A Little Naïve Set Theory

To function in this course, you will need to know some set theory. In this review, I will gather the parts of the subject that you will need to have at your fingertips; I plan to discuss this handout during the first few classes. This handout should begin by defining the subject matter, but this is impossible to do: there is no universal agreement on the precise definition of the term *set*. To Georg Cantor, the principal founder of set theory, a set was a collection of things that could be thought of as a single whole, and it was clear to him that no further definition was needed. This point of view—it is what the word "naïve" in the title means—has turned out to be flawed, but it is the one I will take in this handout. (The topics we will cover in this course include a discussion of the flaws in Cantor's approach and the most widely accepted methods for dealing with them.)

I. Basic Operations.

I will use three examples to introduce the basic operations and vocabulary. Let

$$\begin{cases} G = \{1,2,3,4\}; \\ H = \{2,3,4,5\}; \\ K = \{5,6,7\}. \end{cases}$$

The things gathered together to form the set are the *elements* of the set; thus for example, set G has four elements, namely the integers 1, 2, 3, and 4. One says that thing x is an element of set X by saying "x is in X" or by writing "$x \in X$" and that thing x is not an element of set X by saying "x is not in X" or by writing "$x \notin X$."

Definition. Let $\{S_\alpha : \alpha \in A\}$ be a collection of sets. The *union* of the sets $\{S_\alpha\}$, denoted by "$\bigcup_{\alpha \in A} S_\alpha$" (or by something similar), is the set of all x's that are in at least one of the sets S_α. The *intersection* of the sets $\{S_\alpha\}$, denoted by "$\bigcap_{\alpha \in A} S_\alpha$" (or by something similar) is the set of all x's that are in all of the sets S_α. For example,

$$\begin{cases} G \bigcup H = \{1,2,3,4,5\}; \\ G \bigcap H = \{2,3,4\}. \end{cases}$$

Definition. The *empty set* is the set with no elements; it is denoted "$\{\}$" or "$\emptyset$." There is only one empty set, since it does matter what you leave out to get it. The empty set is a necessary construct if the pieces of the theory are to fit together properly; for example, $G \bigcap K = \emptyset$.[1]

Exercise 1. For each integer $n \geq 2$, let C_n be the closed interval $\left[0, \frac{1}{n}\right]$, and let O_n be the open interval $\left(0, \frac{1}{n}\right)$. Find $\bigcup_{n=2}^{\infty} C_n$; $\bigcup_{n=2}^{\infty} O_n$; $\bigcap_{n=2}^{\infty} C_n$; and $\bigcap_{n=2}^{\infty} O_n$.

Definition. Let X and Y be sets. One says that X is a *subset* of Y, and writes "$X \subseteq Y$," if every element of X is also an element of Y. (Note that the phrase "every element of X" does **not** imply that X must actually have elements. The empty set is in fact a subset of any given set.) You manufacture a subset of X of Y by choosing some of the elements of Y to make the set X. Note that "choosing some" can mean "choosing none" and can mean "choosing all."[2] Thus, for example, the set K has eight subsets, namely

$$\emptyset;\ \{5\};\ \{6\};\ \{7\};\ \{5,6\};\ \{5,7\};\ \{6,7\};\ \{5,6,7\}.$$

(How many subsets does $\emptyset$ have? How many proper subsets?)

[1] If $X \bigcap Y = \emptyset$, one says that X and Y are *disjoint*.
[2] If $X \subseteq Y$ but $X \neq Y$, one says that X is a *proper* subset of Y.

Exercise 2. Show that if a set B has n elements, then B has exactly 2^n subsets.

Definition. The *power set*[3] of a set X, denoted "$P(X)$", is the set of all subsets of X. Thus, for example,

$$P(K) = \big\{\emptyset, \{5\}, \{6\}, \{7\}, \{5,6\}, \{5,7\}, \{6,7\}, \{5,6,7\}\big\}.$$

Definition. Let X and Y be sets. The *difference* $X - Y$ is the set of all things that are in X but are not in Y; that is,

$$X - Y := \{x \in X : x \notin Y\}.$$

(Note that this operation is defined regardless whether or not $X \subseteq Y$. For example, $H - K = \{2,3,4\}$.)

