Cardinality

Michael B. Williams

1 Introduction

Suppose we are given the task or ordering a collection of sets from "smallest" to "largest." If all of the sets are finite, then (in principle) this task is trivial: we order them based on how many elements are in each set, using the ordering of the natural numbers. What if, however, the collection includes infinite sets? How can we make sense of the words "smaller" or "larger" when we can't actually assign to each set a finite number that describes how many elements the set contains? To resolve this issue, we introduce the concept of *cardinality*, which generalizes the notion of "size" of a set to allow for infinite sets. Doing this requires us to get a better picture of exactly what it means for a set to have infinitely many elements, and in fact, we will see that there are "different kinds" of infinities.

2 Countable Sets

Consider a typical finite set, such as the set of lower-case letters in the alphabet:

$$S = \{a, b, c, \ldots, x, y, z\}.$$

What does it mean to *count* the number of elements in the set? What we really do when we count is to assign to each element of the set a unique natural number, generally starting from 1 (or 0 if you're a computer scientist!) and proceeding upward:

$$\begin{array}{ccccccc} a & b & c & \cdots & x & y & z \\ \updownarrow & \updownarrow & \updownarrow & \cdots & \updownarrow & \updownarrow & \updownarrow \\ 1 & 2 & 3 & \cdots & 24 & 25 & 26 \end{array}$$

In this sense, any description of the elements of a set is immaterial; we only care about *how many* elements there are. This process of identifying elements with natural numbers (also called *counting numbers*) can be thought of in another way. When we count we are creating a function from our set S to a subset of the natural numbers that is both one-to-one and onto. A function is *one-to-one* if any two distinct inputs have two distinct outputs. A function is *onto* if every element in the target set (that is, the range) is the output of one or more elements in the domain. A one-to-one and onto function is also called a *bijection*. Another description of a bijection is that every element in the range is the output of exactly one element of the range. In functional notation, the above example is described as follows:

$$\begin{aligned} f : S &\longrightarrow \{1, 2, \ldots, 26\} \\ a &\longmapsto 1 \\ b &\longmapsto 2 \\ &\vdots \\ z &\longmapsto 26 \end{aligned}$$

That f is a bijection means that there is an *inverse function*, written f^{-1}, that undoes the action of f:

$$\begin{aligned} f^{-1} : \{1, 2, \ldots, 26\} &\longrightarrow S \\ 1 &\longmapsto a \\ 2 &\longmapsto b \\ &\vdots \\ 26 &\longmapsto z \end{aligned}$$

When we preform one after the other, the result is that we have done nothing, e.g.,

$$f(f^{-1}(1)) = 1, \quad f^{-1}(f(a)) = a.$$

We say that two sets have the same *cardinality* if there is a bijection from one set to the other. If the sets has finitely many elements, then cardinality corresponds to the number of elements in the sets. We could formally define a *finite set* to be one that is in bijection with a subset of the natural numbers of the form

$$\{1, 2, \ldots, N\} \subset \mathbb{N},$$

for some number N. For example,

$$\{\text{red}, \text{white}, \text{blue}\} \longleftrightarrow \{1, 2, 3\} \longleftrightarrow \{\text{breakfast}, \text{lunch}, \text{dinner}\}.$$

We can generalize the definition of a finite set and say that a set is *countable* if it can be put in bijection with any subset of the natural numbers. If that subset is $\mathbb{N}$ itself, then the original set is *countably infinite*.

Notice that with this definition, all finite sets are countable, and any subset of a countable set is still countable. The natural numbers are by definition countably infinite. Also, we note that ignoring what the elements of the set actually are, any countably infinite set essentially "looks like" a copy of $\mathbb{N}$, just as any finite set "looks like" a copy of the set $\{1, \ldots, N\}$ for some $N \in \mathbb{N}$.

If S and T are sets, recall that the *union* $S \cup T$ of the two sets is the set containing exactly the elements of both S and T. We can also take the union of arbitrarily many sets at once: $\cup_\alpha S_\alpha$, where $\{S_\alpha\}_{\alpha \in A}$ is a collection of sets indexed by a set A. Generally there is no restriction on the cardinality of the index set A, but the we will see that if A is countable and each S_α is countable, then so is the union.

Theorem 2.1. *The union of countably many countable sets is a countable set.*

Proof. Let the countably many sets be denoted

$$S_1, S_2, S_3, \ldots.$$

As a "worst case scenario," we may as well assume that there is a countably infinite number of sets, and that each set itself is countably infinite. Consider the union

$$S = \bigcup_{i=1}^{\infty} S_i = S_1 \cup S_2 \cup S_3 \cup \cdots$$

of these sets. To show that it is countable, we must put it into bijection with $\mathbb{N}$.

