## 13. UNCOUNTABLE SETS AND BEYOND

The notion of countability would not be very useful if all sets were finite or countable. That this is not the case is the content of:

**Theorem 13.1** (Cantor 1891) $\{0,1\}^{\mathbb{N}}$ *is uncountable.*

*Proof.* Recall that $\{0,1\}^{\mathbb{N}}$ is the set of all functions $f\colon \mathbb{N} \to \{0,1\}$ with $\mathrm{Dom}(f) = \mathbb{N}$. Suppose for the sake of contradiction that $\{0,1\}^{\mathbb{N}}$ is countable. Hence there is a bijection $\phi\colon \mathbb{N} \to \{0,1\}^{\mathbb{N}}$. Denoting $f_n := \phi(n)$, the set $\{0,1\}^{\mathbb{N}}$ is thus enumerated into the sequence $\{f_n\}_{n\in\mathbb{N}}$. Now define $h\colon \mathbb{N} \to \{0,1\}$ by

$$\forall k \in \mathbb{N}\colon \quad h(k) := 1 - f_k(k) \tag{13.1}$$

Then $h \in \{0,1\}^{\mathbb{N}}$ with $\mathrm{Dom}(h) = \mathbb{N}$ yet $h$ is distinct from each $f_k$ (as it differs from $f_k$ at $k$). No bijection $\phi$ as above exists and so $\{0,1\}^{\mathbb{N}}$ is uncountable. □

The idea to arrange elements into a two-dimensional array and then produce another element by picking or changing the diagonal terms is often referred to as the *Cantor diagonal argument*. The above statement plus a bit of arithmetic shows:

**Corollary 13.2** *The interval* $[0,1] := \{x \in \mathbb{R}\colon 0 \leqslant x \leqslant 1\}$ *and thus also* $\mathbb{R}$ *are uncountable.*

*Proof.* The proof uses the notion of infinite series that we have not covered systematically. However, all we need is the following definition:

**Definition 13.3** *For any* $I \neq \varnothing$ *and any* $\{x_\alpha\colon \alpha \in I\} \subseteq [0,\infty) := \{x \in \mathbb{R}\colon 0 \leqslant x\}$*, if there is* $C \in [0,\infty)$ *such that*

$$\forall F \subseteq I\colon \ F \text{ finite } \Rightarrow \sum_{\alpha\in F} x_\alpha \leqslant C \tag{13.2}$$

*we set*

$$\sum_{\alpha\in I} x_\alpha := \sup\Big\{\sum_{\alpha\in F} x_\alpha\colon \ F \subseteq I \text{ finite}\Big\} \tag{13.3}$$

*(The supremum exists by the least upper bound property of the reals.)*

Now pick $\sigma \in \{0,1\}^{\mathbb{N}}$ and note that (11.5) gives $(1-q)\sum_{k=0}^{n} q^k = 1 - q^{n+1}$ for all $q \in \mathbb{R}$ and all $n \in \mathbb{N}$. This shows

$$0 \leqslant \sum_{k=0}^{n} \frac{2\sigma_k}{3^{k+1}} \leqslant \frac{2}{3}\sum_{k=0}^{n} 3^{-k} = \frac{2}{3}\frac{1-3^{-n+1}}{1-3^{-1}} \leqslant 1 \tag{13.4}$$

Hence we can set

$$f(\sigma) := \sum_{k\in\mathbb{N}} \frac{2\sigma_k}{3^{k+1}} \tag{13.5}$$

in the sense of Definition 13.3. This gives us a map $f\colon \{0,1\}^{\mathbb{N}} \to [0,1]$.

*Remark* 13.4 Note that $f(\sigma)$ is the real number whose representation in base-3 expansion takes the from $0.\eta_1\eta_2\eta_3\ldots$, where $\eta_i := 2\sigma_{i-1}$. In particular, $\mathrm{Ran}(f)$ is the set of all such numbers that can be written using 0's and 2's only, with no 1's allowed. Restricting to such numbers removes the degeneracy that all these expansions suffer from (e.g., $0.02222222\ldots$ is the same number as $0.10000000\ldots$) which would ruin injectivity of the map; see below. We will return to this later when we discuss *Cantor's ternary set*.