Definition. Sometimes, the sets under consideration will all be subsets of a fixed "universal" set U. When this is the case, then for any given set $X \subseteq U$, one defines the *complement* $\overline{X}$ of X by $\overline{X} := U - X$.

Note that the definition of $\overline{X}$ depends on U as well as on X. For example, if $U = \{1,2,3,4,5,6,7\}$, then $\overline{G} = \{5,6,7\}$, whereas if $U = \{1,2,3,4,5,6,7,8\}$, then $\overline{G} = \{5,6,7,8\}$.

Definition. Let X and Y be sets. The *cartesian product* of X and Y, denoted "$X \times Y$," is the set of all ordered pairs (a,b), where $a \in X$ and $b \in Y$. Thus, for example,

$$H \times K = \{\mathbf{2},\mathbf{3},\mathbf{4},\mathbf{5}\} \times \{5,6,7\} = \left\{\begin{matrix} (\mathbf{2},5), & (\mathbf{2},6), & (\mathbf{2},7), \\ (\mathbf{3},5), & (\mathbf{3},6), & (\mathbf{3},7), \\ (\mathbf{4},5), & (\mathbf{4},6), & (\mathbf{4},7), \\ (\mathbf{5},5), & (\mathbf{5},6), & (\mathbf{2},5) \end{matrix}\right\}$$

Above, I have put the elements of H in boldface to emphasize that for this construction, the occurrence of 5 as an element of B is completely independent of its occurrence as an element of K. The reason for the word "product" in "cartesian product" should be obvious from the example. The reason for the word "cartesian" is that the first example ever constructed, namely the plane

$$\mathbf{R}^2 = \mathbf{R} \times \mathbf{R},$$

was constructed by Descartes.

II. A Few Particular Important Sets.

We will repeatedly need the following sets:

$$\begin{cases} \mathbf{N} = \{0,1,2,\ldots\} \text{ (the natural numbers);} \\ \mathbf{Z} = \{0,\pm 1,\pm 2,\ldots\} \text{ (the integers);} \\ \mathbf{Q} = \text{ the set of all fractions (rational numbers);} \\ \mathbf{R} = \text{ the set of all real numbers (both rational and irrational).} \end{cases}$$

III. Equivalence Relations and Partitions.

Definition. Let S be a nonempty set. A *binary relation* on S is a statement (perhaps true, perhaps false) about each element of $S \times S$. For example, let the set in question be $\mathbf{N}$, and let the statement—about any element $(x,y) \in \mathbf{N} \times \mathbf{N}$—be the statement "$x$ is less than y." Thus "3 is less than 4" is true, "4 is less than 4" is false, *etc.*When binary relation R is true of pair (x,y), one writes "xRy." There are several kinds of binary relations that are singled out for their importance, but the only kind I anticipate needing in this course is the so-called "equivalence relation." An equivalence relation, roughly speaking, is one whose description would include words like "equivalent" or "same." (Equality itself is the simplest example.) It turns out not to be too hard to pin down this idea precisely.

Definition. Let R be a binary relation on set S. R is an *equivalence relation* if it has the following three properties:

$$\begin{cases} [\mathbf{a}] : sRs, \text{ for all elements } s \in S \text{ ["}R\text{ is reflexive"];} \\ [\mathbf{b}] : \text{whenever } sRt \text{ then } tRs, \text{ for all elements } s,t \in S \text{ ["}R\text{ is symmetric"];} \\ [\mathbf{c}] : \text{whenever } sRt \text{ and } tRw \text{ then } sRw, \text{ for all elements } s,t,w \in S \text{ ["}R\text{ is transitive"].} \end{cases}$$

[3] The term "power set" comes from the "2^n" in Exercise 2.

An important binary relation from number theory is this one on $\mathbf{Z}$: for $x, y \in \mathbf{Z}$, say

$$xRy \iff x - y \text{ is a multiple of } 2.$$

(That is: either $x = y$, or you can get from x to y by a series of length-2 hops.)

