To Infinity And Beyond: Countability and Computability

This note ties together two topics that might seem like they have nothing to do with each other. The nature of infinity (and more particularly, the distinction between different levels of infinity) and the fundamental nature of computation and proof. This note can only scratch the surface — if you want to understand this material more deeply, there are wonderful courses in the Math department as well as CS172 and graduate courses like EECS229A that will connect this material to the nature of information and compression as well.

Bijections

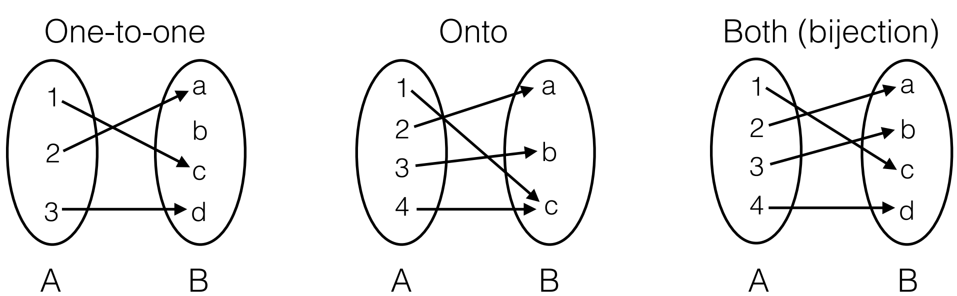
Two finite sets have the same size if and only if their elements can be paired up, so that each element of one set has a unique partner in the other set, and vice versa. We formalize this through the concept of a *bijection*.

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

EECS 70 离散数学与概率论 2020秋季

Note 11

超越无穷：可数性与可计算性

本讲义将两个看似毫无关联的主题联系在一起：无穷的本质（特别是不同层次无穷之间的区别）以及计算和证明的基本性质。本讲义只能触及表面——如果你想更深入理解这些内容，数学系有许多精彩课程，还有 CS172 以及像 EECS229A 这样的研究生课程，会将这些内容与信息和压缩的本质联系起来。

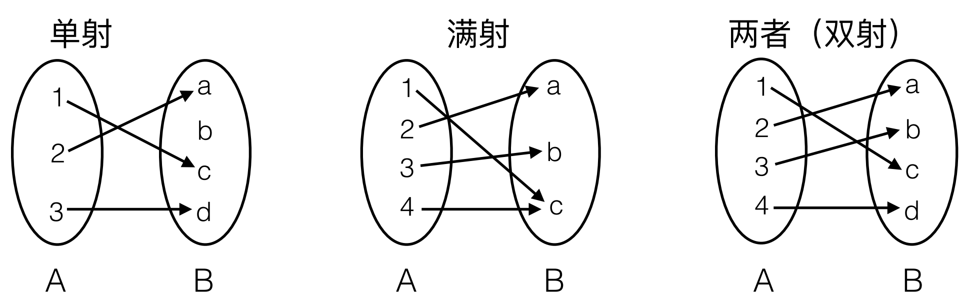
双射

两个有限集合大小相同，当且仅当它们的元素可以一一配对，使得一个集合中的每个元素在另一个集合中有唯一的伙伴，反之亦然。我们通过双射的概念来形式化这一点。考虑一个函数（或映射）f，它将集合 A 中的元素（称为 f 的定义域）映射到集合 B 中的元素（称为 f 的值域）。由于 f 是一个函数，它必须为每个元素 $x \in A$（输入）指定唯一一个元素 $f(x) \in B$（输出）。我们记作 $f : A \to B$。我们说 f 是双射，如果每个元素 $a \in A$ 都有唯一的像 $b = f(a) \in B$，并且每个元素 $b \in B$ 都有唯一的原像 $a \in A : f(a) = b$。

如果 f 将不同的输入映射到不同的输出，则称 f 是单射函数（或称内射）。更严格地说，f 是单射需满足：$x \neq y \Rightarrow f(x) \neq f(y)$。

如果 f 覆盖了值域中的每一个元素（即值域中的每个元素至少有一个原像），则称 f 是满射（或称上映射）。更准确地说，函数 f 是满射需满足：$(\forall y\ \exists x)(f(x) = y)$。

下面是一些简单的例子，帮助可视化单射、满射函数以及双射：



注意，根据上述定义，当且仅当 f 既是单射又是满射时，f 才是双射。

练习。设 $f : A \to B$ 是一个双射。证明 f 存在一个逆函数 $f^{-1} : B \to A$，满足对所有 $a \in A$ 都有 $f^{-1}(f(a)) = a$，并且 f^{-1} 也是一个双射。