Next let $\sigma, \sigma' \in \{0,1\}^{\mathbb{N}}$ be such that $\sigma_k = \sigma'_k$ for all $k = 0, \dots, n-1$ and $\sigma_n = 0$ while $\sigma'_n = 1$. Then the fact that $\sup(b + A) = b + \sup(A)$ for $b + A := \{b + a \colon a \in A\}$ whenever $\sup(A)$ exists shows

$$\begin{aligned} f(\sigma) &= \sum_{k=0}^{n-1} \frac{2\sigma_k}{3^{k+1}} + \sum_{k \in [n+1,\infty)} \frac{2\sigma_k}{3^{k+1}} \\ &\leqslant \sum_{k=0}^{n-1} \frac{2\sigma_k}{3^{k+1}} + \frac{1}{3^{n+1}} \\ &< \sum_{k=0}^{n-1} \frac{2\sigma_k}{3^{k+1}} + \frac{2}{3^{n+1}} = \sum_{k=0}^{n} \frac{2\sigma'_k}{3^{k+1}} \leqslant f(\sigma') \end{aligned} \tag{13.6}$$

thus showing that $f$ is injective. If $[0,1]$ or $\mathbb{R}$ were countable, then Lemma 12.7 would imply that the image $f(\{0,1\}^{\mathbb{N}})$ is countable. But $f$ is a bijection of $\{0,1\}^{\mathbb{N}}$ onto its image, so that would give that $\{0,1\}^{\mathbb{N}}$ is countable, in contradiction with Theorem 13.1. □

The take-away message here is that there are just many more reals than rationals. From Theorem 13.1 and Lemma 12.14 we also conclude:

**Corollary 13.5** *There exists a real which is not algebraic.*

In fact, the argument shows most reals are not algebraic. Non-algebraic real or complex numbers are sometimes called *transcendental*.

We note that, prior to Cantor, Corollary 13.5 was proved by Liouville using the concept of *Liouville numbers* which are those $x \in \mathbb{R}$ such that for each $n \in \mathbb{N}$ there is a rational of the form $p/q$ with $q > 1$ for which $0 < |x - p/q| < q^{-n}$. As it turns out, such numbers exist and are all transcendental.

Cantor continued to develop the notion of cardinality further to include even larger sets than reals. We already encountered his relation $\simeq$ of *equinumerosity* in Definition 3.11 defined for any two sets $A$ and $B$ bz

$$A \simeq B \ := \ \exists f \colon A \to B \text{ bijection} \tag{13.7}$$

This is readily checked to be an equivalence relation on sets (technically, we have to talk about the class of all sets at this point or restrict to subsets of a given set). However, unlike for finite sets, we then define the cardinality of $A$ somewhat differently:

**Definition 13.6** *For any set $A$, the* cardinality *of $A$ is the equivalence class $[A]$ under the equinumerosity relation.*

Thus, the cardinality of the interval $[0,n) := \{k \in \mathbb{N} \colon k < n\}$ includes all sets with exactly $n$ elements. By Lemma 12.7, all infinite countable sets lie in $[\mathbb{N}]$ which (by Theorem 13.1 and Corollary 13.2) includes neither $\{0,1\}^{\mathbb{N}}$ nor $\mathbb{R}$. As each $\sigma \in \{0,1\}^{\mathbb{N}}$ is in one-to-one correspondence with the subset $A := \{n \in \mathbb{N} \colon \sigma(n) = 1\}$ of the naturals, also $\mathcal{P}(\mathbb{N})$ is uncountable and so

$$\mathcal{P}(\mathbb{N}) \notin [\mathbb{N}] \tag{13.8}$$

Another profound discovery made by Cantor is that this holds for all sets:

**Theorem 13.7** (Cantor 1891) *For any set $A$, there is no surjection $f\colon A \to \mathcal{P}(A)$ and so*

$$[A] \neq \big[\mathcal{P}(A)\big] \tag{13.9}$$

*Proof.* The proof is based on a variation of the Cantor diagonal argument albeit in the form that is more reminiscent of Russell's antinomy (see Theorem 2.1). Indeed, assume let $f\colon A \to \mathcal{P}(A)$ be a surjection. Define

$$B := \big\{x \in A\colon x \notin f(x)\big\} \tag{13.10}$$

Then $B \in \mathcal{P}(A)$ and, since $f$ is a surjection, there is $b \in A$ such that $f(b) = B$. But then $b \in B$ implies $b \notin f(b) = B$ while $b \notin B = f(b)$ implies $b \in B$. As at least one of $b \in B$ or $b \notin B$ must be TRUE — remember that $b \notin B$ is a shorthand for $\neg(b \in B)$ — we arrive at a contradiction and so no such $f$ can exist after all. □

Theorem 13.7 shows that applying the powerset to a given set keeps producing sets of different cardinality. Intuitively, $\mathcal{P}(A)$ is larger than $A$ but to formulate that precisely, we need a tool to compare sizes of non-equinumerous sets. This is furnished by the binary relation $\lesssim$ of *size comparison* defined as

$$A \lesssim B \;:=\; \exists f\colon A \to B \text{ injection} \tag{13.11}$$

This relation has many reasonable properties. Indeed, using the identity map we immediately verify the intuitive fact

$$A \subseteq B \Rightarrow A \lesssim B \tag{13.12}$$

The relation $\lesssim$ is also readily checked to be reflexive and transitive, but is not an ordering as antisymmetry cannot hold in general. Indeed, $A \lesssim B$ and $B \lesssim A$ do not imply that $A$ equals $B$ as these could be different sets (take $A := \mathbb{N}$ and $B := \mathbb{Q}$, for instance). However, the following natural alternative to equality does hold:

**Theorem 13.8** (Cantor-Bernstein/Schröder-Bernstein) *Let $A$ and $B$ be sets. Then*

$$A \lesssim B \;\wedge\; B \lesssim A \;\Rightarrow\; A \simeq B \tag{13.13}$$

*In particular, $\lesssim$ is a partial order on equinumerosity classes.*

We will not discuss the proof which can be found in all basic texts of set theory. What matters for us is that, in order to prove equinumerosity of two sets, it suffices to demonstrate an injection from one to the other and *vice versa*.

To see how this works in practice, note that the proof of Corollary 13.2 demonstrated an injection $\{0,1\}^{\mathbb{N}} \to [0,1]$. From $\mathcal{P}(\mathbb{N}) \simeq \{0,1\}^{\mathbb{N}}$ and $[0,1] \subseteq \mathbb{R}$ we get $\mathcal{P}(\mathbb{N}) \lesssim \mathbb{R}$. On the other hand, lifting the bijection between $\mathbb{N}$ and $\mathbb{Q}$ to their power sets shows $\mathcal{P}(\mathbb{N}) \simeq \mathcal{P}(\mathbb{Q})$ while the concept of Dedekind cut gives an injective embedding of $\mathbb{R}$ into $\mathcal{P}(\mathbb{Q})$ thus proving $\mathbb{R} \lesssim \mathcal{P}(\mathbb{N})$. Theorem 13.8 then concludes

$$\mathcal{P}(\mathbb{N}) \simeq \{0,1\}^{\mathbb{N}} \simeq [0,1] \simeq (0,1) \simeq \mathbb{R} \tag{13.14}$$

A special name is reserved for the equivalence class of all these sets:

**Definition 13.9** *The equivalence class of $\mathbb{R}$ under equinumerosity relation is called the* continuum. *Any set in $[\mathbb{R}]$ is said to have* cardinality of the continuum.

Similar arguments as used above in fact show that

$$\mathbb{R} \times \mathbb{R} \lesssim \mathbb{R} \tag{13.15}$$

and so, using also the trivial embedding $\mathbb{R} \lesssim \mathbb{R} \times \mathbb{R}$, Theorem 13.8 and induction give

$$\forall n \in \mathbb{N} \colon n \geqslant 1 \;\Rightarrow\; \mathbb{R}^n \simeq \mathbb{R} \tag{13.16}$$

meaning that all Euclidean spaces are of cardinality of the continuum. (We leave details to homework exercise.)

