2. COUNTABILITY

In this section we introduce some of the simplest ideas about the size or cardinality of a set. (You should probably note, but not be too concerned, that we have not rigorously defined what a set is. Most (but by no means all) mathematicians agree on what a set is, but you will have to wait till the third year B1.2 option *Set Theory* to find out what the current consensus is.)

Almost all the results in this section are due to the German mathematician Georg Cantor (1845-1918). The lectures will focus mainly on the notion of *countability* and in particular that the real numbers are *uncountable*. The cardinality of finite sets was discussed in the *Introduction to University Mathematics* course.

Definition 2.1 *Let A and B be sets. We say A and B are* ***equinumerous****, and write $A \approx B$, if there is a bijection $f\colon A \to B$.*

Note that for any sets A, B, C,

$$\begin{aligned} &A \approx A; \\ A \approx B &\iff B \approx A; \\ A \approx B,\ B \approx C &\implies A \approx C. \end{aligned}$$

These properties rely on the identity map being a bijection, bijections being invertible and the composition of two bijections being a bijection.

Example 2.2 *The sets $A = \{0, 1, 2, 3, \ldots\}$ and $B = \{1, 2, 3, 4, \ldots\}$ are equinumerous despite B being a proper subset of A; we can see this by considering the bijection $f\colon A \to B$ given by $f(n) = n + 1$.*

Definition 2.3 *A set A is called* ***finite*** *if either $A = \varnothing$ or we have that $A \approx \{1, 2, \ldots, k\}$ for some non-zero natural number k. In the former case we say that A has cardinality 0, in the latter has cardinality k. We denote the* ***cardinality*** *of A by $|A|$.*

Remark 2.4 *Note that the cardinality of a finite set is well-defined. There would be issues with the above definition if it were possible to find a set A and distinct k, l such that*

$$A \approx \{1, 2, \ldots, k\}, \quad \text{and} \quad A \approx \{1, 2, \ldots, l\}.$$

To sketch a 'least criminal' proof, consider the smallest k for which $A \approx \{1, 2, \ldots, k\}$ and distinct l with $A \approx \{1, 2, \ldots, l\}$. We could then construct a bijection f from $\{1, \ldots, k\}$ to $\{1, \ldots, l\}$. Remove k and $f(k)$, and (with some adjustment) we'd get two equinumerous sets of sizes $k - 1$ and $l - 1$. But $k - 1 \neq l - 1$ which contradicts the minimality of k.

Exercise 2.5 *How would you prove the following for finite sets A and B?*

- If $A \subseteq B$ then $|A| \leqslant |B|$.
- If $f\colon A \to B$ is a 1-1 map then $|A| \leqslant |B|$.
- If $g\colon A \to B$ is an onto map then $|A| \geqslant |B|$.

Remark 2.6 *(Off-syllabus) More generally given two (possibly infinite) sets A and B we write $|A| \leqslant |B|$ if there is a 1-1 map from A to B. The **Cantor-Bernstein-Schröder Theorem** states that*

$$\text{if } |A| \leqslant |B| \text{ and } |B| \leqslant |A| \text{ then } A \approx B,$$

i.e. if there is a 1-1 maps $A \to B$ and $B \to A$ then there is a bijection from A to B.

Cantor was the first to publish this result, without proof, in 1887. A later proof by Cantor relied on the Axiom of Choice, which is a non-stand axiom of set theory, and unnecessary to this theorem. In 1887 Dedekind proved the theorem, without reference to the Axiom of Choice but did not publish his result. In 1897 Bernstein and Schröder independently published proofs.

Definition 2.7 *A set which is not finite is called **infinite**.*

Remark 2.8 *An equivalent definition for a set to be infinite is that the set has an equinumerous proper subset.*

Example 2.9 *The sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are all infinite.*

Somewhat surprisingly, we will see that the above sets are not all equinumerous.

