Lecture Notes

The topic of this lecture is infinite sets.

Infinite sets can be tricky to think about. For example, last time we saw the Pigeonhole Principle, which says that if a big set is mapped to a small set, then some two items of the big set must be mapped to the same item in the small set. We also saw that the Pigeonhole Principle appears to break down for infinite sets. In particular, the "big" set $\{0, 1, 2, 3, 4, \ldots\}$ of all natural numbers is mapped one-to-one to the "small" set $\{0, 2, 4, 6, 8, \ldots\}$ of even numbers by the function $f(n) = 2n$. This raises confusion about what is meant by "big" and "small" infinite sets. One of the key points of this lecture is sorting out the relative "bigness" of various infinite sets.

This subject is more abstract than others in 6.042. Nevertheless, the techniques we develop for dealing with infinite sets will be quite useful later on in 6.045 when you learn about computability.

1 Finite Sets vs. Infinite Sets

Let's start by distinguishing finite sets from infinite sets.

Definition A set is *finite* if there exists an $n \in \mathcal{N}$ such that $|S| = n$.

Definition A set is *infinite* if for all $n \in \mathcal{N}$, $|S| \geq n$.

These definitions imply that every set is either finite or infinite. For example, for all n, the set $\{0, 1, 2, 3, \ldots, n\}$ is finite. On the other hand, the set of all natural numbers, $\mathcal{N} = \{0, 1, 2, 3, \ldots\}$, is infinite. Here is a more interesting example:

Theorem 1.1 The set P of all primes is infinite.

Proof. The proof is by contradiction. Assume for the purpose of contradiction that P is finite. By the definition of a finite set, there exists an $n \in \mathcal{N}$ such that $|P| = n$. Denote the elements of P by $p_1, p_2, \ldots, p_n$. Now let $m = p_1 p_2 \cdots p_n + 1$. Since $m > p_i$ for all i, m cannot be a prime. Hence m must be divisible by some prime p_i:

$$\begin{aligned} & p_i \mid m \\ \Rightarrow\quad & p_i \mid p_1 p_2 \cdots p_n + 1 \\ \Rightarrow\quad & p_i \mid 1 \end{aligned}$$

The first implication follows by substituting the definition of m. The second follows because p_i divides the product $p_1 p_2 \cdots p_n$, and so must divide 1 in order to divide the sum. But no prime divides 1, so we have a contradiction. Therefore, there are an infinite number of primes. ■

Proving that the set of primes is infinite is relatively easy, but the next example shows that determining whether a set is finite or infinite can be tricky.

Definition A *twin prime* is a pair $(n, n+2)$ such that both n and $n+2$ are prime.

Examples of twin primes include $(3,5)$, $(5,7)$, $(11,13)$, $(17,19)$, and $(101,103)$. Is the set of twin primes finite or infinite? If you settle this question, then you can probably wheedle a few extra credit points from your TA. Mathematicians have tried unsuccessfully to solve this problem for more than *two thousand years*!

2 Countable Sets

A nice property of finite sets is that we can write out their elements in a list. For example, we can write out the set of integers between 1 and n as $\{1, 2, 3, \ldots, n\}$. We can also express some infinite sets in a list format:

$$\begin{aligned} \mathcal{N} &= \{0, 1, 2, 3, \ldots\} \\ \text{even numbers} &= \{0, 2, 4, 6, \ldots\} \\ \text{primes} &= \{2, 3, 5, 7, 11, \ldots\} \end{aligned}$$

Sets whose elements can be listed are called *countable sets.* The reason for the name is that once we list the elements, there is a first element, a second element, etc. We can then count through the items one after another so that every element of the set is eventually counted. Countable sets are formally defined in terms of bijections with the natural numbers.

Definition An infinite set S is said to be *countably infinite* iff there exists a bijection $f : \mathcal{N} \mapsto S$. A set is *countable* if it is finite or countably infinite.

Many familiar sets are countable: $\mathcal{N}$, the even numbers, primes, integers modulo a constant k, etc.

