Day 27: Tuesday, 4/25/2017
Return Homework. Take Questions on 9.2.

Recap:
1. We lump together sets that are linked by a bijection, saying that they have equal cardinality.
2. Sets that have equal cardinality with $\mathbb{N}$ are described as countably infinite, and we say they have cardinality $\aleph_0$.
3. Some examples of countably infinite sets: $\mathbb{N}$, $2\mathbb{N}$, $a\mathbb{N}$ for any natural number a, the integers, the even integers, equivalence classes modulo n for any n, the positive rationals, and all the rationals.
4. Unions of countably many countably infinite sets are countably infinite. The Cartesian product of finitely many countably infinite sets is countably infinite.
5. And yet, there are sets that infinite but not countably infinite.

Overview of today's material:
1. The real numbers is infinite but not countably infinite.
2. For any infinite set A, the power set $\wp(A)$ has a different cardinality from A.
3. We can define Card(A) ≤ Card(B) provided A is equivalent to some subset of B. Then we say that Card(A) < Card(B) provided Card(A) ≤ Card(B) but Card(A) ≠ Card(B).
4. With this definition, we see that for any set A, Card(A) < Card($\wp(A)$).
5. If Card(A) ≤ Card(B) and Card(B) ≤ Card(A) does that imply Card(A) = Card(B)? That seems like a plausible conclusion, but it is not self evident from the definitions. Nevertheless, this is a famous result from set theory, the Cantor-Schröder-Bernstein Theorem. It is a powerful tool for proving two sets are equivalent.

Cardinality of the interval (0,1) is uncountable.
1. This set is infinite because for example, $f(x) = x/2$ is a bijection onto the proper subset (0,1/2).
2. We have to show there is no bijection from $\mathbb{N}$ to (0,1). We take an indirect approach. We prove: For any function $f : \mathbb{N} \to (0,1)$, range(f) ≠ (0,1). Thus f is not a bijection. Since this is true for all such functions, no bijection from $\mathbb{N}$ to (0,1) can exist.
3. To show that range(f) ≠ (0,1), it suffices to show the existence of a number z in (0,1) that is not equal to $f(n)$ for any n in $\mathbb{N}$.
4. Before doing so, a short digression regarding decimal expansions.
   a. every real number in (0,1) has a decimal expansion of the form
      $.d_1d_2d_3d_4 \ldots$
      where each d_j is a decimal digit.
   b. This is really a statement about infinite series. We interpret $.d_1d_2d_3d_4 \ldots$ to mean $\sum_{j=1}^{\infty} d_j 10^{-j}$.
   c. This form is valid for finite decimals – we simply tack on an infinite string of zeros after the last nonzero digit.

d. The infinite decimal .9999… is also equal to 1.000… More generally, any decimal expansion of the form $.d_1d_2d_3d_4 \ldots d_m9999\ldots$, where we assume that $d_m \neq 9$ is also equal to to the decimal $.d_1d_2d_3d_4 \ldots (d_m + 1)0000\ldots$. These results can be used by analyzing the infinite series $\sum_{j=1}^{\infty} 9 \cdot 10^{-j} = 9\sum_{j=1}^{\infty} 10^{-j} = .9\ \sum_{j=0}^{\infty}\ .1^j = .9\ \frac{1}{1-.1} = 1$.

e. We define a *normal* decimal expansion to be one that does *not* end in an infinite string of 9's. Then every number in (0,1) has a normal decimal expansion. Moreover, two numbers can have the same normal decimal expansion iff the two numbers are equal. That is, two normal decimal expansions $.d_1d_2d_3d_4 \ldots$ and $.e_1e_2e_3e_4 \ldots$ are equal iff for every natural number j, $d_j = e_j$.

5. Now we can proceed with our proof. Consider any function $f : \mathbb{N} \to (0,1)$, and for each natural number k, express $f(k)$ as a normal decimal expansion with the following notation:

   $f(k) = .d_{k:1}d_{k:2}d_{k:3}d_{k:4} \ldots$.

   That is, the jth digit of the normal decimal expansion of $f(k)$ is $d_{k:j}$.