Definition 2.10 *A set A is called **countably infinite** (or denumerable) if $\mathbb{N} \approx A$ We say A is **countable** if A is finite or countably infinite. A set which is not countable is called **uncountable.** (Note some authors use 'countable' to mean 'countably infinite'.)*

We then have:

Proposition 2.11 *A set A is countable if and only if there is a 1-1 map $f\colon A \to \mathbb{N}$.*

Corollary 2.12 *If $B \subseteq A$ and A is countable then B is also countable. Equivalently if B is uncountable then A is uncountable.*

Example 2.13 *The set of integers is countably infinite.*

Solution. A bijection from $\mathbb{N} = \{0, 1, 2, \ldots\}$ to $\mathbb{Z}$ can be described using the list

$$0, 1, -1, 2, -2, 3, -3, \ldots$$

or more formally by setting

$$f(n) = \begin{cases} (n+1)/2 & n \text{ is odd} \\ -n/2 & n \text{ is even} \end{cases}$$

■

We can generalise this approach to show:

Proposition 2.14 *Suppose that A_1 and A_2 are disjoint and countably infinite. Then $A_1 \cup A_2$ is countably infinite.*

Proof. As A_i is countably infinite then there is a bijection $f_i \colon \mathbb{N} \to A_i$. We define the map $g \colon \mathbb{N} \to A_1 \cup A_2$ by

$$g(n) = \begin{cases} f_1\left(\frac{n}{2}\right) & \text{if } n \text{ is even} \\ f_2\left(\frac{n-1}{2}\right) & \text{if } n \text{ is odd} \end{cases}$$

This map can be readily checked to be a bijection onto $A_1 \cup A_2$. ■

Remark 2.15 *The above proposition still holds even if A_1 and A_2 are not disjoint; essentially the same g can be used to list A_1 and A_2 but skipping over any repetitions as they occur.*

Proposition 2.16 *Suppose that A and B are countably infinite. Then the Cartesian prodcut $A \times B$ is countably infinite.*

Proof. As both sets are countably infinite then they can be listed as

$$a_0, a_1, a_2, \ldots \qquad b_0, b_1, b_2, \ldots$$

The elements (a_i, b_j) of $A \times B$ can be put into a grid as below

$$\begin{array}{ccccccccc}
(a_0,b_0) & \to & (a_1,b_0) & & (a_2,b_0) & \to & (a_3,b_0) & & (a_4,b_0) \\
 & \swarrow & & \nearrow & & \swarrow & & \nearrow & \\
(a_0,b_1) & & (a_1,b_1) & & (a_2,b_1) & & (a_3,b_1) & & (a_4,b_1) \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow & \\
(a_0,b_2) & & (a_1,b_2) & & (a_2,b_2) & & (a_3,b_2) & & (a_4,b_2) \\
 & \swarrow & & \nearrow & & \swarrow & & \nearrow & \\
(a_0,b_3) & & (a_1,b_3) & & (a_2,b_3) & & (a_3,b_3) & & (a_4,b_3) \\
\downarrow & \nearrow & & \swarrow & & \nearrow & & \swarrow & \\
(a_0,b_4) & & (a_1,b_4) & & (a_2,b_4) & & (a_3,b_4) & & (a_4,b_4)
\end{array}$$

and then can themselves be listed, in accordance with the arrows, as

$$(a_0,b_0), (a_1,b_0), (a_0,b_1), (a_0,b_2), (a_1,b_1), (a_2,b_0), (a_3,b_0), \ldots$$

■

Corollary 2.17 *If $A_0, A_1, A_2, \ldots$ are countable sets then so is their union $\bigcup_{i=0}^{\infty} A_i$.*

Proof. As each set A_i is countable then it can be listed as

$$a_{i0}, a_{i1}, a_{i2}, \ldots$$

By placing the a_{ij} into a square grid as in the previous proof then these can be counted in a similar fashion, omitting any repetitions of elements that arise. ■

Remark 2.18 *For those with a particular interest in set theory, note that the above proof relies on the Axiom of Choice in a subtle way. In listing each set A_i we are effectively choosing a bijection $f_i \colon \mathbb{N} \to A_i$ and to do so for each i requires the Axiom of Choice.*