The formal definition of a countable set is equivalent to the notion of a listable set, since listing the elements of a set gives a bijection with $\mathcal{N}$ and vice-versa. If we can list the elements of S as $s_0, s_1, s_2, \ldots$, then we can construct a bijection $f : \mathcal{N} \mapsto S$ defined by $f(i) = s_i$. In the reverse direction, if we have such a bijection, then $f(0), f(1), f(2), \ldots$ is a list of all the elements of S.

Sometimes we can list the elements of a set S, but the corresponding bijection $f : \mathcal{N} \mapsto S$ is very hard to compute. For example, in principle we could list the elements in any subset of $\mathcal{N}$. However, most of these subsets are bizarre and difficult to describe; writing a program to compute $f(i)$ might be difficult or even impossible. The good news is that to prove that S is countable, we must only prove that a bijection $f : \mathcal{N} \mapsto S$ exists; we do not have to write f explicitly or to give an algorithm to compute $f(i)$.

3 Uncountable Sets

There would be little need for intuition and definitions about countable sets if all sets were countable; unfortunately, many important sets are not countable. Sets that are not countable are called *uncountable*.

Theorem The set $\mathcal{R}$ of real numbers is uncountable.

This theorem says that there is no way to make a list of all the real numbers. Naturally, one is immediately tempted to try anyway:

$$0,\ 1,\ 0.5,\ 0.75,\ 0.1,\ 0.05,\ 0.001,\ \ldots$$

This is not going to be so easy! For example, where are we going to stick in $\sqrt{2}$ and π^{-173}? Later we will see clever ways of listing the elements of some seemingly "big" sets like the rational numbers, but for the real numbers we can actually prove that no such listing exists.

We'll first work through the proof informally, and then write it up. The argument is by contradiction; we assume for the purpose of contradiction that $\mathcal{R}$ is countable. This means that there exists a bijection $f : \mathcal{N} \mapsto \mathcal{R}$. We can then make a list of all the real numbers that looks something like the one shown below. (The extra spaces only serve to highlight a diagonal sequence of digits discussed below.)

```
f(0) =  76. 5 430000000...
f(1) = 163.7 7 77777777...
f(2) =   2.78 7 8787878...
f(3) =  -1.112 1 231234...
f(4) =   3.1415 9 26535...
f(5) = -15.37273 2 3928...
        ...etc...
```

The plan is to reach a contradiction by constructing a real number that appears nowhere in the list; this will prove that f is not surjective and contradict the assumption that f is a bijection.

Consider the number defined by the diagonal digits in the above table, $0.577192\ldots$. Now suppose that we add 5 to every digit, working modulo 10. This gives $0.022647\ldots$; call this value r. This quantity r is totally bizarre, but it is a well-defined real number, since we have specified all the digits.

By our construction of r, $f(0)$ differs from r in the first position to the right of the decimal. Similarly, $f(1)$ differs from r in the second decimal position, $f(2)$ differs in the third position, and so forth. In general, $f(i)$ differs from r in the $(i+1)$-st position. So for all i, $f(i)$ is not equal to r. This implies that the real number r does not appear anywhere in our list!

There is one irritating problem in the preceding paragraph that must be fixed. Some numbers have two decimal representations. For example, $1 = 0.999\ldots$. Therefore, distinct decimal representations do not necessarily correspond to distinct numbers. Even though r disagrees with every $f(i)$

in at least one digit, the two might be equal. However, our construction ensures that r actually differs from $f(i)$ by at least $4 \cdot 10^{-(i+1)}$. For example, $1.0002999\ldots$ differs from $1.0007000\ldots$ by $4 \cdot 10^{-4}$.

Here is how the same argument might be written as a formal proof.