Suppose a set S_i has elements

$$a_{i1}, a_{i2}, a_{i3}, a_{i4}, \ldots.$$

Then we can arrange the elements of the union of the sets in a doubly-infinite array:

$$\begin{array}{ccccc}
a_{11} & a_{12} & a_{13} & a_{14} & \cdots \\
a_{21} & a_{22} & a_{23} & a_{24} & \cdots \\
a_{31} & a_{32} & a_{33} & a_{34} & \cdots \\
a_{41} & a_{42} & a_{43} & a_{44} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{array}$$

We create a bijection by counting each element as we follow the arrows:

$$\begin{array}{ccccccc}
a_{11} & \rightarrow & a_{12} & & a_{13} & \rightarrow & a_{14} \\
 & \swarrow & & \nearrow & & \swarrow & \\
a_{21} & & a_{22} & & a_{23} & & \\
\downarrow & \nearrow & & \swarrow & & & \\
a_{31} & & a_{32} & & & & \\
 & \swarrow & & & & & \\
a_{41} & & & & & & \\
\vdots & & & & & &
\end{array}$$

In other words, we have a pairing

$$\begin{array}{cccccccc}
a_{11} & a_{12} & a_{21} & a_{31} & a_{22} & a_{13} & a_{14} & \cdots \\
\updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\
1 & 2 & 3 & 4 & 5 & 6 & 7 & \cdots
\end{array}$$

of each element of S with a natural number. Clearly, each element in the union is accounted for at some point, and so S is countably infinite. $\square$

There are several immediate corollaries regarding the cardinalities of some familiar sets.

Corollary 2.2. *The set $\mathbb{Z}$ of integers is countably infinite.*

Proof. We exhibit $\mathbb{Z}$ as a countable union of countable sets:

$$\mathbb{Z} = \{1, 2, 3, \dots\} \cup \{0\} \cup \{-1, -2, -3, \dots\},$$

and apply the theorem. $\square$

Corollary 2.3. *The set $\mathbb{Q}$ of rational numbers is countably infinite.*

Proof. We exhibit $\mathbb{Q}$ as a countable union of countable sets. For each $k \in \mathbb{N}$, let

$$A_k = \left\{ \cdots, -\frac{2}{k}, -\frac{1}{k}, \frac{0}{k}, \frac{1}{k}, \frac{2}{k}, \dots \right\}.$$

Each of these sets is clearly in bijection with the integers, which is a countable set. But

$$\mathbb{Q} = \bigcup_{k \in \mathbb{N}} A_k,$$

so it is a countable union of countable sets, and the theorem applies. $\square$

There are many ways to prove this last fact, and here is another. Let

$$B_k = \left\{ \frac{1}{k}, \frac{2}{k}, \dots \right\},$$

for $k \in \mathbb{Z}$, with $B_0 = \{0\}$. Then

$$\mathbb{Q} = \bigcup_{k \in \mathbb{Z}} B_k.$$

Corollary 2.4. *The cartesian product of finitely many countable sets is countable.*

Proof. It is enough to consider the cartesian product

$$\mathbb{N}^d = \underbrace{\mathbb{N} \times \mathbb{N} \times \cdots \times \mathbb{N}}_{d \text{ times}} = \{(n_1, \dots, n_d) \mid n_1, \dots, n_d \in \mathbb{N}\},$$

since any countable set is in bijection with a subset of $\mathbb{N}$. The proof is by induction on d. The base case is $d = 1$, and is true by definition. Now suppose that $\mathbb{N}^{d-1}$ is countable. We can list its elements:

$$a_1, a_2, a_3, \dots$$

To show that $\mathbb{N}^d$ is countable, we simply write

$$\mathbb{N}^d = \mathbb{N}^{d-1} \times \mathbb{N},$$

and use the argument from the previous theorem. That is, we write this set as a doubly infinite array:

$$\begin{array}{ccccccc}
(a_1, 1) & \rightarrow & (a_1, 2) & & (a_1, 3) & \rightarrow & (a_1, 4) \\
& \swarrow & & \nearrow & & \swarrow & \\
(a_2, 1) & & (a_2, 2) & & (a_2, 3) & & \\
\downarrow & \nearrow & & \swarrow & & & \\
(a_3, 1) & & (a_3, 2) & & & & \\
& \swarrow & & & & & \\
(a_4, 1) & & & & & & \\
\vdots & & & & & &
\end{array}$$

which shows $\mathbb{N}^d$ is countable. $\square$

3 Uncountable Sets

At this point, one might wonder if in fact all sets are countable. We have shown that several familiar infinite sets all have the same cardinality, but in fact, there exist sets that are not countable. We call such sets *uncountable.*