6. Next we define the decimal expansion a number $z = .e_1e_2e_3e_4 \ldots$ according to this rule:

$$e_k = \begin{cases} 5, & d_{k:k} \neq 5 \\ 6, & d_{k:k} = 5 \end{cases}.$$

7. Now let k be any element of $\mathbb{N}$. Then the kth digit of z cannot equal the kth digit of $f(k)$, because by our rule when the kth digit of $f(k)$ is 5, the digit $e_k = 6$, and when the kth digit of $f(k)$ is not 5, then e_k is 5. Thus, z cannot be equal to $f(k)$. And since k is an arbitrary element of x, z is not in the range of f – it is not equal to $f(k)$ for any k in the domain.

8. Thus we have shown that the f under discussion is not a surjection, and hence not a bijection. Since f is an arbitrary function from $\mathbb{N}$ to (0,1), no such function can be a bijection. This proves that card((0,1)) $\neq$ card($\mathbb{N}$).

Cardinality of the continuum.

1. We know that card((0,1)) is different from $\aleph_0$. But we do not know (yet) whether it is the next infinite cardinal after $\aleph_0$. So let us denote this new cardinality with the symbol c.
2. We can easily see that every finite open interval (a, b) has the same cardinal as (0,1) by defining a linear function taking 0 to a and 1 to b. That will be a straight line through the points $(0, a)$ and $(1, b)$ in the xy plane, and so it has slope $b - a$ and y intercept a. The equation for this function is $f(x) = (b - a)x + a$.
3. As a specific example, take $a = -1$ and $b = 1$. Then the function is $f(x) = 2x - 1$, and we can see using methods of calculus that this is a bijection from the interval (0,1) to the interval (-1,1).
4. Now consider the function $g(x) = \frac{1}{(1-x)} - \frac{1}{x}$ with domain (0,1) and codomain $\mathbb{R}$. This is actually a bijection, as can be shows with methods of calculus. The derivative is positive for x in (0,1), so the function is increasing and therefore an injection. Considering the limiting values as x approaches 0 from above or 1 from below, also shows that f is a surjection. Thus, we see that (0,1) and $\mathbb{R}$ have the same cardinality. The symbol c stands for *continuum*, meaning the real line, and c is known as the cardinality of the continuum.

Cantor's Theorem: Card(A) ≠ Card($\wp(A)$).

1. If A is a finite set with cardinality n , then we know that card($\wp(A)$) = 2^n. So in this case Card(A) ≠ Card($\wp(A)$).

2. But what about for an infinite set? We know that an infinite set can be equivalent to a proper subset, so maybe it is possible for the power set of A to be equivalent to A.
3. NO: Cantor's theorem shows this. In fact, the proof is reminiscent of both the proof that (0,1) has a different cardinality than ℕ, and also the proof that there cannot be a set of all sets.
   a. Again, we take an indirect approach. We will prove, for any function $f : A \to \wp(A)$, f cannot be a surjection.
   b. Notice that we can think of f as a set of ordered pairs $(a, f(a))$ where a can be any *element* of A, and $f(a)$ must be some *subset* of A.
   c. In some cases, it may happen that $a \in f(a)$. And in other cases it might turn out that $a \notin f(a)$. Let's agree to call a an *inside element* if a is inside $f(a)$, and call a an *outside element* if a is not inside $f(a)$. For a given function f there may be no inside elements, or no outside elements.
   d. We use this idea to define a special subset of A, the outside set for f. That is, we define $Z = \{\text{all outside elements for } f\} = \{a \in A \mid a \notin f(a)\}$.
   e. Now we can show that Z is not in the range of f. We argue by contradiction. Suppose that Z IS in the range of f. Then there must be some element z in A such that $f(z) = Z$. Now there are two possibilities. Either z is an inside element or it is an outside element. But each of these possibilities leads to a contradiction. For, if z is an inside element, then $z \in f(z)$. But $f(z) = Z = \{\text{all outside elements for } f\}$. Thus, z must be an outside element. This contradicts the assumption that z is an inside element.

      On the other hand, if z is an outside element, then by definition, z is outside $f(z)$, or equivalently, $z \notin f(z)$. But $f(z) = Z = \{\text{all outside elements for } f\}$. Thus $z \notin \{\text{all outside elements for } f\}$ so z is not an outside element. That contradicts the assumption that z is an outside element.

      In either case, we reach a contradiction. This shows that it is impossible for Z to equal $f(z)$ for any z in A. That proves that f is not a surjection. And since f is an arbitrary function from A to $\wp(A)$, no such function can be a surjection. This shows that there is no bijection $f : A \to \wp(A)$. In particular, card(A) ≠ card($\wp(A)$).