Proof. We must show that there is no bijection $f : \mathcal{N} \mapsto \mathcal{R}$. The proof is by contradiction. Assume for the purpose of contradiction that such a bijection f exists. Let r be the real number with decimal expansion $0.r_0r_1r_2r_3\ldots$ where r_i is 5 + the $(i+1)$-st decimal digit of $f(i)$, taken modulo 10. For all i, $f(i) \neq r$, since the $(i+1)$-st digits of $f(i)$ and r disagree by 5 and so $|r - f(i)| \geq 4 \cdot 10^{-(i+1)}$. This implies that f is not surjective, contradicting the assumption that f is a bijection. ■

This is called a "diagonalization argument", because of the construction based on diagonal entries. Diagonalization arguments are used all the time to prove that there is no bijection between two infinite sets. The generic idea is to assume for the purpose of contradiction that a bijection f exists. Then an element r of the range is constructed that differs a little bit from $f(\alpha)$ for every α in the domain. (In this case, for example, the real number r differs from $f(k)$ in the $(k+1)$-st decimal position for every natural k.) The existence of such an r shows that f is not surjective, contradicting the assumption that f is a bijection. We conclude that no bijection between the two sets exists.

Diagonalization arguments are important in logic and computer science. For example, this type of argument is used to show that there are some problems that can not be solved by a computer no matter how much time you give them. A similar method is used in logic to prove Gödel's Theorem, which says that there are some true theorems for which no proof exists!

4 Countable or Uncountable?

We now know that the natural numbers are countable and the real numbers are uncountable. In general, the first question about an infinite set is whether it is countable or uncountable. Let's answer this question for some simple sets.

The Integers

Theorem 4.1 The set $\mathcal{Z}$ of integers is countable.

Proof. To prove that that the integers are countable, we must prove that there exists a bijection $f : \mathcal{N} \mapsto \mathcal{Z}$. The integers can be listed as follows: 0, 1, -1, 2, -2, 3, -3, 4, -4, etc. This listing defines the desired bijection, e.g. $f(0) = 0$, $f(1) = 1$, $f(2) = -1$, $f(3) = 2$, etc. ■

This is the typical way to prove that a set is countable— just show that you can list all of the elements. This implicitly defines a bijection with $\mathcal{N}$. As noted above, we only have to prove that a bijection f exists; f need not be a function that we can actually compute. In this case, however, f can be written simply:

$$f = \begin{cases} -\frac{i}{2} & i \text{ even} \\ \frac{i+1}{2} & i \text{ odd} \end{cases}$$

The method of proving a set countable by listing elements has one potential pitfall. As an example, the following listing of the integers is incorrect: $0, -1, -2, -3, \ldots, 1, 2, 3, \ldots$. The problem is that the "$\ldots$" in the middle of the list represents a infinite number of elements. As a result, we never actually get around to listing the positive integers; in particular, there is no i such that $f(i) = 1$. This means that the function f defined by this listing is not a bijection.

The Half Integers

Define the set $Z^{1/2}$ as follows:

$$\begin{aligned} Z^{1/2} &= \{\frac{x}{2} \mid x \in \mathcal{Z}\} \\ &= \{0, 1, -1, 2, -2, \ldots\} \cup \{\frac{1}{2}, -\frac{1}{2}, \frac{3}{2}, -\frac{3}{2}, \ldots\} \end{aligned}$$

In other words, $Z^{1/2}$ consists of all integers and all half integers. This set is countable. We have written $Z^{1/2}$ above as the union of two countable sets. We can list all the elements of $Z^{1/2}$ by interleaving these two lists:

$$Z^{1/2} = \{0, \frac{1}{2}, 1, -\frac{1}{2}, -1, \frac{3}{2}, 2, -\frac{3}{2}, -2, \ldots\}$$

This trick generalizes to a theorem.

Theorem 4.2 If A and B are countable, then $A \cup B$ is countable.