Exercise 3. Show that R has all three properties **[a]**, **[b]**, and **[c]**.

Exercise 4. The properties **[a]**, **[b]**, and **[c]** are logically independent (in the sense that no two of them imply the third). Show this by demonstrating that each of the following three binary relations on $\mathbf{Z}$ is missing exactly one of the three properties and that no two of them are missing the same property.

$xR_1y \iff |x - y| \leq 2$

$xR_2y \iff x \leq y$

$xR_3y \iff x \leq 2 \;\&\; y \leq 2$

Exercise 5. Find the flaw in the following "proof" that any binary relation that is symmetric and transitive must also be reflexive.

> For any $a \in S$, I claim that aRa. Proof: Take any b such that aRb. Since R is symmetric, also bRa. But then since R is transitive, aRa. ∎

The most important fact about equivalence relations on a given set is that they correspond to what are called "partitions" of the set. Roughly speaking, a partition of a nonempty set is a way of breaking it apart into disjoint pieces. Here is the precise definition:

Definition. Let S be a nonempty set. A **partition** $\mathcal{P}$ of S is a collection $\mathcal{P} = \{A_\alpha\}$ of nonempty subsets of A with the property that each element of A is contained in exactly one of the subsets A_α. In other words:

[**i**] : $S = \bigcup_\alpha A_\alpha$; and

[**ii**] : If $A_{\alpha_1} \neq A_{\alpha_2}$, then $A_{\alpha_1} \bigcap A_{\alpha_2} = \emptyset$.

Now, any partition $\mathcal{P} = \{A_\alpha\}$ of S induces a binary relation on S, namely

$$xR_{\mathcal{P}}y \iff x \text{ and } y \text{ are in the same subset } A_\alpha, \qquad (*)$$

and it is easy to see that $R_{\mathcal{P}}$ is an equivalence relation. Theorem 1 establishes the interesting fact that these "$R_{\mathcal{P}}$" relations are the *only* equivalence relations on S; there are no others.

Theorem 1. *Let R be any equivalence relation on (nonempty) set S. There is a partition $\mathcal{P}$ of S for which $R = R_{\mathcal{P}}$.*

Proof. We are given R (reflexive, symmetric, and transitive); we need to find $\mathcal{P}$. Condition $(*)$ shows the way here: for each $a \in S$, put

$$[a] := \{x \in S \colon aRx\};$$

then let

$$\mathcal{P} := \{[a] \colon a \in S\}\,.$$

(The set $[a]$ is called the *equivalence class containing a*.) We first must show that this collection $\mathcal{P}$ is really is a partition—that is, that the sets $[a]$ are all nonempty and that **[i]** and **[ii]** are satisfied. Since $a \in [a]$,[4] each set $[a]$ is nonempty and $\bigcup_{a \in S} [a] = S$ (so that **[i]** is satisfied). To establish **[ii]**, which in this context says

$$\text{If } [a] \neq [b], \text{ then } [a] \bigcap [b] = \emptyset,$$

[4] This is the reflexive property of R.

we prove the logically equivalent assertion

$$\text{If } [a]\bigcap[b] \neq \emptyset, \text{ then } [a] = [b]. \tag{†}$$

The argument is tangled but easy. Suppose that $[a]\bigcap[b] \neq \emptyset$, and let z be any element in the intersection. Then

$$z \in [a] \Longrightarrow aRz \underset{\substack{\uparrow\\ \text{(symmetry of R)}}}{\Longrightarrow} zRa,$$

and

$$z \in [b] \Longrightarrow bRz \underset{\substack{\uparrow\\ \text{(symmetry of R)}}}{\Longrightarrow} zRb.$$

Next, since R is transitive, the above four relations yield two more:

$$bRz \text{ and } zRa \Longrightarrow bRa, \tag{1}$$

and

$$aRz \text{ and } zRb \Longrightarrow aRb. \tag{2}$$

I now **claim** that $[a] \subseteq [b]$ and $[b] \subseteq [a]$; proving the claim will establish (†) and complete the proof of that $\mathcal{P}$ is a partition.