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:

$$\begin{array}{ccccccccc}
\mathbb{N} & 0 & 1 & 2 & 3 & 4 & 5 & \ldots & \\
f\downarrow & \searrow & \searrow & \searrow & \searrow & \searrow & \searrow & & \\
\mathbb{Z}^+ & & 1 & 2 & 3 & 4 & 5 & 6 & \ldots
\end{array}$$

Why is this mapping a bijection? Clearly, the function $f : \mathbb{N} \to \mathbb{Z}^+$ 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 $2\mathbb{N} = \{0, 2, 4, 6, \ldots\}$? 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:

$$\begin{array}{cccccccc}
\mathbb{N} & 0 & 1 & 2 & 3 & 4 & 5 & \ldots \\
f\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\
2\mathbb{N} & 0 & 2 & 4 & 6 & 8 & 10 & \ldots
\end{array}$$

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:

基数

我们如何判断两个集合是否具有相同的基数（或大小）？这个问题的答案令人欣慰地回到了小学的记忆：通过展示两个集合元素之间的配对。更正式地说，我们需要在两个集合之间展示一个双射 f。这个双射建立了两个集合元素之间的一一对应关系或配对。我们上面已经看到了有限集合的情况。在本讲的其余部分，我们将了解这对无限集合意味着什么。

我们关于无限集合的第一个问题是：自然数集 N 比正整数集 Z+ 更大吗？很容易会回答“是”，因为每个正整数都是自然数，但自然数集多了一个额外的元素 0 /∈Z+。然而更仔细观察后，我们发现可以在自然数和正整数之间定义如下映射：

$$
\begin{array}{lccccccc}
\mathbb{N} & 0 & 1 & 2 & 3 & 4 & 5 & \dots \\
f\downarrow & \searrow & \searrow & \searrow & \searrow & \searrow & \searrow & \\
\mathbb{Z}^+ & & 1 & 2 & 3 & 4 & 5 & 6 \quad \dots
\end{array}
$$

为什么这个映射是双射？显然，函数 f : N →Z+ 是满射，因为每个正整数都被覆盖了。它也是单射，因为没有两个自然数具有相同的像。（n 的像是 f(n) = n+1，所以如果 f(n) = f(m)，那么必须有 n = m。）因为我们已经建立了 N 和 Z+ 之间的双射，这说明自然数和正整数一样多！（非常）非正式地，我们证明了 ∞+1 = ∞。

偶数自然数集 2N = {0,2,4,6,...} 呢？在前面的例子中，差别只是一个元素。但在这个例子中，自然数的数量似乎是偶数自然数的两倍。显然，N 的基数应该比 2N 大，因为 N 还包含了所有奇数自然数？尽管这看起来更困难，但我们尝试用以下映射在两个集合之间建立一个双射：

$$
\begin{array}{lccccccc}
\mathbb{N} & 0 & 1 & 2 & 3 & 4 & 5 & \dots \\
f\downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\
2\mathbb{N} & 0 & 2 & 4 & 6 & 8 & 10 & \dots
\end{array}
$$

这个例子中的映射也是一个双射。f 显然是单射，因为不同的自然数被映射到不同的偶数自然数（因为 f(n) = 2n）。f 也是满射，因为范围内的每个 n 都能被覆盖：它的原像是 $\frac{n}{2}$。由于我们在两个集合之间找到了一个双射，这告诉我们实际上 N 和 2N 具有相同的基数！

整数集 Z 呢？乍一看，似乎整数集比自然数集大，因为它包含了无限多个负数。然而，事实证明，我们可以在两个集合之间建立一个双射，这意味着这两个集合大小相同！考虑以下映射 f：

$$0 \to 0,\ 1 \to -1,\ 2 \to 1,\ 3 \to -2,\ 4 \to 2,\ \ldots,\ 124 \to 62,\ \ldots$$

In other words, our function is defined as follows:

$$f(x) = \begin{cases} \frac{x}{2}, & \text{if } x \text{ is even} \\ \frac{-(x+1)}{2}, & \text{if } x \text{ is odd} \end{cases}$$

We will prove that this function $f : \mathbb{N} \to \mathbb{Z}$ is a bijection, by first showing that it is one-to-one and then showing that it is onto.

Proof (one-to-one): Suppose $f(x) = f(y)$. Then they both must have the same sign. Therefore either $f(x) = \frac{x}{2}$ and $f(y) = \frac{y}{2}$, or $f(x) = \frac{-(x+1)}{2}$ and $f(y) = \frac{-(y+1)}{2}$. In the first case, $f(x) = f(y) \Rightarrow \frac{x}{2} = \frac{y}{2} \Rightarrow x = y$. Hence $x = y$. In the second case, $f(x) = f(y) \Rightarrow \frac{-(x+1)}{2} = \frac{-(y+1)}{2} \Rightarrow x = y$. So in both cases $f(x) = f(y) \Rightarrow x = y$, so f is injective.

