

## Infinity and Countability

In this note we'll discuss the question of when two sets have the same cardinality, or size. This is a simple issue for finite sets, but for infinite sets it becomes quite tricky. We'll see how to formulate the question precisely, and then see several quite surprising consequences. To set the scene, we begin with the simple case of finite sets.

### Bijections

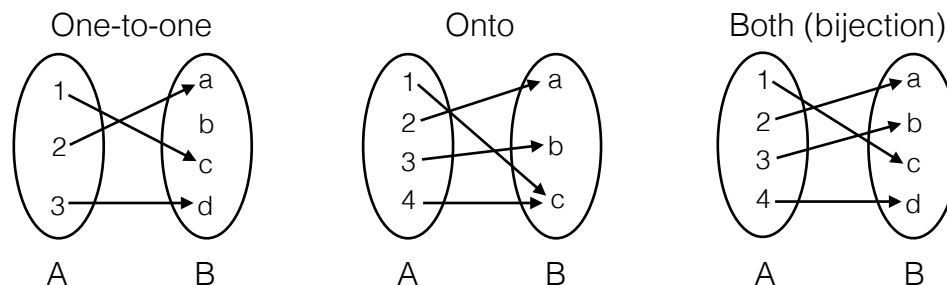
Two finite sets have the same size if and only if their elements can be paired up, so that each element of one set has a unique partner in the other set, and vice versa. We formalize this through the concept of a *bijection*.

Consider a function (or mapping)  $f$  that maps elements of a set  $A$  (called the *domain* of  $f$ ) to elements of set  $B$  (called the *range* of  $f$ ). Since  $f$  is a function, it must specify, for each element  $x \in A$  ("input"), exactly one element  $f(x) \in B$  ("output"). Recall that we write this as  $f : A \rightarrow B$ . We say that  $f$  is a *bijection* if every element  $a \in A$  has a unique *image*  $b = f(a) \in B$ , and every element  $b \in B$  has a unique *pre-image*  $a \in A : f(a) = b$ .

$f$  is a *one-to-one function* (or an *injection*) if  $f$  maps distinct inputs to distinct outputs. More rigorously,  $f$  is one-to-one if the following holds:  $x \neq y \Rightarrow f(x) \neq f(y)$ .

$f$  is *onto* (or *surjective*) if it "hits" every element in the range (i.e., each element in the range has at least one pre-image). More precisely, a function  $f$  is onto if the following holds:  $(\forall y \exists x)(f(x) = y)$ .

Here are some simple examples to help visualize one-to-one and onto functions, and bijections:



Note that, according to the above definitions,  $f$  is a bijection if and only if it is both one-to-one and onto.

---

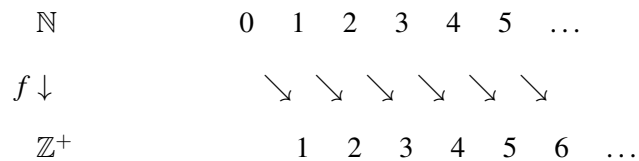
*Exercise.* Let  $f : A \rightarrow B$  be a bijection. Show that  $f$  has an *inverse*  $f^{-1} : B \rightarrow A$  that satisfies  $f^{-1}(f(a)) = a$  for all  $a \in A$ , and that  $f^{-1}$  is also a bijection.

---

# Cardinality

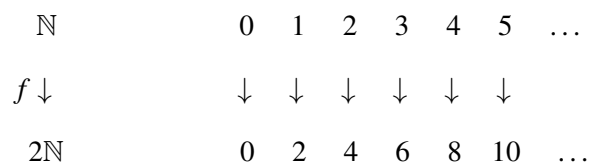
How can we determine whether two sets have the same *cardinality* (or “size”)? The answer to this question, reassuringly, lies in early grade school memories: by demonstrating a *pairing* between elements of the two sets. More formally, we need to demonstrate a *bijection*  $f$  between the two sets. The bijection sets up a one-to-one correspondence, or pairing, between elements of the two sets. We’ve seen above how this works for finite sets. In the rest of this lecture, we will see what it tells us about *infinite* sets.