Proof of claim. For all $t \in S$,

$$t \in [a] \Longrightarrow aRt \underset{\substack{\uparrow\\ \text{(transitivity and (1))}}}{\Longrightarrow} bRt \Longrightarrow t \in [b];$$

and

$$t \in [b] \Longrightarrow bRt \underset{\substack{\uparrow\\ \text{(transitivity and (2))}}}{\Longrightarrow} aRt \Longrightarrow t \in [a]. \ \blacksquare\textbf{(Claim)}$$

Finally, to see that $R = R_{\mathcal{P}}$ for this partition $\mathcal{P}$, observe that for all $a, b \in S$,

$$aRb \Longleftrightarrow [a] = [b] \Longleftrightarrow aR_{\mathcal{P}}b. \ \blacksquare$$

Exercise 6. The equivalence relation of exercise **3** generates two equivalence classes. What are they?

IV. Functions.

Definition. Let A and B be nonempty sets. A *function from A into B* is a rule that associates to each element a of A exactly one element b of B. If f is the name of the function, then one often writes "$f\colon A \to B$." For any element $a \in A$, the unique $b \in B$ that f associates to a is denoted "$f(a)$." A is called the *domain* of f, and B is called the *codomain* of f. The *range* of f is the set

$$\{f(a)\colon a \in A\} = \{b \in B\colon b = f(a) \text{ for at least one } a \in A\}.$$

(Obviously, the range of f must be a subset of B.)

Definition. Let A, B, and C be nonempty sets. If g is a function from A into B and f be a function from B into C, then the *composition* $f \circ g\colon A \to C$ is the function defined by

$$(f \circ g)(a) := f\big(g(a)\big) \text{ (for all } a \in A).$$

Definition. A function $f\colon A \to B$ is *one-to-one* (or is *injective*, or is an *injection*) if it has the property:

$$a_1 \neq a_2 \in A \Longrightarrow f(a_1) \neq f(a_2) \in B.$$

Definition. A function $f: A \to B$ is *onto* (or is *surjective*, or is a *surjection*) if the range of f is *all* of B; that is, for each $b \in B$, there is at least one $a \in A$ for which $f(a) = b$. (One can indicate that f is onto by saying that f is a function from A *onto* B.)

Definition. If a function $f: A \to B$ is both one-to-one and onto, one says that f is *bijective* (or is a *bijection*, or is a *one-to-one correspondence*).

Inverse function. When there is a bijection $f: A \to B$, then for each $b \in B$, there is *exactly one* $a \in A$ such that $f(a) = b$ (why?). One can then define the function $g: B \to A$ that retraces f from b back to a for each $b \in B$. g is called the *inverse* of f. The function g is also a bijection (why?), and f is the inverse of g. These functions satisfy the relations

$$\begin{cases} (g \circ f)(a) = a, & \text{all } a \in A, \text{ and} \\ (f \circ g)(b) = b, & \text{all } b \in B; \end{cases}$$

that is, each function undoes the other one.

Exercise 7. Let $g: A \to B$ and $f: B \to C$. Show that if f and g are both injective, then so is $f \circ g$.

Exercise 8. Let $g: A \to B$ and $f: B \to C$. Show that if f and g are both surjective, then so is $f \circ g$.

Exercise 9. Let $g: A \to B$ and $f: B \to C$. Show that if f and g are both bijective, then so is $f \circ g$.

V. Cardinality.

It requires concentration to understand and learn the set theory presented so far, but the content has been straightforward and has held no surprises. That state of affairs changes when the notion of cardinality (size) of a set is introduced. It turns out that it is possible to say precisely what the size of a set is, *regardless of whether that set is finite or infinite*. Moreover, there are bewilderingly many different sizes of infinite set. In this handout, I will not define this notion of "size" precisely; but I will discuss what it means for two infinite sets to be of the same size as well as what it means for one infinite set to be larger than another.