Proof (onto): If $y \in \mathbb{Z}$ is non-negative, then $f(2y) = y$. Therefore, y has a pre-image. If y is negative, then $f(-(2y+1)) = y$. Therefore, y has a pre-image. Thus every $y \in \mathbb{Z}$ has a preimage, so f is onto.

Since f is a bijection, this tells us that $\mathbb{N}$ and $\mathbb{Z}$ have the same size.

Now for an important definition. We say that a set S is **countable** if there is a bijection between S and $\mathbb{N}$ or some subset of $\mathbb{N}$. Thus any finite set S is countable (since there is a bijection between S and the subset $\{0, 1, 2, \ldots, m-1\}$, where $m = |S|$ is the size of S). And we have already seen three examples of countable infinite sets: $\mathbb{Z}^+$ and $2\mathbb{N}$ are obviously countable since they are themselves subsets of $\mathbb{N}$; and $\mathbb{Z}$ is countable because we have just seen a bijection between it and $\mathbb{N}$.

What about the set of all rational numbers? Recall that $\mathbb{Q} = \{\frac{x}{y} \mid x, y \in \mathbb{Z},\ y \neq 0\}$. Surely there are more rational numbers than natural numbers? After all, there are infinitely many rational numbers between any two natural numbers. Surprisingly, the two sets have the same cardinality! To see this, let us introduce a slightly different way of comparing the cardinality of two sets.

If there is a one-to-one function $f : A \to B$, then the cardinality of A is less than or equal to that of B. Now to show that the cardinality of A and B are the same we can show that $|A| \leq |B|$ and $|B| \leq |A|$. This corresponds to showing that there is a one-to-one function $f : A \to B$ and a one-to-one function $g : B \to A$. The existence of these two one-to-one functions implies that there is a bijection $h : A \to B$, thus showing that A and B have the same cardinality. The proof of this fact, which is called the Cantor-Bernstein theorem, is actually quite hard, and we will skip it here.

Back to comparing the natural numbers and the integers. First it is obvious that $|\mathbb{N}| \leq |\mathbb{Q}|$ because $\mathbb{N} \subseteq \mathbb{Q}$. So our goal now is to prove that also $|\mathbb{Q}| \leq |\mathbb{N}|$. To do this, we must exhibit an injection $f : \mathbb{Q} \to \mathbb{N}$. The following picture of a spiral conveys the idea of this injection:

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?).

$$0 \to 0,\ 1 \to -1,\ 2 \to 1,\ 3 \to -2,\ 4 \to 2,\ \ldots,\ 124 \to 62,\ \ldots$$

换句话说，我们的函数定义如下：

$$f(x) = \begin{cases} \frac{x}{2}, \text{如果 x 是偶数} \\ \frac{-(x+1)}{2} \text{如果 x 是奇数} \end{cases}$$

我们将证明这个函数 f : N →Z 是一个双射，首先证明它是单射，然后证明它是满射。

证明（单射）：假设 f(x) = f(y)。那么它们必须具有相同的符号。因此要么 f(x) = $\frac{x}{2}$ 且 f(y) = $\frac{y}{2}$, 或 f(x) = $\frac{-(x+1)}{2}$ 且 f(y) = $\frac{-(y+1)}{2}$。在第一种情况下，f(x) = f(y) ⇒$\frac{x}{2} = \frac{y}{2}$ ⇒x = y。因此 x = y。在第二种情况下，f(x) = f(y) ⇒$\frac{-(x+1)}{2} = \frac{-(y+1)}{2}$ ⇒x = y。所以在两种情况下 f(x) = f(y) ⇒x = y, 所以 f 是单射。

证明（满射）：如果 y ∈Z 是非负的，则 f(2y) = y。因此 y 有原像。如果 y 是负数，则 f(−(2y+1)) = y。因此 y 有原像。于是每个 y ∈Z 都有原像，所以 f 是满射。

由于 f 是双射，这告诉我们 N 和 Z 具有相同的大小。

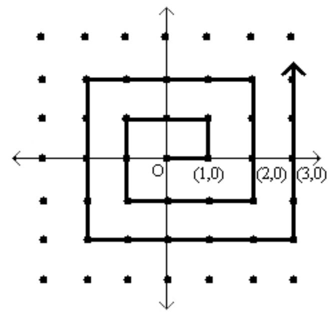
现在是一个重要的定义。我们说一个集合 S 是可数的，如果在 S 与 N 或 N 的某个子集之间存在一个双射。因此任何有限集合 S 都是可数的（因为在 S 与子集 {0,1,2,...,m−1} 之间存在双射，其中 m = |S| 是 S 的大小）。我们已经看到了三个可数无限集合的例子：Z+ 和 2N 显然是可数的，因为它们本身就是 N 的子集；而 Z 是可数的，因为我们刚刚看到了它与 N 之间的双射。