Theorem 3.1. *The set $\mathbb{R}$ of real numbers is uncountable.*

Proof. It is enough to show that the interval $(0, 1) \subset \mathbb{R}$ is uncountable, since if $\mathbb{R}$ contains an uncountable subset, it certainly cannot itself be countable. We represent elements of $(0, 1)$ by their infinite decimal expansions, e.g.,

$$0.229384720983473\ldots$$

We also agree not to use any numbers ending in an infinite string of nines, to avoid ambiguity. Now, for a contradiction, suppose that this set is countable. We will construct an element that is in the set, but which was not counted. Write out our countable list of numbers as follows:

$$\begin{array}{c}
0\,.\,\underline{a_{11}}\,a_{12}\,a_{13}\,a_{14}\ldots \\
0\,.\,a_{21}\,\underline{a_{22}}\,a_{23}\,a_{24}\ldots \\
0\,.\,a_{31}\,a_{32}\,\underline{a_{33}}\,a_{34}\ldots \\
0\,.\,a_{41}\,a_{42}\,a_{43}\,\underline{a_{44}}\ldots \\
\vdots
\end{array}$$

To construct the missing number, let

$$b = 0\,.\,b_1\,b_2\,b_3\,b_4\ldots,$$

where each b_k can be any digit *except* a_{kk}, i.e., the underlined diagonal digits above. Thus, b is not in the list, since it differs from each number in the list in at least one decimal place. This is a contradiction, so the set $(0, 1)$ is uncountable. $\square$

The interval $(0,1)$, and thus $\mathbb{R}$, is uncountable, but in fact these two sets have the same cardinality. We can construct a bijection between them. Define

$$\begin{aligned} f : (0,1) &\longrightarrow \mathbb{R} \\ x &\longmapsto \tan\left(\pi x - \frac{\pi}{2}\right) \end{aligned}$$

and

$$\begin{aligned} f^{-1} : \mathbb{R} &\longrightarrow (0,1) \\ y &\longmapsto \frac{1}{\pi}\tan^{-1}(y) + \frac{1}{2} \end{aligned}$$

It is easy to check that these functions are inverses, and so they are bijections.

Since $\mathbb{Q}$ is a countable subset of $\mathbb{R}$, which is uncountable, we see that infinitely many real numbers are not rational. In fact, the rationals are a vanishingly small subset of the reals. Even taking cartesian products of arbitrary (finite) arity yields only countable sets, not even approaching uncountability. Uncountable therefore represents a "bigger" infinity than countable.

4 Other Cardinalities

Having discovered the uncountability of the reals is, a natural question is, "Do there exist sets with cardinality greater than that of the reals?" As before, we could attempt to find a larger cardinality by taking cartesian products; intuitively, the plane $\mathbb{R} \times \mathbb{R}$ has "more points" than the real line, so one might suspect that the cardinality of $\mathbb{R} \times \mathbb{R}$ is "bigger than" the cardinality of $\mathbb{R}$. This turns out to be false!

We will consider the open interval $(0,1) \subset \mathbb{R}$ and the open square $(0,1) \times (0,1) \subset \mathbb{R}^2$. It is cumbersome to construct a bijection, so we will instead construct a one-to-one map $f : (0,1) \times (0,1) \to (0,1)$, which will at least show that the cardinality of $(0,1) \times (0,1)$ is no larger than that of $(0,1)$ and hence of $\mathbb{R}$.

Represent real numbers by their decimal expansions, this time taking them to be non-terminating, that is, such that the number of non-zero digits is infinte. Similarly, represent points in the plane by pairs of real number decimal expansions, e.g.,

$$(x,y) = (0\,.\,a_1\,a_2\,a_3\cdots, 0\,.\,b_1\,b_2\,b_3\cdots) \in \mathbb{R} \times \mathbb{R}.$$

Now define

$$f(x,y) = f(0\,.\,a_1 a_2\cdots, 0\,.\,b_1\,b_2\cdots) = 0\,.\,a_1\,b_1\,a_2\,b_2\cdots$$

It is not hard to see that this is one-to-one.

It is easy to see that $(0,1) \times (0,1)$ and $\mathbb{R} \times \mathbb{R}$ have the same cardinality, so indeed $\mathbb{R} \times \mathbb{R}$ is really no larger than $\mathbb{R}$. In fact, the cardinalities are the same (but we haven't proven it!). This can also be extended to cartesian products: $\mathbb{R}^d$ has the same cardinality as $\mathbb{R}$ for any natural number d.

Despite this result, the answer to the above question, "Do there exist sets with cardinality greater than that of the reals?" is "yes." Given any set, we provide a method for constructing a set with strictly larger cardinality.