Observe first that there are two different ways to decide whether two *finite* sets are the same size. **First**, we can count the elements in each set and see whether we get the same answer for both. **Second**, we can see whether there is a bijection between the sets. The second method is quite different from the first; if we can find a bijection, we have learned that the sets are the same size, even though we may not know exactly what that size is. (For example: the number of five-element subsets of $\{1, 2, ..., 17\}$ equals the number of twelve-element subsets of $\{1, 2, ..., 17\}$, because we can correspond each five-element subset to its twelve-element complement.)

Now suppose we have two infinite sets, and we wish to know whether they are the same size—or if the question even makes sense. Indeed, the first method above makes no sense in this context; you can't count the elements of either set. The second method, however, makes perfect sense, and it is used to define the notion of "same size" for any two nonempty sets regardless whether finite or infinite:

Definition. Let X and Y be nonempty sets. We say that X *is equivalent to* Y (and write "$X \approx Y$") if there is a bijection $f: X \to Y$.

It is clear that if X and Y are finite sets, then "$X \approx Y$" means exactly that X and Y have the same number of elements. When X and Y are infinite, however, "$X \approx Y$" is a new concept, and one must be very careful not to assume that $\approx$ will have every property for infinite sets that it does for finite sets. I will investigate the behavior of $X \approx Y$ below; first, I need two more closely-related definitions.

Definition. Let X and Y be nonempty sets. We say that X *is no larger than* Y or that Y *is at least as large as* X (and write "$X \preceq Y$" or "$Y \succeq X$") if there is an injection from X into Y.

Since the composition of injections is injective, $\preceq$ is transitive. Furthermore, $X \preceq Y$ precisely if $X \approx W$ for some subset W of Y. The reason is as follows. If $X \preceq Y$, then the injection $f: X \to Y$ gives a bijection from X onto the range of f, so the range of f is the subset W. Conversely, if $X \approx W \subseteq Y$, then the bijection $X \leftrightarrow W$ is an injection $X \to Y$.

Definition. Let X and Y be nonempty sets. We say that X *is smaller than* Y or that Y *is larger than* X (and write "$X \prec Y$" or "$Y \succ X$") if there is an injection from X to Y but there is no bijection $X \leftrightarrow Y$.

Example. $\mathbf{N} \approx (\mathbf{N} - \{0\})$ by the bijection

$$\begin{array}{ccc} 0 & \longleftrightarrow & 1 \\ 1 & \longleftrightarrow & 2 \\ 2 & \longleftrightarrow & 3 \\ 3 & \longleftrightarrow & 4 \\ 4 & \longleftrightarrow & 5 \\ \vdots & & \vdots \end{array}$$

Thus, $\mathbf{N}$ is equivalent to a proper subset of itself.[5]

Exercise 10. Show that $\mathbf{N} \approx \mathbf{Z}$.

Exercise 11. Show that $\approx$ is reflexive, symmetric and transitive. That is, for nonempty sets X, Y, and Z: show that $X \approx X$; show that if $X \approx Y$, then $Y \approx X$; and show that if $X \approx Y$ and $Y \approx Z$, then $X \approx Z$.

Definition. An infinite set X is called *denumerable* if $\mathbf{N} \approx X$.[6] Note that an infinite set X is denumerable precisely if the elements of X can be listed $(x_0, x_1, x_2, \ldots)$, since this listing is another way of writing down the bijection

$$\begin{array}{ccc} 0 & \longleftrightarrow & x_0 \\ 1 & \longleftrightarrow & x_1 \\ 2 & \longleftrightarrow & x_2 \\ \vdots & & \vdots \end{array}$$

Exercise 12. Let Y be a denumerable set, and let X be an infinite subset of Y. Show that X is denumerable. (*Hint*: Since Y is denumerable, we can list the elements $(y_0, y_1, y_2, \ldots)$. Go along the list and look for the elements of X.)

Exercise **12** suggests that denumerable sets are the smallest infinite sets. Exercise **13** shows that this is indeed true. (Why? How does this assertion follow from that of exercise **13**?)