Our first question about infinite sets is the following: Are there more natural numbers  $\mathbb{N}$  than there are positive integers  $\mathbb{Z}^+$ ? It is tempting to answer yes, since every positive integer is also a natural number, but the natural numbers have one extra element  $0 \notin \mathbb{Z}^+$ . Upon more careful observation, though, we see that we can define a mapping between the natural numbers and the positive integers as follows:



Why is this mapping a bijection? Clearly, the function  $f : \mathbb{N} \rightarrow \mathbb{Z}^+$  is onto because every positive integer is hit. And it is also one-to-one because no two natural numbers have the same image. (The image of  $n$  is  $f(n) = n + 1$ , so if  $f(n) = f(m)$  then we must have  $n = m$ .) Since we have shown a bijection between  $\mathbb{N}$  and  $\mathbb{Z}^+$ , this tells us that there are as many natural numbers as there are positive integers! (Very) informally, we have proved that “ $\infty + 1 = \infty$ .”

What about the set of *even* natural numbers  $2\mathbb{N} = \{0, 2, 4, 6, \dots\}$ ? In the previous example the difference was just one element. But in this example, there seem to be twice as many natural numbers as there are even natural numbers. Surely, the cardinality of  $\mathbb{N}$  must be larger than that of  $2\mathbb{N}$  since  $\mathbb{N}$  contains all of the odd natural numbers as well? Though it might seem to be a more difficult task, let us attempt to find a bijection between the two sets using the following mapping:



The mapping in this example is also a bijection.  $f$  is clearly one-to-one, since distinct natural numbers get mapped to distinct even natural numbers (because  $f(n) = 2n$ ).  $f$  is also onto, since every  $n$  in the range is hit: its pre-image is  $\frac{n}{2}$ . Since we have found a bijection between these two sets, this tells us that in fact  $\mathbb{N}$  and  $2\mathbb{N}$  have the same cardinality!

What about the set of all integers,  $\mathbb{Z}$ ? At first glance, it may seem obvious that the set of integers is larger than the set of natural numbers, since it includes infinitely many negative numbers. However, as it turns out, it is possible to find a bijection between the two sets, meaning that the two sets have the same size! Consider the following mapping  $f$ :

$$0 \rightarrow 0, 1 \rightarrow -1, 2 \rightarrow 1, 3 \rightarrow -2, 4 \rightarrow 2, \dots, 124 \rightarrow 62, \dots$$



Each rational number  $\frac{a}{b}$  (written in its lowest terms, so that  $\gcd(a, b) = 1$ ) is represented by the point  $(a, b)$  in the infinite two-dimensional grid shown (which corresponds to  $\mathbb{Z} \times \mathbb{Z}$ , the set of all pairs of integers). Note that not all points on the grid are valid representations of rationals: e.g., all points on the  $x$ -axis have  $b = 0$  so none are valid (except for  $(0, 0)$ , which we take to represent the rational number 0); and points such as  $(2, 8)$  and  $(-1, -4)$  are not valid either as the rational number  $\frac{1}{4}$  is represented by  $(1, 4)$ . But  $\mathbb{Z} \times \mathbb{Z}$  certainly contains all rationals under this representation, so if we come up with an injection from  $\mathbb{Z} \times \mathbb{Z}$  to  $\mathbb{N}$  then this will also be an injection from  $\mathbb{Q}$  to  $\mathbb{N}$  (why?).

The idea is to map each pair  $(a, b)$  to its position along the spiral, starting at the origin. (Thus, e.g.,  $(0, 0) \rightarrow 0$ ,  $(1, 0) \rightarrow 1$ ,  $(1, 1) \rightarrow 2$ ,  $(0, 1) \rightarrow 3$ , and so on.) It should be clear that this mapping maps every pair of integers injectively to a natural number, because each pair occupies a unique position along the spiral.

This tells us that  $|\mathbb{Q}| \leq |\mathbb{N}|$ . Since also  $|\mathbb{N}| \leq |\mathbb{Q}|$ , as we observed earlier, by the Cantor-Bernstein Theorem  $\mathbb{N}$  and  $\mathbb{Q}$  have the same cardinality.

*Exercise.* Show that the set  $\mathbb{N} \times \mathbb{N}$  of all ordered pairs of natural numbers is countable.