Proof. Since A and B are countable, we can list the elements:

$$\begin{aligned} A &= \{a_1, a_2, a_3, \ldots\} \\ B &= \{b_1, b_2, b_3, \ldots\} \end{aligned}$$

Therefore, we can list the elements of the union as follows:

$$A \cup B = \{a_1, b_1, a_2, b_2, a_3, b_3, \ldots\}$$

There are two complications. First, sets A and B may contain common elements, so an element should be omitted from the list if it has already appeared. Second, either or both sets may be finite. If at some point all elements of one set have been listed, continue listing elements of the other set without interleaving. ∎

The Rationals

Recall that the set $\mathcal{Q}$ of rational numbers consists of all ratios of integers with non-zero denominator:

$$\mathcal{Q} = \{a/b \mid a, b \in \mathcal{Z}, b \neq 0\}$$

There are a lot of rational numbers! It is easy to imagine that we could not list all the rational numbers just as we could not list all the real numbers. In fact, however, there is a clever way to list all the rationals, and this proves that the set $\mathcal{Q}$ is countable.

Theorem 4.3 The set $\mathcal{Q}$ of rational numbers is countable.

Proof. *(Failed attempt.)* To prove that that the rational numbers are countable, we must prove that there exists a bijection $f : \mathcal{N} \mapsto \mathcal{Q}$. The rational numbers can be listed as follows:

$$0, \frac{1}{1}, \frac{2}{1}, \frac{3}{1}, \ldots, \frac{1}{2}, \frac{2}{2}, \frac{3}{2}, \ldots, \frac{1}{3}, \frac{2}{3}, \frac{3}{3}, \ldots$$

Omit from the listing all numbers that have already appeared. (For example, $2/2$ and $3/3$ should be omitted, since the rational number 1 already appeared as $1/1$.) This listing implicitly defines the desired bijection f. ∎

There are two problems with this proof, one minor and one serious. The minor problem is that we completely forgot the negative rational numbers. But the negative rationals are easy to add; we will show how in the correct proof given below. The killer problem is that we used "..." in the middle of the list to represent an infinite number of elements. This means that we will never reach, say, $1/2$; that is, there is no i such that $f(i) = 1/2$. This means that f is not a bijection as claimed. We need to list the rational numbers in a completely different order to make the proof work.

Proof. *(Successful attempt.)* We can list the positive rational numbers in the order indicated in Figure 1, omitting numbers that have already appeared in the list. This proves that the positive rationals are countable.

Negating each number in the figure establishes that the negative rationals are countable. The set of all rationals is the union of positive rationals, negative rationals, and $\{0\}$. Since these three sets are countable, two applications of Theorem 4.2 show that the set of all rationals is countable. ∎

The ordering technique used in the proof and indicated in Figure 1 is called "dovetailing". This technique is often used in 6.045. We can use dovetailing to prove that the product of any two countable sets is countable:

Theorem 4.4 If A and B are countable, then $A \times B$ is countable.

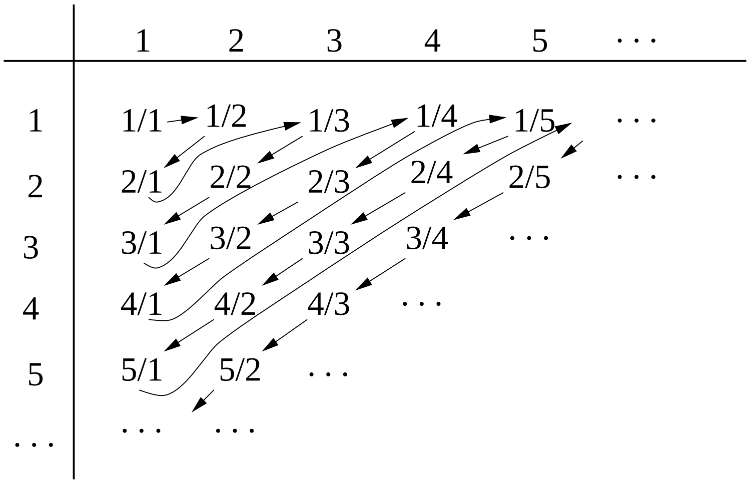

Figure 1: *This digram shows the ordering of rationals used in the proof of Theorem 4.3. This ordering technique is called "dovetailing".*