Exercise 13. Let Y be any infinite set. Show that Y has a subset W such that $\mathbf{N} \approx W$. (*Hint*: Argue that it is possible to pull a sequence of distinct elements $(y_0, y_1, y_2, \ldots)$ out of Y; then let W be the set of elements in the sequence.)

Exercise 14. Show that if X and Y are both denumerable, then $X \bigcup Y$ is also denumerable. (*Hint*: Obviously $X \bigcup Y$ is not finite. First: assume that $X \bigcap Y = \emptyset$ (that is, that there are no elements in common). Show that $(X \bigcup Y) \approx \mathbf{Z}$, and apply exercises **10** and **11**. Next: show that if $(X \bigcap Y) \neq \emptyset$, then $(X \bigcup Y)$ is equivalent to an infinite subset of $\mathbf{Z}$, and apply exercises **10**, **11** and **12**.)

Exercise 15. Let X be in infinite set. Show that X has a *proper* subset $\widehat{X}$ such that $X \approx \widehat{X}$. (*Hint*: The result you proved in Exercise **13** tells you where to start.)

Now consider the cartesian product of to denumerable sets (say, $\mathbf{N} \times \mathbf{N}$). You can picture the elements of this set as the points in the first quadrant with integer coordinates (or equivalently, as the set of boxes below).

---

[5] In fact, *every* infinite set is equivalent to a proper subset of itself. See Exercise **15**.

[6] Sometimes such sets are called *countable*. Some texts make a distinction between "denumerable" and "countable," allowing one of the two terms to refer to finite as well as to infinite sets. There is no universally accepted convention on this.

$(0,3)$	$(1,3)$	$(2,3)$	$(3,3)$	$\cdots$
$(0,2)$	$(1,2)$	$(2,2)$	$(3,2)$	$\cdots$
$(0,1)$	$(1,1)$	$(2,1)$	$(3,1)$	$\cdots$
$(0,0)$	$(1,0)$	$(2,0)$	$(3,0)$	$\cdots$

Is this set denumerable? Constructing a bijection between **N** and this set is the same as distributing the natural numbers $0, 1, 2, \cdots$ among the boxes in such a way that every box gets its own natural number. It is not immediately obvious how to do this. It is no good, for example, to distribute the natural numbers along the first row,

$(0,3)$	$(1,3)$	$(2,3)$	$(3,3)$	$\cdots$
$(0,2)$	$(1,2)$	$(2,2)$	$(3,2)$	$\cdots$
$(0,1)$	$(1,1)$	$(2,1)$	$(3,1)$	$\cdots$
0 $(0,0)$	1 $(1,0)$	2 $(2,0)$	3 $(3,0)$	$\cdots$ $\cdots$

because all the natural numbers wind up on the first row. (Above, we have an injection from **N** into $(\mathbf{N} \times \mathbf{N})$ which is not a surjection. If we were working with finite sets X and Y, we'd be done. As soon as *one* injection from finite set X to finite set Y fails to be onto, we know that $X \prec Y$, so that *all other* injections from X to Y also fail to be onto. For infinite sets, this is clearly not the case—*cf*. exercise **15**—so that we need to ask whether any other injection from **N** to $(\mathbf{N} \times \mathbf{N})$ might turn out to be a bijection. In fact, bijections $\mathbf{N} \leftrightarrow (\mathbf{N} \times \mathbf{N})$ do exist! Here is the one that Cantor found; he is distributing the natural numbers along lines of slope (-1).

6	$\cdots$			
$(0,3)$	$(1,3)$	$(2,3)$	$(3,3)$	$\cdots$
3	7	$\cdots$		
$(0,2)$	$(1,2)$	$(2,2)$	$(3,2)$	$\cdots$
1	4	8	$\cdots$	
$(0,1)$	$(1,1)$	$(2,1)$	$(3,1)$	$\cdots$
0	2	5	9	$\cdots$
$(0,0)$	$(1,0)$	$(2,0)$	$(3,0)$	$\cdots$

With this bijection, Cantor proved a result that surprised him:

Theorem 2. $(\mathbf{N} \times \mathbf{N}) \approx \mathbf{N}$. ∎

Theorem 2 has some important consequences. The first is Theorem **3**, which says that a denumerable union of denumerable sets is denumerable; a second is exercise **16**.