那么所有有理数的集合呢？回忆一下 Q = {x $\bar{y}$ | x,y ∈Z, y ≠ 0}。当然有更多的有理数比自然数多吗？毕竟，任意两个自然数之间都有无穷多个有理数。令人惊讶的是，这两个集合具有相同的基数！为了看到这一点，让我们引入一种稍微不同的比较集合基数的方法。

如果存在一个一对一函数 f : A →B，则 A 的基数小于或等于 B 的基数。现在要证明 A 和 B 的基数相同，我们可以证明 |A| ≤|B| 且 |B| ≤|A|。这相当于证明存在一个一对一函数 f : A →B 和一个一对一函数 g : B →A。这两个单射的存在意味着存在一个双射 h : A →B，从而表明 A 和 B 具有相同的基数。这个事实的证明被称为康托尔–伯恩斯坦定理，实际上相当困难，我们这里略过。

回到比较自然数和整数的问题。首先显然 |N| ≤|Q| 因为 N ⊆Q。所以我们现在的目标是证明 |Q| ≤|N|。为此，我们必须构造一个单射 f : Q →N。下面这个螺旋图传达了这一单射的思想：

每个有理数 ab（以最简形式书写，使得 gcd(a,b) = 1）由无限二维网格中的点 (a,b) 表示（对应于 Z ×Z，即所有整数对的集合）。注意并非网格上的所有点都有效表示有理数：例如，x 轴上的所有点都有 b = 0，因此都不合法（除了 (0,0)，我们用它表示有理数 0）；像 (2,8) 和 (−1,−4) 这样的点也不合法，因为有理数 14 由 (1,4) 表示。但 Z×Z 当然包含了所有有理数的这种表示，因此如果我们能构造一个从 Z×Z 到 N 的单射，那它也将是从 Q 到 N 的单射（为什么？）。



The idea is to map each pair (a,b) to its position along the spiral, starting at the origin. (Thus, e.g.,$(0,0) \to 0$, $(1,0) \to 1$, $(1,1) \to 2$, $(0,1) \to 3$, 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 $f : \{0,1\}^* \to \mathbb{N}$ 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 nth 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

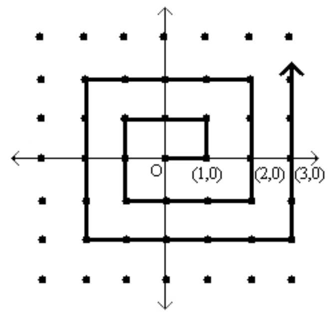
$$\varepsilon, 0, 1, 00, 01, 10, 11, 000, 001, 010, 011, 100, 101, 110, 111, 1000, \ldots,$$

where ε 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 $f : \mathbb{N}(x) \to \{0,1,2\}^*$, 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:

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 $f : \mathbb{N}(x) \to \{0,1,2\}^*$ is not onto (and hence not a bijection) since many ternary strings will not be



思路是将每一对 (a,b) 映射到螺旋线上的位置，从原点开始。（例如，(0,0) →0, (1,0) →1, (1,1) →2, (0,1) →3，依此类推。）显然，这种映射将每一对整数单射地映射到一个自然数，因为每一对占据螺旋线上唯一的位置。

这告诉我们 |Q| ⩽|N|。又因为我们之前观察到 |N| ⩽|Q|，根据康托尔–伯恩斯坦定理，N 和 Q 具有相同的基数。

练习：证明所有自然数有序对的集合 N×N 是可数的。

我们的下一个例子涉及所有有限长度的二进制字符串集合，记作 {0,1}*。尽管这个集合包含长度无界的字符串，但它实际上与 N 具有相同的基数。为此，我们建立一个直接的双射 f : {0,1}*→N 如下。注意到只要以每个字符串恰好出现一次的方式枚举 {0,1}* 中的元素即可。然后我们通过令 f(n) 等于列表中第 n 个字符串来得到双射。我们如何枚举 {0,1}* 中的字符串？很自然地按长度递增排序，然后在相同长度内按字典序（或等价地，视为二进制数时的数值升序）排列。这意味着列表看起来像