4. This means we can form a chain of sets: ℕ, $\wp(\mathbb{N})$, $\wp(\wp(\mathbb{N}))$, $\wp(\wp(\wp(\mathbb{N})))$, etc. No two consecutive elements in the chain can have the same cardinality. And in fact these all have different cardinalities, though it requires a little work to prove that. Indeed, we can define an ordering on these infinite cardinals such that each set has a smaller cardinality than its own power set. Thus, we can conclude that there is an infinite chain of increasing cardinalities.

So there are infinitely many different sizes of infinite sets, and there is no largest size of infinite set. A powerful tool for developing these (and other) ideas in the Cantor-Schröder-Bernstein Theorem. We look at the next.

The Cantor-Schröder-Bernstein Theorem

1. Statement: Suppose A and B are sets, and that there exist injective functions $f : A \to B$ and $g : B \to A$. Then there exists a bijection $h : A \to B$.

2. An application: card((0,1)) = card((0,1]).
   a. This seems obvious but is difficult to prove by direct construction of a bijection
   b. The earlier proof we saw with countable sets, where we just stick in one additional element is not valid (without developing a lot more machinery) because we cannot create a roster for an uncountable set.
   c. To prove this using CSB theorem, define $f : (0,1) \to (0,1]$ with equation $f(x) = x$, so that f is the identity function. This is clearly an injection. Also, define g: $(0,1] \to (0,1)$ according to the equation $g(x) = x/2$. Again, this is easily seen to be an injection. Now the CSB theorem tells us that a bijection exists, although we are given no guidance in how to find one.
   d. Actually, understanding the proof of the CSB provides a way to actually construct the bijection, but it is a little complicated.
   e. Here is one bijection. We separate (0,1) into two sets. The first set is all the fractions of the form $1/n$ with n in ℕ. This set is {1/1, 1/2, 1/3, 1/4, 1/5, etc}. The other set is all the elements of (0,1) that are not in the first set. Now we can define a bijection h: $(0,1] \to (0,1)$ as follows.
$$h(x) = \begin{cases} x, & x \notin \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots\right\} \\ \frac{1}{n+1}, & x = \frac{1}{n} \text{ for some } n \in \mathbb{N} \end{cases}.$$
   f. This is a bijection, though you have to think about it a little bit.

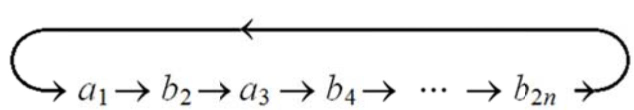
3. Suppose you have a cycle of injections like so:
$$A_0 \overset{f_0}{\to} A_1 \overset{f_1}{\to} A_2 \overset{f_2}{\to} \cdots \overset{f_{n-2}}{\longrightarrow} A_{n-1} \overset{f_n}{\to} A_0.$$
   Then, because compositions of injections are injections, we have an injection from any member of the cycle to any other member. For example, $A_0 \overset{f_0}{\to} A_1 \overset{f_1}{\to} A_2$ gives us an injection from A_0 to A_2, and $A_2 \overset{f_2}{\to} \cdots \overset{f_{n-2}}{\longrightarrow} A_{n-1} \overset{f_n}{\to} A_0$ gives us an injection from A_2 to A_0. By the CSB theorem, that proves there is a bijection from A_2 onto A_0, so card(A_2) = card(A_0).
4. We can apply this to the chain ℕ, ℘(ℕ), ℘(℘(ℕ)), ℘(℘(℘(ℕ))), etc. Note that every adjacent pair is of the form B, ℘(B). There is an obvious injection for any such pair, from B to ℘(B): for every b in B, define $f(b) = \{b\}$. That is, the image of any *element* b is the

subset $\{b\}$ containing just that one element. So in the chain $\mathbb{N}$, $\wp(\mathbb{N})$, $\wp(\wp(\mathbb{N}))$, $\wp(\wp(\wp(\mathbb{N})))$, etc, we have injections from each element to the next. Now suppose that two elements of the chain are equivalent. For example, suppose $\wp^{(5)}(\mathbb{N}) \approx \wp^{(13)}(\mathbb{N})$. Then there is a bijection g from $\wp^{(13)}(\mathbb{N})$ to $\wp^{(5)}(\mathbb{N})$, which in particular is an injection. Thus we have a cycle of injections as in the prior paragraph:

$$\wp^{(5)}(\mathbb{N}) \overset{f_5}{\to} \wp^{(6)}(\mathbb{N}) \overset{f_6}{\to} \wp^{(7)}(\mathbb{N}) \overset{f_7}{\to} \cdots \overset{f_{12}}{\to} \wp^{(13)}(\mathbb{N}) \overset{g}{\to} \wp^{(5)}(\mathbb{N}).$$