Proof. Since A and B are countable, we can list the elements of A as $\{a_1, a_2, a_3, \dots\}$ and the elements of B as $\{b_1, b_2, b_3, \dots\}$. We can obtain a listing of $A \times B$ by dovetailing:

$$\begin{aligned}
&(a_1, b_1),\\
&(a_1, b_2), (a_2, b_1),\\
&(a_1, b_3), (a_2, b_2), (a_3, b_1),\\
&(a_1, b_4), (a_2, b_3), (a_3, b_2), (a_4, b_1),\\
&\qquad\dots \qquad \dots
\end{aligned}$$

If either or both sets A and B are finite, omit from the listing all pairs containing non-existent elements a_i or b_j. ■

Note that the ordering of elements in this proof corresponds to the ordering in Figure 1. In fact, the proof that the rationals are countable used the idea of the above proof with $A = B = \mathcal{Z}^+$.

A Paradox

Here is a set that appears to be both countable and uncountable! This is impossible, so one argument must be incorrect, but which one?

Define S_n to be the set of n-digit decimal numbers. For example:

$$\begin{aligned}
S_1 &= \{.0, .1, .2, \dots, .9\}\\
S_2 &= \{.00, .01, .02, \dots, .99\}
\end{aligned}$$

Now let $S = \bigcup_{n=1}^{\infty} S_n$. Is S countable?

Argument that S is uncountable: We will prove that S is uncountable by the same diagonalization argument that we used to prove that the set $\mathcal{R}$ of real numbers is uncountable. Assume for the purpose of contradiction that we can list the elements of S in some order $s_0, s_1, s_2, \ldots$. Define $r = .r_0 r_1 r_2 \ldots$ where r_i is 5 + the $(i+1)$-st decimal digit of s_i taken modulo 10. Since S contains decimal numbers of every length, r must be in S. However, r differs from every element of S by at least one digit. Therefore, r is not in S, and we have a contradiction.

Argument that S is countable: We can list the elements of S by listing the elements of S_1, followed by S_2, followed by S_3, etc. Since each of these sets is finite, every set is fully listed eventually.

The first argument is incorrect, and the second argument is correct.

The false statement in the first argument is, "Since S contains decimal numbers of every length, r must be in S." The set S does contain decimal numbers of every *finite* length, but no numbers with an infinite number of non-zero digits to the right of the decimal. For example, the number $1/9 = 0.11111\ldots$ is not contained in any set S_i and therefore is not contained in the union S. As a result, there are many real numbers not in S. For this reason, the statement, "r must be in S" is false. This eliminates the contradiction from the proof by contradiction!

The second argument is correct; the union of a countable number of finite sets is always countable. In fact, for homework you will show that a union of a countable number of countable sets is always countable.

The distinction between countable and uncountable sets is so important that it distinguishes entire fields of mathematics. Discrete math is the study of countable sets; calculus, analysis, and continuous math all study uncountable sets. In this course, we will stick to discrete mathematics after today; all sets will be countable.

5 Other Infinite Sets

Two sets are considered to be the same size if there is a bijection between them. For example, the set of natural numbers is the same size as the set of rational numbers because there is a bijection $f : \mathcal{N} \mapsto \mathcal{Q}$. On the other hand, the set of natural numbers is not the same size as the set of real numbers because we proved there is no bijection between these sets.

More generally, we say that:

- $|A| \le |B|$ iff there exists an injection $f : A \mapsto B$
- $|A| \ge |B|$ iff there exists a surjection $f : A \mapsto B$
- $|A| = |B|$ iff there exists a bijection $f : A \mapsto B$
- $|A| < |B|$ iff $|A| \le |B|$ and $|A| \ne |B|$
- $|A| > |B|$ iff $|A| \ge |B|$ and $|A| \ne |B|$

For example, $|\mathcal{N}| < |\mathcal{R}|$. This fact follows from $|\mathcal{N}| \leq |\mathcal{R}|$ and $|\mathcal{N}| \neq |\mathcal{R}|$. The first statement holds because there is an injection from $\mathcal{N}$ to $\mathcal{R}$; namely, each natural number is mapped to itself. The second statement holds because we have proven that there is no bijection between $\mathcal{N}$ and $\mathcal{R}$.