Pushing this further, the fact that $\mathbb{R} \simeq \{0,1\}^{\mathbb{N}}$ and $\mathbb{N} \times \mathbb{N} \simeq \mathbb{N}$ imply

$$\mathbb{R}^{\mathbb{N}} \simeq \{0,1\}^{\mathbb{N}\times\mathbb{N}} \simeq \{0,1\}^{\mathbb{N}} \simeq \mathbb{R} \tag{13.17}$$

so even the space of all real-valued sequences is of cardinality of the continuum. However, by Theorem 13.7, the space $\mathbb{R}^{\mathbb{R}}$ of all functions $\mathbb{R} \to \mathbb{R}$ and, by $\mathbb{N} \times \mathbb{R} \simeq \mathbb{R}$ (prove this!), also the space of just zero-one valued functions on $\mathbb{R}$ are strictly larger:

$$\mathbb{R} \lesssim \mathcal{P}(\mathbb{R}) \simeq \{0,1\}^{\mathbb{R}} \simeq \{0,1\}^{\mathbb{N}\times\mathbb{R}} \simeq \mathbb{R}^{\mathbb{R}} \tag{13.18}$$

where we used

$$A \lesssim B := A \lesssim B \wedge \neg(A \simeq B) \tag{13.19}$$

We again leave checking the details to the reader.

An attentive reader will notice that the various "orders of infinity" encountered above (namely, the naturals $\mathbb{N}$, the reals $\mathbb{R}$ in (13.14) and $\mathbb{R}$-valued functions on $\mathbb{R}$ in (13.18)) can be constructed from $\mathbb{N}$ by powerset operation. This suggests that we push this further: Using Theorem 13.7 along with the natural injection $x \mapsto \{x\}$ of every set in its power set demonstrates an infinite sequence

$$\mathbb{N} \lesssim \mathcal{P}(\mathbb{N}) \simeq \mathbb{R} \lesssim \mathcal{P}(\mathcal{P}(\mathbb{N})) \lesssim \mathcal{P}(\mathcal{P}(\mathbb{N})) \lesssim \mathcal{P}\big(\mathcal{P}(\mathcal{P}(\mathbb{N}))\big) \lesssim \dots \tag{13.20}$$

of infinite sets with distinct cardinalities. A natural question is whether other "orders of infinity" might exist as well. An immediate answer to this is in the affirmative: The union of the infinite sequence above (which is defined precisely via the Recursion principle) is not equinumerous to any member thereof. We have thus discovered yet another way (besides powerset) to produce "set larger than before" by taking the union of an infinite family of all "previous" infinities. (A formal construction requires the notion of a *limit ordinal*.)

That being said, more important for analysis (which is concerned mostly with the first two or three terms in the above sequence) is the question whether other "orders of infinity" are missing because the powerset construction "jumped" over them. This already concerns the first and second term and is the basis of:

**Continuum hypothesis:** (Cantor 1878) *Every infinite subset of the reals is either countable or of cardinality of the continuum. In short,*

$$\forall A \subseteq \mathbb{R} \colon\; A \text{ infinite } \;\Rightarrow\; \big(A \simeq \mathbb{N} \;\vee\; A \simeq \mathbb{R}\big) \tag{13.21}$$

This "hypothesis" has been a subject of much debate in the first half of the 20th century for it seemed to have demonstrable consequences for what could/could not be done in mathematics. A striking resolution was presented by K. Gödel in 1940 and P. Cohen in 1963 who showed that the Continuum hypothesis can be neither disproved (Gödel) nor proved (Cohen) in Zermelo's theory (assuming the latter is free of contradictions). In

other words, adding the Continuum hypothesis to Zermelo's axioms creates one mathematical universe, adding its negation creates another and yet another universe is produced by leaving the TRUE/FALSE value of (13.21) undecided.

For this reason, just as with the Axiom of choice, mathematicians are careful to regard results that rely on the Continuum hypothesis as conditional on this axiom to be added and often mark that by adding the acronym "CH" to the statement of the theorem. ("AC" is used to mark the use of the Axiom of choice.)