This shows that all of the sets in the chain are equivalent. In particular card($\wp^{(5)}(\mathbb{N})$) = card($\wp^{(6)}(\mathbb{N})$). But that contradicts Cantor's theorem, because the second set is the power set of the first. In this way, we argue that no two sets in the chain $\mathbb{N}$, $\wp(\mathbb{N})$, $\wp(\wp(\mathbb{N}))$, $\wp(\wp(\wp(\mathbb{N})))$, etc. can be equivalent. Therefore, we have an infinite collection of sets of different cardinalities.

5. Time permitting: a proof of the CSB theorem
    a. Assume A and B are sets, and that there are injections $f : A \to B$ and $g : B \to A$. Also, assume that A and B have no elements in common. (A minor modification of the proof below shows that the conclusion holds even when A and B do have elements in common.)
    b. We now define a collection of chains of elements of the set $A \cup B$. For any element a of A, we can apply f to get an element of B, then apply g to get an element of A, and so forth. Thus we create a chain of the form a, $f(a)$, $g(f(a))$, $f(g(f(a)))$, $g(f(g(f(a))))$, etc. Similarly, starting with a b in B, we can form the chain b, $g(b)$, $f(g(b))$, $g(f(g(b)))$, etc.
    c. Now define an equivalence relation on $A \cup B$, by defining two elements to be equivalent if they are both members of some chain. This relation is reflexive and symmetric by the way it is defined. It is also transitive, though that takes a little thought to show. But in any case, it is an equivalence relation.
    d. The equivalence classes are chains of maximal size: Each is a chain that is not a proper subsets of any other chain. We now show that these chains come in four different types.
        i. First there maybe cycles of the form a_1, b_2, a_3, b_4, $\cdots$, a_{2n-1}, b_{2n}, a_1. This type of chain occurs if we start with an element of A or B and after a finite number of steps we return to the starting point. It is also clearly an equivalence class, because all the elements are in the same chain with each other, and they cannot be in chains with any other elements. A chain of this first type can be shown in a diagram like so:



        ii. Second, there may be chains that start with a definite first element in A, and can be extended forward without ever repeating a prior value. For this kind of chain, the starting value a can be any element that is not in the range of g. Thus, there is no element we can place before a to extend the chain to the left. A chain of this type can be diagrammed in the form $a_1 \to b_2 \to a_3 \to b_4 \to \cdots$.

iii. Third, we may have a chain that begins with an element in *B* that is not in the range of *f*. This is similar to the prior case, but the starting point is in *B* rather than in *A*. A chain of this type can be diagrammed in the form $b_1 \to a_2 \to b_3 \to a_4 \to \cdots$.

iv. Finally, there may be chains that extend infinitely far in both directions. This type of chain occurs if we start with an initial value and propagate the chain forward without ever hitting a repetition, and also if every element in the chain is either in the range of *f* or in the range of *g*, so that from any element in the chain it is possible to propagate it to the left. Such a chain can be diagrammed as
$\cdots \to a_n \to b_{n+1} \to a_{n+2} \to b_{n+3} \to a_{n+4} \to \cdots$

e. Now these equivalence classes partition $A \cup B$, and in particular each element of *A* can be in only one such class. If *a* is in a class of type iii, then it is not the first element in the chain, and so we can find a unique preceding element *b* in the same chain. We define that *b* to be $h(a)$. In other words, for *a*'s in an equivalence class of type iii, the effect of *h* is to move one position along the chain from *a* in the opposite direction from the arrows. In all other cases, we define $h(a) = f(a)$. In these cases, the effect of *h* is to move along *a*'s chain one position in the same direction as the arrows. Since every element of *A* is in some chain, and since we have defined the effect of *h* on every chain, we have in fact defined a function from *A* to *B*. It is an injection because in each case we have only one choice for $h(a)$. To see that it is a surjection, consider any *b* in *B*. If *b* is in a class of type iii, then there is always an element *a* of *A* to the right of *b*, and $b = h(a)$ for that *a*. Thus every *b* in a type iii class is in the range of *h*. In all the other classes, every *b* has an element *a* to the left, and by definition, $b = h(a)$ for that *a*. This shows that *h* is a surjection, and hence a bijection.

End of Day