Our next example concerns the set of all binary strings (of any finite length), denoted  $\{0, 1\}^*$ . Despite the fact that this set contains strings of unbounded length, it turns out to have the same cardinality as  $\mathbb{N}$ . To see this, we set up a direct bijection  $f : \{0, 1\}^* \rightarrow \mathbb{N}$  as follows. Note that it suffices to *enumerate* the elements of  $\{0, 1\}^*$  in such a way that each string appears exactly once in the list. We then get our bijection by setting  $f(n)$  to be the  $n$ th string in the list. How do we enumerate the strings in  $\{0, 1\}^*$ ? Well, it's natural to list them in increasing order of length, and then (say) in *lexicographic* order (or, equivalently, numerically increasing order when viewed as binary numbers) within the strings of each length. This means that the list would look like

$\epsilon, 0, 1, 00, 01, 10, 11, 000, 001, 010, 011, 100, 101, 110, 111, 1000, \dots$ ,

where  $\epsilon$  denotes the empty string (the only string of length 0). It should be clear that this list contains each binary string once and only once, so we get a bijection with  $\mathbb{N}$  as desired.

Our final countable example is the set of all polynomials with natural number coefficients, which we denote  $\mathbb{N}(x)$ . To see that this set is countable, we will make use of (a variant of) the previous example. Note first that, by essentially the same argument as for  $\{0, 1\}^*$ , we can see that the set of all *ternary* strings  $\{0, 1, 2\}^*$  (that is, strings over the alphabet  $\{0, 1, 2\}$ ) is countable. To see that  $\mathbb{N}(x)$  is countable, it therefore suffices to exhibit an injection  $f : \mathbb{N}(x) \rightarrow \{0, 1, 2\}^*$ , which in turn will give an injection from  $\mathbb{N}(x)$  to  $\mathbb{N}$ . (It is obvious that there exists an injection from  $\mathbb{N}$  to  $\mathbb{N}(x)$ , since each natural number  $n$  is itself trivially a polynomial, namely the constant polynomial  $n$  itself.)

How do we define  $f$ ? Let's first consider an example, namely the polynomial  $p(x) = 5x^5 + 2x^4 + 7x^3 + 4x + 6$ . We can list the coefficients of  $p(x)$  as follows:  $(5, 2, 7, 0, 4, 6)$ . We can then write these coefficients as binary strings:  $(101, 10, 111, 0, 100, 110)$ . Now, we can construct a ternary string where a "2" is inserted as a separator between each binary coefficient (ignoring coefficients that are 0). Thus we map  $p(x)$  to a ternary string as illustrated below:

$$\begin{array}{c}
 5x^5 + 2x^4 + 7x^3 + 4x + 6 \\
 \downarrow \\
 1012102111221002110
 \end{array}$$

It is easy to check that this is an injection, since the original polynomial can be uniquely recovered from this ternary string by simply reading off the coefficients between each successive pair of 2's. (Notice that this mapping  $f : \mathbb{N}(x) \rightarrow \{0, 1, 2\}^*$  is not onto (and hence not a bijection) since many ternary strings will not be the image of any polynomials; this will be the case, for example, for any ternary strings that contain binary subsequences with leading zeros.)

Hence we have an injection from  $\mathbb{N}(x)$  to  $\mathbb{N}$ , so  $\mathbb{N}(x)$  is countable.

## Cantor's Diagonalization

We have established that  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$  all have the same cardinality. What about  $\mathbb{R}$ , the set of real numbers? Surely they are countable too? After all, the rational numbers, like the real numbers, are dense (i.e., between any two rational numbers  $a, b$  there is a rational number, namely  $\frac{a+b}{2}$ ). In fact, between any two *real* numbers there is always a rational number. It is really surprising, then, that there are more real numbers than rationals. That is, there is *no* bijection between the rationals (or the natural numbers) and the reals. We shall now prove this, using a beautiful argument due to Cantor that is known as *diagonalization*. In fact, we will show something even stronger: the real numbers in the interval  $[0, 1]$  are uncountable!

---

*Exercise.* Show how to find a rational number between any two (distinct) real numbers.

---

In preparation for the proof, recall that any real number can be written out uniquely as an infinite decimal with no trailing zeros. In particular, a real number in the interval  $[0, 1]$  can be written as  $0.d_1d_2d_3\dots$ . In this representation, we write for example 1 as  $0.999\dots^1$ , and 0.5 as  $0.4999\dots$ . (Thus rational numbers will always be represented as recurring decimals, while irrational ones will be represented as non-recurring ones. Importantly for us, all of these expressions will be infinitely long and unique.)