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

其中 ε 表示空字符串（唯一长度为 0 的字符串）。显然这个列表包含了每个二进制字符串一次且仅一次，因此我们如愿得到了与 N 的双射。

我们最后一个可数的例子是所有系数为自然数的多项式构成的集合，记作 N(x)。为了说明这个集合是可数的，我们将利用（前一个例子的某种变体）。首先注意到，与 {0,1}*的情况基本相同，我们可以看出所有三进制字符串的集合 {0,1,2}*（即字母表 {0,1,2} 上的字符串）是可数的。为了说明 N(x) 是可数的，只需构造一个单射 f : N(x) →{0,1,2}*，这反过来就能给出从 N(x) 到 N 的单射。（显然存在从 N 到 N(x) 的单射，因为每个自然数 n 本身平凡地就是一个多项式，即常数多项式 n。）

我们如何定义 f？先来看一个例子，即多项式 p(x) = 5x5+2x4+7x3+4x+6。我们可以列出 p(x) 的系数如下：(5,2,7,0,4,6)。然后我们可以把这些系数写成二进制字符串：(101,10,111,0,100,110)。现在，我们可以构造一个三进制字符串，其中在每个二进制系数之间插入一个 2 作为分隔符（忽略值为 0 的系数）。因此我们把 p(x) 映射到一个三进制字符串，如下图所示：

容易验证这是一个单射，因为原始多项式可以通过简单读取每对连续的 2 之间的系数唯一恢复出来。（注意这个映射 f : N(x) →{0,1,2}*不是满射（因此不是双射），因为许多三进制字符串不会是

$$5x^5 + 2x^4 + 7x^3 + 4x + 6$$
$$\downarrow$$
$$101210211221002110$$

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 $0.d_1d_2d_3\ldots$. In this representation, we write for example 1 as $0.999\ldots$[1], and 0.5 as $0.4999\ldots$. (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 $f : \mathbb{N} \to \mathbb{R}[0,1]$. Then, we can enumerate the real numbers in an infinite list $f(0), f(1), f(2), \ldots$ as follows:

$$\begin{aligned}
f(0) &= 0\,.\,5\,2\,1\,4\,9\,3\,5\,6\,\ldots \\
f(1) &= 0\,.\,1\,4\,1\,6\,2\,9\,8\,5\,\ldots \\
f(2) &= 0\,.\,9\,4\,7\,8\,2\,7\,1\,2\,\ldots \\
f(3) &= 0\,.\,5\,3\,0\,9\,8\,1\,7\,5\,\ldots \\
&\vdots \qquad\qquad \vdots
\end{aligned}$$

The number circled in the diagonal can be viewed as some real number $r = 0.5479\ldots$, 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 $d + 2 \bmod 10$; thus in our example above, $s = 0.7691\ldots$. 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^{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.

[1]To see this, write $x = .999\ldots$. Then $10x = 9.999\ldots$, so $9x = 9$, and thus $x = 1$.

$$5x^5 + 2x^4 + 7x^3 + 4x + 6$$
$$\downarrow$$
$$1012102111221002110$$

任何多项式的像；例如，对于包含带前导零二进制子序列的三进制字符串就会出现这种情况。

因此我们得到了从 N(x) 到 N 的一个单射，所以 N(x) 是可数的。

Cantors Diagonalization

我们已经确定 N, Z, Q 都具有相同的基数。那么 R，即实数集合呢？它们也应该是可数的吧？毕竟，有理数和实数一样都是稠密的（即在任意两个有理数 a,b 之间总存在另一个有理数，比如 a+b $\frac{}{2}$). 实际上，在任意两个实数之间那么在这两个实数之间总存在一个有理数。因此，尽管如此令人惊讶，实数的数量仍然多于有理数。也就是说，在有理数（或自然数）与实数之间不存在双射。我们现在就来证明这一点，使用康托尔提出的一个优美论证，称为对角化方法。事实上，我们将证明更强的结果：区间 [0,1] 内的实数是不可数的！

练习：说明如何在任意两个（不同的）实数之间找到一个有理数。

在准备证明之前，请回想一下，任何实数都可以唯一地写成一个没有尾随零的无限小数。特别地，区间 [0,1] 内的实数可以写作 0.d1d2d3.... 在
在这种表示中，我们例如将 1 写成 0.999...1，将 0.5 写成 0.4999...。（因此有理数总是表示为循环小数，而无理数则表示为非循环小数。对我们来说重要的是，所有这些表达式都是无限长且唯一的。）

定理：实数区间 R[0,1]（因而整个实数集 R）是不可数的。

证明：假设存在一个双射 f : N →R[0,1]，我们将导出矛盾。于是我们可以将 [0,1] 区间内的实数按无限列表 f(0), f(1), f(2),... 枚举如下：

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

对角线上圈出的数字可以看作某个实数 r = 0.5479...，因为它是一个无限小数展开。现在考虑通过修改 r 的每一位数字得到的实数 s，例如将每一位数字 d 替换为 (d + 2) mod 10；因此在上面的例子中，s = 0.7691...。我们断言 s 不出现在我们这个无限的实数列表中。假设相反，它确实出现过，并且是列表中的第 n 个数，即 f(n)。但根据构造，s 与 f(n) 在第 (n+1) 位不同，因此这两个数不可能相等！所以我们构造了一个不在 f 值域中的实数 s。但这与 f 是双射的说法矛盾。因此实数是不可数的。

1要理解这一点，设 x = .999.... 那么 10x = 9.999...，因此 9x = 9，从而 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... $= 1.000....$ 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 $\{1, 2, \dots, 8\}$.)