Notation 2.19 *The symbol* $\aleph_0$ *is used to denote the cardinality* $|\mathbb{N}|$ *of* $\mathbb{N}$, *or any countably infinite set.* $\aleph$ *is aleph, the first letter of the Hebrew alphabet and* $\aleph_0$ *is read as 'aleph-null' or 'aleph-nought'.*

Example 2.20 *The set* $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$ *is countable as we have seen. An explicit example of an injection* $f\colon \mathbb{N}^2 \to \mathbb{N}$ *is the map*

$$f(m,n) = 2^m 3^n.$$

Example 2.21 *The set* $\mathbb{Q}^+$ *of positive rationals is countable as the map taking a rational* m/n *in its lowest form to* (m,n) *is an injection from* $\mathbb{Q}^+$ *to* $\mathbb{N}^2 \approx \mathbb{N}$. *So* $\mathbb{Q} = \{0\} \cup \mathbb{Q}^+ \cup \{-q \mid q \in \mathbb{Q}^+\}$ *is also countable.*

However it turns out that not all infinite sets are countable. In particular it is a fact of considerable importance that $\mathbb{R}$ is uncountable. This was first shown in 1874 by Cantor, producing a second more intuitive proof using his *diagonal argument* in 1891. Below we give two proofs. The second is Cantor's diagonal argument which makes use of decimal expansions – something we are yet to define and construct (see Example 5.12) – whilst the first proof uses results we have so far demonstrated.

Theorem 2.22 $\mathbb{R}$ *is uncountable.*

Proof. Proof 1: If $\mathbb{R}$ were countable, then so too would be $[0,1]$. Clearly $[0,1]$ is not finite as it contains all $\frac{1}{k}$ where $k \geqslant 1$. We proceed now with a proof by contradiction to show that $[0,1]$ is not countably infinite. Suppose $f : \mathbb{N} \to [0,1]$ is a bijection and we write $x_k = f(k)$.

- We choose distinct a_0, b_0 so that $x_0 \notin [a_0, b_0]$. If $x_0 \neq 1$ then we can find a_0 and b_0 such that $x_0 < a_0 < b_0 < 1$ and if $x_0 = 1$ then we can take the interval $[0, 1/2]$.

- Having chosen a_0, b_0 we then select a_1, b_1 so that $a_0 < a_1 < b_1 < b_0$ and $x_1 \notin [a_1, b_1]$. In a similar fashion to the above if $x_1 < b_0$ we can find a_1 and b_1 so that

$$\max(a_0, x_1) < a_1 < b_1 < b_0$$

  and if $x_1 \geqslant b_0$ then we can take the interval $[(2a_0 + b_0)/3, (a_0 + 2b_0)/3]$, i.e. the middle third of the previous interval.

- We repeat this process producing reals a_i and b_j such that

$$0 \leqslant a_0 < a_1 < a_2 < \cdots < b_2 < b_1 < b_0 \leqslant 1$$

  and $x_i \notin [a_i, b_i]$ for each i.

Now set $S = \{a_j \mid j \in \mathbb{N}\}$ which is bounded above by 1 and $T = \{b_j \mid j \in \mathbb{N}\}$ is bounded below by 0. So we may define

$$\lambda = \sup S \quad \text{and} \quad \mu = \inf T.$$

For all m, n we have $a_m \leqslant b_n$. In particular, each b_n is an upper bound of S and so $\lambda \leqslant b_n$ for all n as λ is the least upper bound of S. So λ is lower bound of T which means $\lambda \leqslant \mu$ as μ is the greatest lower bound of T. Then

$$a_n \leqslant \lambda \leqslant \mu \leqslant b_n \quad \text{for all } n.$$

For all n we have $\lambda \in [a_n, b_n]$ and $x_n \notin [a_n, b_n]$ and so $\lambda \neq x_n$ for all n which contradicts the fact that f is a bijection. ■

Proof. Proof 2 (diagonal argument): We will prove $\mathbb{R}$ is uncountable by showing that the interval $(0, 1]$ is uncountable. To each x in this interval corresponds a unique decimal expansion $0.a_1a_2a_3\ldots$ which does not end in a string of zeros.

