.1](#).

### 3.3.7.3 C. Going deeper

#### Exercise. C1: Exit check: Cantor

Prove [Theorem 3.3.2](#): for every set  $X$ , there is no surjection  $X \rightarrow \mathcal{P}(X)$ . Deduce that  $\mathcal{P}(\mathbb{N})$  is uncountable.

*Solution.* Let  $f: X \rightarrow \mathcal{P}(X)$  be a function, and let  $D := \{x \in X : x \notin f(x)\}$ , a subset of  $X$ , so  $D \in \mathcal{P}(X)$ . Suppose  $f$  is surjective. Then  $D = f(d)$  for some  $d \in X$ . If  $d \in D$ , then  $d \notin f(d) = D$  by the definition of  $D$ , a contradiction; so  $d \notin D$ . But then  $d \notin f(d)$  and  $d \in X$ , so  $d \in D$  by the definition of  $D$ , a contradiction. Hence  $f$  is not surjective, and since  $f$  was arbitrary, no surjection  $X \rightarrow \mathcal{P}(X)$  exists.

For  $X = \mathbb{N}$ : by Definition 3.1.1,  $\mathcal{P}(\mathbb{N})$  is countable only if it is empty or there is a surjection  $\mathbb{N} \rightarrow \mathcal{P}(\mathbb{N})$ . It is not empty, since  $\emptyset \in \mathcal{P}(\mathbb{N})$ , and the theorem excludes the surjection. So  $\mathcal{P}(\mathbb{N})$  is uncountable.

**Marking scheme.** (10 points) - 2: takes an arbitrary  $f: X \rightarrow \mathcal{P}(X)$  and either assumes it onto, towards a contradiction, or sets out to show that it misses a set. - 3: defines  $D = \{x \in X : x \notin f(x)\}$  and notes that it is a subset of  $X$ . - 2: uses surjectivity to obtain  $d$  with  $f(d) = D$  (or, in the direct form, fixes an arbitrary  $x$ ). - 2: refutes both cases,  $d \in D$  and  $d \notin D$  (or shows that  $D$  and  $f(x)$  disagree about  $x$ ). - 1: the deduction for  $\mathcal{P}(\mathbb{N})$ , handling both clauses of “countable”: non-empty, and no surjection from  $\mathbb{N}$ . The direct form (“ $D \neq f(x)$  for every  $x$ , since they disagree about  $x$ ”) earns the same points: 2 for the arbitrary  $f$ , 3 for  $D$ , 4 for the disagreement at every  $x$ , 1 for the deduction. Fatal errors (0 for the step):  $D = \{x \in X : x \in f(x)\}$  (it may well be a value of  $f$ ); checking the claim for one particular  $f$  and concluding for all, as in (N1) of Logic and Proof, Example (What is not allowed) (0 for the first step).

#### Exercise. C2: Permutations of the naturals

Prove that the set of all bijections  $\mathbb{N} \rightarrow \mathbb{N}$  is uncountable.

*Solution (Hint 1).* Consider the bijections that only swap  $2n$  and  $2n + 1$ , for some of the  $n \in \mathbb{N}$ , and fix every other number.

*Solution (Hint 2).* Such bijections correspond to  $\{0, 1\}^{\mathbb{N}}$ : the  $n$ -th entry says whether the pair  $2n, 2n + 1$  is swapped.

*Solution.* Every  $m \in \mathbb{N}$  is  $2q + r$  for exactly one pair  $(q, r)$  with  $q \in \mathbb{N}$  and  $r \in \{0, 1\}$  (Theorem 2.2.3 with divisor 2, the quotient lying in  $\mathbb{N}$  as in the proof of Proposition 3.1.5). For  $b \in \{0, 1\}^{\mathbb{N}}$  define  $\pi_b: \mathbb{N} \rightarrow \mathbb{N}$  by

$$\pi_b(2q + r) := \begin{cases} 2q + r & \text{if } b(q) = 0, \\ 2q + 1 - r & \text{if } b(q) = 1, \end{cases}$$

so  $\pi_b$  swaps  $2q$  and  $2q + 1$  when  $b(q) = 1$  and fixes them otherwise. It maps each pair  $\{2q, 2q + 1\}$  to itself, and applying it twice returns every number to its place, so  $\pi_b \circ \pi_b = \text{id}_{\mathbb{N}}$ , and  $\pi_b$  is a bijection by Theorem 1.6.4.

Let  $B$  be the set of bijections  $\mathbb{N} \rightarrow \mathbb{N}$ ; it is non-empty, since  $\text{id}_{\mathbb{N}} \in B$ . Let  $f: \mathbb{N} \rightarrow B$  be any function, and write  $f_n := f(n)$ . Define  $b \in \{0, 1\}^{\mathbb{N}}$  by  $b(n) := 1$  if  $f_n(2n) = 2n$ , and  $b(n) := 0$  otherwise, and let  $\pi := \pi_b \in B$ . For every  $n \in \mathbb{N}$ ,  $\pi(2n) \neq f_n(2n)$ : if  $f_n(2n) = 2n$ , then  $b(n) = 1$  and  $\pi(2n) = 2n + 1$ ; if  $f_n(2n) \neq 2n$ , then  $b(n) = 0$  and  $\pi(2n) = 2n$ . So  $\pi \neq f_n$  for every  $n$  (Proposition 1.5.1), and  $f$  is not surjective. Since  $f$  was arbitrary and  $B \neq \emptyset$ ,  $B$  is uncountable (Definition 3.1.1).

This is the diagonal again, with the  $n$ -th entry tested at the place  $2n$ . Alternatively,  $b \mapsto \pi_b$  is an injection  $\{0, 1\}^{\mathbb{N}} \rightarrow B$ , since  $\pi_b(2q)$  is  $2q$  or  $2q + 1$  according to  $b(q)$ ; if  $B$  were countable, composing with an injection  $B \rightarrow \mathbb{N}$  (Theorem 3.1.2) would make  $\{0, 1\}^{\mathbb{N}}$  countable, against Theorem 3.3.1.

**Marking scheme.** (10 points) - 3: defines, for each  $b \in \{0, 1\}^{\mathbb{N}}$  (or each set of pairs), the bijection that swaps  $2n$  and  $2n + 1$  exactly when  $b(n) = 1$ , and proves it is a bijection. - 4: given an arbitrary list of bijections, defines  $b$  from the  $n$ -th entry's value at a place belonging to the  $n$ -th pair, and shows that the resulting bijection differs from the  $n$ -th entry; or shows that  $b \mapsto \pi_b$  is injective. - 3: concludes, with the non-emptiness of the set of bijections or through Theorem 3.3.1 and the closure properties of countable sets. Fatal errors (0 for the step): the plain diagonal  $g(n) := f_n(n) + 1$ , which is not a bijection; a "bijection" that moves numbers between different pairs without checking that it stays one-to-one and onto.