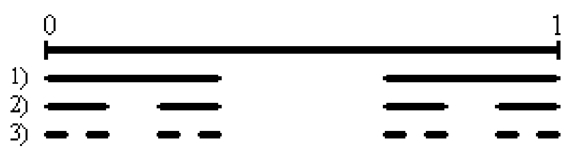
With Cantor's diagonalization method, we proved that $\mathbb{R}$ is uncountable. What happens if we apply the same method to $\mathbb{Q}$, in a (futile) attempt to show the rationals are uncountable? Well, suppose for contradiction that our bijective function $f : \mathbb{N} \to \mathbb{Q}[0,1]$ produces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

$$\begin{aligned} f(0) &= 0\,.\,1\;4\;0\;0\;0\;\dots \\ f(1) &= 0\,.\,5\;9\;2\;4\;5\;\dots \\ f(2) &= 0\,.\,2\;1\;4\;2\;1\;\dots \\ &\vdots \qquad\qquad \vdots \end{aligned}$$

This time, let us consider the number q obtained by modifying every digit of the diagonal, say by replacing each digit d with $d + 2$ mod 10. Then in the above example q = 0.316..., 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:

$$1 \longrightarrow \tfrac{2}{3} \longrightarrow \tfrac{2}{3} \times \tfrac{2}{3} \longrightarrow \tfrac{2}{3} \times \tfrac{2}{3} \times \tfrac{2}{3} \longrightarrow \cdots \longrightarrow \lim_{n \to \infty} (\tfrac{2}{3})^n = 0$$

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![2]

[2]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?)

让我们指出，之所以我们将每一位修改方式设为加2模10而不是加1，是因为同一个实数可能有两种小数展开；例如0.999... = 1.000...。但如果两个实数在某一位上相差超过1，则它们不可能相等。因此我们的断言是完全安全的。（避免这一潜在陷阱的另一种方法是将每一位替换为从范围{1,2,...,8}中选出的不同数字。）

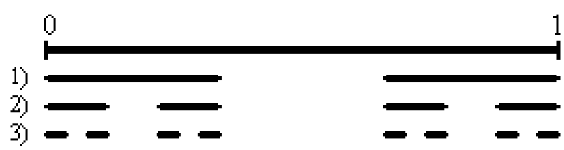
使用康托尔的对角化方法，我们证明了R是不可数的。如果我们尝试将同样的方法应用于Q，企图（徒劳地）证明有理数是不可数的，会发生什么？假设为了产生矛盾，我们的双射函数f：N →Q[0,1]产生了如下映射：

$$
\begin{array}{l}
f(0) = 0\,.\,\textcircled{1}\,4\;0\;0\;0\;\ldots \\
f(1) = 0\,.\,5\,\textcircled{9}\,2\;4\;5\;\ldots \\
f(2) = 0\,.\,2\;1\,\textcircled{4}\,2\;1\;\ldots \\
\vdots \qquad\qquad\quad \vdots
\end{array}
$$

这次，让我们考虑通过修改对角线上每一位得到的数q，比如说将每一位d替换为(d +2) mod 10。那么在上面的例子中q = 0.316...，我们试图证明它不会出现在我们无限的有理数列表中。然而，我们并不知道q是有理数（事实上，q的小数展开极不可能是周期性的）。这就是为什么这种方法应用于有理数时会失败。当处理实数时，修改后的对角线数保证是一个实数。

康托尔集

康托尔集是一个涉及[0,1]区间内实数的非凡集合构造。该集合通过从原始区间开始无限次重复移除线段的中间三分之一来定义。例如，第一次迭代将涉及移除开区间$(\frac{1}{3},\frac{2}{3})$，留下$[0,\frac{1}{3}]\cup[\frac{2}{3},1]$。然后我们继续移除这两个剩余区间各自的中间三分之一，，依此类推。前三次迭代如下图所示：