**Theorem:** The real interval  $\mathbb{R}[0, 1]$  (and hence also the set of real numbers  $\mathbb{R}$ ) is uncountable.

**Proof:** Suppose towards a contradiction that there is a bijection  $f : \mathbb{N} \rightarrow \mathbb{R}[0, 1]$ . Then, we can enumerate the real numbers in an infinite list  $f(0), f(1), f(2), \dots$  as follows:

$$\begin{array}{l} f(0) = 0.\textcircled{5}2\ 1\ 4\ 9\ 3\ 5\ 6\ \dots \\ f(1) = 0.\ 1\textcircled{4}1\ 6\ 2\ 9\ 8\ 5\ \dots \\ f(2) = 0.\ 9\ 4\textcircled{7}8\ 2\ 7\ 1\ 2\ \dots \\ f(3) = 0.\ 5\ 3\ 0\textcircled{9}8\ 1\ 7\ 5\ \dots \\ \vdots \qquad \qquad \qquad \vdots \end{array}$$

The number circled in the diagonal can be viewed as some real number  $r = 0.5479\dots$ , since it is an infinite decimal expansion. Now consider the real number  $s$  obtained by modifying every digit of  $r$ , say by replacing each digit  $d$  with  $d + 2 \pmod{10}$ ; thus in our example above,  $s = 0.7691\dots$ . We claim that  $s$  does not occur in our infinite list of real numbers. Suppose for contradiction that it did, and that it was the  $n^{\text{th}}$  number in the list,  $f(n)$ . But by construction  $s$  differs from  $f(n)$  in the  $(n + 1)$ th digit, so these two numbers cannot be equal! So we have constructed a real number  $s$  that is not in the range of  $f$ . But this contradicts the assertion that  $f$  is a bijection. Thus the real numbers are not countable.

---

<sup>1</sup>To see this, write  $x = .999\dots$ . Then  $10x = 9.999\dots$ , so  $9x = 9$ , and thus  $x = 1$ .

Let us remark that the reason that we modified each digit by adding 2 mod 10 as opposed to adding 1 is that the same real number can have two decimal expansions; for example  $0.999\dots = 1.000\dots$ . But if two real numbers differ by more than 1 in any digit they cannot be equal. Thus we are completely safe in our assertion. (An alternative way of avoiding this potential pitfall is to replace each digit by some different digit chosen from the range  $\{1, 2, \dots, 8\}$ .)

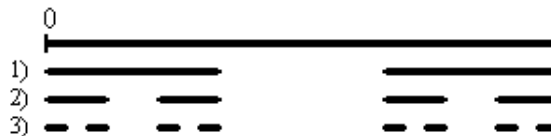
With Cantor's diagonalization method, we proved that  $\mathbb{R}$  is uncountable. What happens if we apply the same method to  $\mathbb{Q}$ , in a (futile) attempt to show the rationals are uncountable? Well, suppose for contradiction that our bijective function  $f : \mathbb{N} \rightarrow \mathbb{Q}[0, 1]$  produces the following mapping:

$$\begin{aligned} f(0) &= 0.\textcircled{1}4\ 0\ 0\ 0\ \dots \\ f(1) &= 0.5\textcircled{2}2\ 4\ 5\ \dots \\ f(2) &= 0.2\ 1\textcircled{4}2\ 1\ \dots \\ &\vdots \qquad \qquad \qquad \vdots \end{aligned}$$

This time, let us consider the number  $q$  obtained by modifying every digit of the diagonal, say by replacing each digit  $d$  with  $d + 2 \pmod{10}$ . Then in the above example  $q = 0.316\dots$ , and we want to try to show that it does not occur in our infinite list of rational numbers. However, we do not know that  $q$  is rational (in fact, it is extremely unlikely for the decimal expansion of  $q$  to be periodic). This is why the method fails when applied to the rationals. When dealing with the reals, the modified diagonal number was guaranteed to be a real number.