Suppose for a contradiction that $f\colon \mathbb{N} \to (0, 1]$ is a bijection. Then we may uniquely write out the decimal expansions of $f(1), f(2), \ldots$. Say:

$$\begin{aligned} f(1) &= 0.r_{11}r_{12}r_{13}r_{14}\ldots \\ f(2) &= 0.r_{21}r_{22}r_{23}r_{24}\ldots \\ f(3) &= 0.r_{31}r_{32}r_{33}r_{34}\ldots \end{aligned}$$

Cantor then created a real α not on the list by setting

$$\alpha = 0.a_1a_2a_3\ldots$$

where

$$a_k = \begin{cases} 6 & \text{if } r_{kk} \neq 6 \\ 7 & \text{if } r_{kk} = 6 \end{cases}$$

The decimal expansion of α is allowed (in that it doesn't conclude in a string of 0s) and we see, for any k, that $\alpha \neq f(k)$ as α and $f(k)$ disagree in the kth decimal position. This contradicts the surjectivity of f. ■

Notation 2.23 *The symbol $\mathfrak{c}$, which stands for 'continuum' (an old name for the real line), denotes the cardinality of* $\mathbb{R}$.

Corollary 2.24 $\mathbb{C}$ *is uncountable. (In fact,* $\mathbb{C} \approx \mathbb{R}$*, which can be proved using the Cantor-Bernstein-Schröder theorem)*

The following result, known as Cantor's Theorem. It shows that any set has more subsets than elements. It further proves that there are ever increasingly large sets that can be formed.

Theorem 2.25 *(**Cantor's Theorem**, 1891) Let A be a set, and let $\mathcal{P}(A)$ be the **power set** of A, that is the set of subsets of A. Then*

$$|A| < |\mathcal{P}(A)|.$$

This means there is an injection from A to $\mathcal{P}(A)$ but there is no bijection from A to $\mathcal{P}(A)$.

Proof. The map $A \to \mathcal{P}(A)$ given by $a \mapsto \{a\}$ is an injection.

Suppose we have a map $f\colon A \to P(A)$. We show that f cannot be a surjection, and so cannot be a bijection. We consider the set

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

and will show that $X \notin f(A)$. Hence f is not onto.

Suppose to the contrary that $X = f(x)$ for some $x \in A$. Then either $x \in X$ or $x \notin X$. From the definition of X, we know that $x \in X$ if and only if $x \notin f(x) = X$. This is the required contradiction and so $X \notin f(A)$, as claimed. ■

Example 2.26 *Let $A = \{1, 2, 3\}$ and define $f\colon A \to \mathcal{P}(A)$ by*

$$f(1) = \varnothing, \qquad f(2) = A, \qquad f(3) = \{1, 2\}.$$

Find the set $X \subseteq A$ guaranteed by Cantor's theorem not to be in $f(A)$.

Solution. As $1 \notin f(1)$, $2 \in f(2)$, $3 \notin f(3)$ then $X = \{1, 3\}$. ■

Example 2.27 $\mathcal{P}(\mathbb{N}) \approx \mathbb{R}$. *(This can be proved using the Cantor-Bernstein-Schröder theorem.)*

Remark 2.28 *In the remainder of the course there will be very few explicit references to the uncountability of the real numbers. Having said that, it is the uncountability of $\mathbb{R}$ that characterises how we describe real numbers and impacts the nature of analysis.*

The integers, rational numbers, algebraic numbers (Sheet 2, Exercise 6) are all countable sets. Further the ***computable numbers*** *can be shown to be countable.*

A real number is said to be computable if there is a finite length computer program, written in a finite alphabet, that can (in principle) calculate that real number to any required accuracy. Essentially the set of computable numbers comprises all real numbers that can be described by finite means. However Cantor's proofs can be readily adapted to show that there are countably many such programs and so countably many computable numbers. This means, to describe the uncountably many real numbers, some infinite description is necessary – such as infinite decimal expansions.

Quite what this means is somewhat contentious. In this course we will consider arbitrary decimal expansions involving $0, 1, \ldots, 9$, but some logicians and mathematical philosophers take issue with this. In particular, 'intuitionists' would be content only with a decimal expansion that is defined constructively.