康托尔集包含所有未被移除的点：C = {x : x未被移除}。在此过程无限重复后，原始单位区间还剩下多少？我们从长度为1的区间开始，第一次迭代后我们移除了$\frac{1}{3}$它的部分，留下我们$\frac{2}{3}$。对于第二次迭代，我们保留$\frac{2}{3}\times\frac{2}{3}$原始区间的部分。当我们无限次重复这些迭代时，最终剩下的是：

$$1 \longrightarrow \tfrac{2}{3} \longrightarrow \tfrac{2}{3}\times\tfrac{2}{3} \longrightarrow \tfrac{2}{3}\times\tfrac{2}{3}\times\tfrac{2}{3} \longrightarrow \cdots \longrightarrow \lim_{n\to\infty}(\tfrac{2}{3})^n = 0$$

根据计算，我们已经从原始区间移除了所有内容！这是否意味着康托尔集为空？不，不是。这意味着康托尔集的测度为零；康托尔集由孤立点组成，不包含任何非平凡区间。事实上，不仅康托尔集非空，它还是不可数的！[2]

---

2 实际上很容易看出C至少包含可数多个点，即构造过程中区间的端点——例如像1, 23, 1, 127等数。不太明显的是C还包含其他各种点，比如1和 3 − 9 — 4 $\frac{3}{10}$。（为什么？）

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 $0.0222\ldots$, and $\frac{2}{3}$ can be written as 0.2 or as $0.1222\ldots$.) 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 $0.02222\ldots$ 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 $0.\text{xxxxx}\ldots$ 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).[3] 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 $\mathscr{P}(S)$, is the set of all subsets of S. More formally, it is defined as: $\mathscr{P}(S) = \{T : T \subseteq S\}$. For example, if $S = \{1,2,3\}$, then $\mathscr{P}(S) = \{\{\},\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\},\{1,2,3\}\}$.

What is the cardinality of $\mathscr{P}(S)$? If $|S| = k$ is finite, then $|\mathscr{P}(S)| = 2^k$. To see this, let us think of each subset of S corresponding to a k bit string, where a 1 in the ith position indicates that the ith 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 $|\mathscr{P}(S)| = 2^k$. 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 $\mathscr{P}(S)$, so $\mathscr{P}(S)$ 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: $|\mathscr{P}(\mathbb{N})| > |\mathbb{N}|$.

Proof: Suppose towards a contradiction that there is a bijection $f : \mathbb{N} \to \mathscr{P}(\mathbb{N})$. 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 $\mathbb{N}$:

[3] Note that f is *not* injective; for example, the ternary strings $0.20222\ldots$ and 0.22 map to binary strings $0.10111\ldots$ 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 $\mathbb{R}$).

要理解原因，让我们先对三进制字符串做一些观察。在三进制记法中，所有字符串由来自集合{0,1,2}的数字（称为三进制位）组成。区间[0,1]内的所有实数都可以用三进制记法表示。（例如，$\frac{1}{3}$ 可以写成0.1，或者等价地写成0.0222...，而$\frac{2}{3}$ 可以写成0.2 或者0.1222....）因此，在第一次迭代中，被移除的中间三分之一包含所有形如0.1xxxxx。第一次移除后剩下的三进制数都可以表示为0.0xxxxx... 或或0.2xxxxx...（这里我们必须稍微小心处理区间的端点；但我们可以通过将1写成$\frac{1}{3}$ 作为0.02222... 和 $\frac{2}{3}$ 作为0.2。）第二次迭代移除形如0.01xxxxx 和 0.21xxxxx（即第二位为1的任何数）。第三次迭代移除第三位为1的数，依此类推。因此剩下的就是所有仅包含0和2的三进制数。因此我们已经证明

$$C = \{x \in [0,1] : x\text{的三进制表示仅包含0和2}\}。$$

最后，利用这种表征，我们可以建立一个从C到[0,1]的满射f。因为我们已经知道[0,1]是不可数的，这意味着C也是不可数的。映射f定义如下：对于$x \in C$，f(x)定义为将x的三进制表示中每一位除以2后得到的二进制小数。因此，例如，如果x = 0.0220（三进制），则f(x)是二进制小数0.0110。但所有形如0.xxxxx...的二进制小数与实数区间[0,1]一一对应，且映射f是满射，因为每个二进制小数都是某个三进制字符串在f下的像（通过将每个二进制位加倍获得）。[3] 这就完成了C是不可数的证明。

幂集与更高阶的无穷

设S为任意集合。则S的幂集，记为P(S)，是S所有子集的集合。更正式地，它定义为：$P(S) = \{T : T \subseteq S\}$。例如，若S = {1,2,3}，则P(S) = {{},{1},{2},{3},{1,2},{1,3},{2,3},{1,2,3}}。