If S is a set, the *power set* of S, denoted $P(S)$, is the collection of all subsets of S. For example, if

$$S = \{a, b, c\},$$

then

$$P(S) = \Big\{\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\Big\}.$$

(Recall that the empty set $\varnothing$ is a subset of *every* set, even itself.)

In the finite case, it is obvious that $P(S)$ has strictly more elements than does S. In fact, if S has n elements, then $P(S)$ has 2^n elements. One quick proof of this fact relies on simple combinatorial principles. To select a subset with k elements from S is to select k arbitrary elements from S without order. There are

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

ways to do this for any given k. To find all the subsets, we repeat this procedure for each value of k, from 0 to n. Summing, and using the binomial formula yields

$$\sum_{k=0}^{n} \binom{n}{k} = (1+1)^n = 2^n$$

subsets.

It is a more interesting fact that the analagous statement is true for infinite sets. This should not be obvious; after all, infinite sets can display behavior that is completely different from finite sets. For example, an infinite set can be put in bijection with a proper subset of itself (e.g., the integers and the even integers), something that is certainly not true for finite sets!

Two sets have the same cardinality if there is a one-to-one and onto function from one to the other. Therefore, a necessary condition for two sets to have the same cardinality is the existance of an onto function from one set to the other.

Theorem 4.1. *For any set S, there is no onto function $S \to P(S)$.*

Proof. The proof is by contradiction. Suppose $f : S \to P(S)$ is onto. This means that for each $A \in P(S)$ (so A is a subset of S!), there is an element $a \in S$ such that $f(a) = A$. There are two possibilities:

$$(1)\ \ a \in A \qquad (2)\ \ a \notin A.$$

Let's define a new set:

$$B = \{a \in S \mid a \notin f(a)\} \subset S.$$

Since f is onto, there must be some $b \in S$ such that $f(b) = B$. Now, the question is "Is b an element of B?"

Suppose so. Then by definition of B, we must have

$$b \notin f(b) = B,$$

which is impossible. On the other hand, suppose not. Then again by the definition of B, we have

$$b \in f(b) = B,$$

which is also impossible. Since both possibilities lead to a contradiction, we conclude that such a function f cannot exist. □

The main idea of this proof has many non-mathematical representations. Here is one:

> Suppose that there is a barber in Seville who shaves only those men who do not shave themselves.
>
> Who shaves the barber?

Any attempt to answer this question lead to a contradiction, as in the proof.

The main implication of this theorem is that we can construct a sequence of sets, each with "larger" cardinality than the one before:

$$S, P(S), P(P(S)), P(P(P(S))), \ldots.$$

To make the word "larger" precise, say that the cardinality of a set S is *larger* than the cardinality of a set T if there is a one-to-one function $T \to S$, but there is no such onto function. Thus, if we start with an infinte set S, then we will have larger and larger infinte sets, that is, larger and larger cardinalities. A natural question at this point is, "Does this sequence ever end?" Or, to put it another way, is there a "largest" set, or a "set of everything" that ends the sequence?

Let us call a set *pleasant* if $A \notin A$. This seems like a strange property for a set to have, but if there is a set of everything, then "the set of everything" is also a thing, so it must be a member of itself. Let R be the set of all pleasant sets. As in the proof of the previous theorem, we have two possibilities:

$$(1)\ \ R \in R \qquad (2)\ \ R \notin R.$$

If $R \in R$, then R is not pleasant, by the definition of pleasant. But since R is the set of all pleasant sets, $R \notin R$. On the other hand, if $R \notin R$, then by definition of R, R is not pleasant. By the definition of pleesent, $R \in R$. Either way, we have a contradiction.

This is known as *Russell's Paradox.* It implies that the answer to the above question is "no," there cannot be a set of everytyhing, and this indicates that care is needed when speaking of "the set of all" of anything.

Additionally, it means that the chain of increasingly large sets above does not end, so in fact, there are infinitely many cardinalities. Or, put another way, there are infinitely many infinities!

5 A Brain Teaser

Suppose that a hotel (possibly in an alternate universe) has a countably infinite number of rooms, and suppose they are labeled with the natural numbers:

$$\text{Room } 1, \text{Room } 2, \ldots.$$

On some dark and stormy night, all the rooms in the hotel are filled.

(1) A couple comes to the hotel, looking for a room. How can the hotel, despite being full, accomodate the couple (that is, give them a normal hotel room) without kicking out any other hotel guests?

(2) A group of N people comes to the hotel, each looking for a room. How can the hotel, despite being full, accomodate the new guests (that is, give them a normal hotel room) without kicking out any other hotel guests?

(3) A countably infinite number of people come to the hotel, each looking for a room. How can the hotel, despite being full, accomodate each new guest (that is, give them a normal hotel room) without kicking out any other hotel guests?