In this sense, $\mathcal{N}$ and $\mathcal{R}$ are two different kinds of infinities. In fact, there are many types of infinities, and there is a whole branch of mathematics to deal with them. The most common types of infinities are based on power sets. (For now, we will denote the power set of S by $\mathcal{P}(S)$.) These types of infinity are denoted by the Hebrew letter aleph, $\aleph$, with a subscript. A bigger subscript implies a bigger type of infinity. Here are the first few:

$$\begin{aligned}
\aleph_0 &= |\mathcal{N}| \\
\aleph_1 &= |\mathcal{P}(\mathcal{N})| = |\mathcal{R}| \\
\aleph_2 &= |\mathcal{P}(\mathcal{P}(\mathcal{N}))| = |\mathcal{P}(\mathcal{R})| \\
&\dots
\end{aligned}$$

Each new type of infinity is obtained by taking the power set of the preceding type of infinity. (We will not prove the assertion on the second line that the set of real numbers is the same size as the power set of the natural numbers.) It is not immediately clear, however, that sets of size $\aleph_2$ are really any larger than sets of size $\aleph_1$, or more generally that taking the power set of S gives a larger set than S itself. This fact can be proven, however.

Theorem 5.1 For all sets S, $|S| < |\mathcal{P}(S)|$.

This implies that $\aleph_0 < \aleph_1 < \aleph_2 < \dots$, and so there is a whole infinite hierarchy of infinite sets! The proof of the theorem is really convoluted; the main idea is drawn from "Russel's Paradox".

Proof. We have to prove that $|S| \leq |\mathcal{P}(S)|$ and $|S| \neq |\mathcal{P}(S)|$.

$|S| \leq |\mathcal{P}(S)|$: We must show that there exists an injection $f : S \mapsto \mathcal{P}(S)$. The function $f(\alpha) = \{\alpha\}$ for all $\alpha \in S$ is just such an injection.

$|S| \neq |\mathcal{P}(S)|$: The proof is by contradiction. Assume for the purpose of contradiction that $|S| = |\mathcal{P}(S)|$. Then there is a surjection $f : S \mapsto \mathcal{P}(S)$. (In fact, there is a bijection, but we do need this fact.) To show a contradiction, we will define an item in $\mathcal{P}(S)$ that is not covered by f; this will contradict the assertion that f is a surjection. In particular, let $R = \{a \mid a \in S, a \notin f(a)\}$. The set R is a subset of S, and so R is an element of the power set of S. Since f is a surjection, there must be an $a \in S$ such that $f(a) = R$. Now there are two cases: either $a \in R$ or $a \notin R$.

($a \in R$): By the definition of the set R, this implies that $a \notin f(a) = R$, a contradiction.

($a \notin R$): Now the definition of the set R implies that $a \in f(a) = R$, another contradiction.

The contradictions in both cases imply that our assumption was wrong; therefore, $|S| \neq |\mathcal{P}(S)|$ and so $|S| < |\mathcal{P}(S)|$. ■

We can construct many types of infinity using power sets, but what about completely different types of infinity? For example, is there an infinite set S that is larger than the natural numbers, but smaller than the real numbers? This question was posed by Cantor in the 1870's and was not resolved for almost a century.

In 1900 David Hilbert posed 24 unsolved problems for the new century and this was first on his list. Hilbert conjectured that no such set existed, and this conjecture became known as the Continuum Hypothesis.

In 1963, Paul Cohen answered the question. He did not prove that such a set existed or did not exist, but rather showed that *both* scenarios are consistent with the axioms of set theory. As a consequence, if the axioms of set theory are consistent, then there is no way to prove the Continuum Hypothesis true or false!