P(S)的基数是多少？如果|S| = k是有限的，则$|P(S)| = 2^k$。要理解这一点，让我们将S的每个子集对应于一个k位的二进制字符串，其中第i位上的1表示S的第i个元素在该子集中，0表示不在。在上面的例子中，子集{1,3}对应字符串101。现在长度为k的二进制字符串的数量是2^k，因为每一位有两个选择。因此$|P(S)| = 2^k$。所以对于有限集合S，S的幂集的基数比S本身的基数呈指数级更大。那么对于无限（可数）集合呢？我们声称从S到P(S)之间不存在双射，因此P(S)是不可数的。因此例如，所有自然数子集的集合是不可数的，即使自然数集合本身是可数的。

定理：|P(N)| > |N|。

证明：假设存在一个双射$f : N \to P(N)$，从而导致矛盾。回忆我们可以用一个二进制字符串来表示一个子集，每一位对应N的一个元素。（所以，由于N是无限的，该字符串将是无限长的。对比之前讨论过的$\{0,1\}^*$，它由所有有限长度的二进制字符串组成。）考虑如下对角化图示，其中函数f将自然数x映射到对应N子集的二进制字符串：

[3]注意f不是单射；例如，三进制字符串0.20222...和0.22分别映射到二进制字符串0.10111...和0.11，它们表示同一个实数。因此f不是双射。然而，当前证明表明C的基数至少等于[0,1]的基数，而由于$C \subset [0,1]$，显然C的基数至多等于[0,1]的基数。因此C与[0,1]（以及R）具有相同的基数。

$$
\begin{array}{c|cccccccc}
 & 0 & 1 & 2 & 3 & 4 & 5 & \dots \\
\hline
0 & 1 & 0 & 0 & 0 & 0 & 0 & \dots \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\
2 & 1 & 0 & 1 & 0 & 0 & 0 & \dots \\
3 & & & & & & & \\
\vdots & & & & & & &
\end{array}
$$

In this example, we have assigned the following mapping: $0 \to \{0\}$, $1 \to \{\}$, $2 \to \{0{,}2\}$, $\dots$ (i.e., the nth 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: $1 \to 0$, $0 \to 1$, 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 $\mathscr{P}(\mathbb{N})$ is uncountable.

Thus we have seen that the cardinality of $\mathscr{P}(\mathbb{N})$ (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 $\mathscr{P}(\mathbb{N})$ is denoted $2^{\aleph_0}$. It turns out that in fact $\mathscr{P}(\mathbb{N})$ has the same cardinality as $\mathbb{R}$ (the real numbers), and indeed as the real numbers in $[0,1]$. This cardinality is known as $\mathbf{c}$, the "cardinality of the continuum." So we know that $2^{\aleph_0} = \mathbf{c} > \aleph_0$. 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 $\mathbf{c} = \aleph_1$ (which is equivalent to saying that there are no sets with cardinality between that of the natural numbers and that of the real numbers).

$$
\begin{array}{c|cccccccc}
 & 0 & 1 & 2 & 3 & 4 & 5 & \ldots & \\
\hline
0 & 1 & 0 & 0 & 0 & 0 & 0 & \ldots \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & \ldots \\
2 & 1 & 0 & 1 & 0 & 0 & 0 & \ldots \\
3 & & & & & & & \\
\end{array}
$$

……

在这个例子中，我们分配了如下映射：0 →{0}, 1 →{}, 2 →{0,2}, ...（即第n行描述第n个子集如下：如果第k列有一个1，则k属于这个子集，否则不属于。）使用类似之前的对角化论证，将对角线上的每一位翻转：1 →0, 0 → 1，并让b表示得到的二进制字符串。首先，我们必须证明这个新元素是N的一个子集。显然它是，因为b是一个对应于N子集的无限二进制字符串。现在假设b是第n个二进制字符串。但这种情况不可能发生，因为b的第n位与对角线上的第n位不同（位被翻转了）。因此它不在我们的列表中，但它应该在，因为我们假设该列表枚举了N的所有可能子集。于是我们得到了一个矛盾，说明P(N)是不可数的。

因此我们已经看到，P(N)（自然数的幂集）的基数严格大于N本身的基数。N的基数记为ℵ0（读作阿列夫零），而P(N)的基数记为2ℵ0。事实上，P(N)与R（实数）具有相同的基数，甚至与[0,1]区间内的实数也具有相同的基数。这个基数被称为c，即连续统的基数。所以我们知道2ℵ0 = c > ℵ0。使用集合论的工具可以定义更大的无限基数（或无穷阶），记为ℵ1, ℵ2,...；它们遵循（对初学者来说有些奇怪）算术规则。现代数学中的几个基本问题涉及这些对象。例如，著名的连续统假设断言c = ℵ1（这等价于说不存在基数介于自然数和实数之间的集合）。