## The Cantor Set

The Cantor set is a remarkable set construction involving the real numbers in the interval  $[0, 1]$ . The set is defined by repeatedly removing the middle thirds of line segments infinitely many times, starting with the original interval. For example, the first iteration would involve the removal of the (open) interval  $(\frac{1}{3}, \frac{2}{3})$ , leaving  $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ . We then proceed to remove the middle third of each of these two remaining intervals, and so on. The first three iterations are illustrated below:



The Cantor set contains all points that have *not* been removed:  $C = \{x : x \text{ not removed}\}$ . How much of the original unit interval is left after this process is repeated infinitely? Well, we start with an interval of length 1, and after the first iteration we remove  $\frac{1}{3}$  of it, leaving us with  $\frac{2}{3}$ . For the second iteration, we keep  $\frac{2}{3} \times \frac{2}{3}$  of the original interval. As we repeat these iterations infinitely often, we are left with:

$$1 \longrightarrow \frac{2}{3} \longrightarrow \frac{2}{3} \times \frac{2}{3} \longrightarrow \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} \longrightarrow \dots \longrightarrow \lim_{n \rightarrow \infty} \left(\frac{2}{3}\right)^n = 0$$

According to the calculations, we have removed everything from the original interval! Does this mean that the Cantor set is empty? No, it doesn't. What it means is that the *measure* of the Cantor set is zero; the Cantor set consists of isolated points and does not contain any non-trivial intervals. In fact, not only is the Cantor set non-empty, it is uncountable!<sup>2</sup>

<sup>2</sup>It's actually easy to see that  $C$  contains at least countably many points, namely the endpoints of the intervals in the construction—i.e., numbers such as  $\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{1}{27}$  etc. It's less obvious that  $C$  also contains various other points, such as  $\frac{1}{4}$  and  $\frac{3}{10}$ . (Why?)

To see why, let us first make a few observations about ternary strings. In ternary notation, all strings consist of digits (called “trits”) from the set  $\{0, 1, 2\}$ . All real numbers in the interval  $[0, 1]$  can be written in ternary notation. (E.g.,  $\frac{1}{3}$  can be written as 0.1, or equivalently as 0.0222..., and  $\frac{2}{3}$  can be written as 0.2 or as 0.1222....) Thus, in the first iteration, the middle third removed contains all ternary numbers of the form 0.1xxxxx. The ternary numbers left after the first removal can all be expressed either in the form 0.0xxxxx... or 0.2xxxxx... (We have to be a little careful here with the endpoints of the intervals; but we can handle them by writing  $\frac{1}{3}$  as 0.02222... and  $\frac{2}{3}$  as 0.2.) The second iteration removes ternary numbers of the form 0.01xxxxx and 0.21xxxxx (i.e., any number with 1 in the second position). The third iteration removes 1’s in the third position, and so on. Therefore, what remains is all ternary numbers with only 0’s and 2’s. Thus we have shown that

$$C = \{x \in [0, 1] : x \text{ has a ternary representation consisting only of 0's and 2's}\}.$$

Finally, using this characterization, we can set up an *onto* map  $f$  from  $C$  to  $[0, 1]$ . Since we already know that  $[0, 1]$  is uncountable, this implies that  $C$  is uncountable also. The map  $f$  is defined as follows: for  $x \in C$ ,  $f(x)$  is defined as the binary decimal obtained by dividing each digit of the ternary representation of  $x$  by 2. Thus, for example, if  $x = 0.0220$  (in ternary), then  $f(x)$  is the binary decimal 0.0110. But the set of all binary decimals 0.xxxxx... is in 1-1 correspondence with the real interval  $[0, 1]$ , and the map  $f$  is onto because every binary decimal is the image of some ternary string under  $f$  (obtained by doubling every binary digit).<sup>3</sup> This completes the proof that  $C$  is uncountable.

## Power Sets and Higher Orders of Infinity

Let  $S$  be any set. Then the *power set* of  $S$ , denoted by  $\mathcal{P}(S)$ , is the set of all subsets of  $S$ . More formally, it is defined as:  $\mathcal{P}(S) = \{T : T \subseteq S\}$ . For example, if  $S = \{1, 2, 3\}$ , then  $\mathcal{P}(S) = \{\{\}, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$ .

What is the cardinality of  $\mathcal{P}(S)$ ? If  $|S| = k$  is finite, then  $|\mathcal{P}(S)| = 2^k$ . To see this, let us think of each subset of  $S$  corresponding to a  $k$  bit string, where a 1 in the  $i$ th position indicates that the  $i$ th element of  $S$  is in the subset, and a 0 indicates that it is not. In the example above, the subset  $\{1, 3\}$  corresponds to the string 101. Now the number of binary strings of length  $k$  is  $2^k$ , since there are two choices for each bit position. Thus  $|\mathcal{P}(S)| = 2^k$ . So for finite sets  $S$ , the cardinality of the power set of  $S$  is exponentially larger than the cardinality of  $S$ . What about infinite (countable) sets? We claim that there is no bijection from  $S$  to  $\mathcal{P}(S)$ , so  $\mathcal{P}(S)$  is not countable. Thus for example the set of all subsets of natural numbers is not countable, even though the set of natural numbers itself is countable.

**Theorem:**  $|\mathcal{P}(\mathbb{N})| > |\mathbb{N}|$ .

**Proof:** Suppose towards a contradiction that there is a bijection  $f : \mathbb{N} \rightarrow \mathcal{P}(\mathbb{N})$ . Recall that we can represent a subset by a binary string, with one bit for each element of  $\mathbb{N}$ . (So, since  $\mathbb{N}$  is infinite, the string will be infinitely long. Contrast the case of  $\{0, 1\}^*$  discussed earlier, which consists of all binary strings of *finite* length.) Consider the following diagonalization picture in which the function  $f$  maps natural numbers  $x$  to binary strings which correspond to subsets of  $\mathbb{N}$ :

<sup>3</sup>Note that  $f$  is *not* injective; for example, the ternary strings 0.20222... and 0.22 map to binary strings 0.10111... and 0.11 respectively, which denote the same real number. Thus  $f$  is not a bijection. However, the current proof shows that the cardinality of  $C$  is at least that of  $[0, 1]$ , while it is obvious that the cardinality of  $C$  is at most that of  $[0, 1]$  since  $C \subset [0, 1]$ . Hence  $C$  has the same cardinality as  $[0, 1]$  (and as  $\mathbb{R}$ ).

	0	1	2	3	4	5	...
0	1	0	0	0	0	0	...
1	0	1	0	0	0	0	...
2	1	0	1	0	0	0	...
3							
⋮							

In this example, we have assigned the following mapping:  $0 \rightarrow \{0\}$ ,  $1 \rightarrow \{\}$ ,  $2 \rightarrow \{0,2\}$ , ... (i.e., the  $n$ th row describes the  $n$ th subset as follows: if there is a 1 in the  $k$ th column, then  $k$  is in this subset, else it is not.) Using a similar diagonalization argument to the earlier one, flip each bit along the diagonal:  $1 \rightarrow 0$ ,  $0 \rightarrow 1$ , and let  $b$  denote the resulting binary string. First, we must show that the new element is a subset of  $\mathbb{N}$ . Clearly it is, since  $b$  is an infinite binary string which corresponds to a subset of  $\mathbb{N}$ . Now suppose  $b$  were the  $n$ th binary string. This cannot be the case though, since the  $n$ th bit of  $b$  differs from the  $n$ th bit of the diagonal (the bits are flipped). So it's not on our list, but it should be, since we assumed that the list enumerated all possible subsets of  $\mathbb{N}$ . Thus we have a contradiction, implying that  $\mathcal{P}(\mathbb{N})$  is uncountable.

Thus we have seen that the cardinality of  $\mathcal{P}(\mathbb{N})$  (the power set of the natural numbers) is strictly larger than the cardinality of  $\mathbb{N}$  itself. The cardinality of  $\mathbb{N}$  is denoted  $\aleph_0$  (pronounced "aleph null"), while that of  $\mathcal{P}(\mathbb{N})$  is denoted  $2^{\aleph_0}$ . It turns out that in fact  $\mathcal{P}(\mathbb{N})$  has the same cardinality as  $\mathbb{R}$  (the real numbers), and indeed as the real numbers in  $[0, 1]$ . This cardinality is known as  $\mathfrak{c}$ , the "cardinality of the continuum." So we know that  $2^{\aleph_0} = \mathfrak{c} > \aleph_0$ . Even larger infinite cardinalities (or "orders of infinity"), denoted  $\aleph_1, \aleph_2, \dots$ , can be defined using the machinery of set theory; these obey (to the uninitiated somewhat bizarre) rules of arithmetic. Several fundamental questions in modern mathematics concern these objects. For example, the famous "continuum hypothesis" asserts that  $\mathfrak{c} = \aleph_1$  (which is equivalent to saying that there are no sets with cardinality between that of the natural numbers and that of the real numbers).