Theorem 3. *Let* $\{A_n : n = 0, 1, 2, \ldots\}$ *be a collection of denumerable sets. Then* $\bigcup_{n=0}^{\infty} A_n$ *is also denumerable.*

Proof. Distribute the elements of A_1 along the first row of boxes, and the elements of A_2 along the second row of boxes, *etc*. (This distributes the elements of the union among all the boxes in the first quadrant, but there will be a repeat whenever any element appears in more than one A_n.) If the sets $\{A_n\}$ are pairwise disjoint—that is, if $n \neq m \Longrightarrow A_n \bigcap A_m = \emptyset$—then there are no repeated elements in the boxes, so going through the boxes in Cantor's order will create a one-to-one correspondence $\mathbf{N} \leftrightarrow \bigcup_{n=1}^{\infty} A_n$. If the sets $\{A_n\}$ are not pairwise disjoint, then you can get the bijection by going through the boxes and writing down each element of the union *the first time you encounter it*, ignoring any repeat occurrences. ∎

Exercise 16. Show that $\mathbf{Q}$ is denumerable.

VI. Nondenumerable Infinite Sets.

The facts that $\mathbf{N} \approx \mathbf{Z}$ and that $\mathbf{N} \approx \mathbf{Q}$ may leave the impression that there is only one size of infinite set; after all, the set of rational numbers *looks* larger than the set of integers, since between any two rational numbers there are infinitely many others, so maybe infinitely big is just infinitely big, period. Cantor's most striking discovery is that this is not the case:

Theorem 4. *The set* $\mathbf{N}$ *is smaller than the set* $[0,1) = \{x \in \mathbf{R} : 0 \leq x < 1\}$.

Proof. The proof must show that there is an injection but no surjection from $\mathbf{N}$ to $[0,1)$. By exercise **13**, there are injections from $\mathbf{N}$ into $[0,1)$; so one must show is that every function $f : \mathbf{N} \to [0,1)$ fails to be a surjection. This is what Cantor did. Let f be any function from $\mathbf{N}$ into $[0,1)$. Since each real number on the interval $[0,1)$ has a unique representation as a base-ten decimal,[7] The function f can be represented by

[7] To make the representation unique, I stipulate that every decimal must contain infinitely many digits $d \neq 9$. Thus, for example, $\frac{3}{4}$ has only one representation, namely $0.75\overline{0}$; the representation $0.74\overline{9}$ is excluded.

a table

$$\begin{aligned} f(0) &= d_{00}d_{01}d_{02}d_{03}d_{04}\cdots \\ f(1) &= d_{10}d_{11}d_{12}d_{13}d_{14}\cdots \\ f(2) &= d_{20}d_{21}d_{22}d_{23}d_{24}\cdots \\ f(3) &= d_{30}d_{31}d_{32}d_{33}d_{34}\cdots \\ f(4) &= d_{40}d_{41}d_{42}d_{43}d_{44}\cdots \\ &\vdots \end{aligned}$$

Cantor showed that f is not a surjection by constructing an infinite decimal $x = 0.e_0e_1e_2e_3e_4\cdots$ that is not in the range of f. The construction goes as follows. Choose e_0 to be any digit that is different both from d_{00} and from 9; this will guarantee that $x \neq f(0)$, regardless what the rest of the digits are. Then choose e_1 to be any digit that is different both from d_{11} and from 9; choose e_2 to be any digit that is different both from d_{22} and from 9; *etc*. By the time you have chosen the k^{th} digit, you will have guaranteed that x is different from all of $f(0), f(1), \cdots, f(k)$. Thus, by the time you have chosen *all* of the digits, you will have guaranteed that x is different from $f(k)$ for *all* $k \in \mathbf{N}$—that is, that x is not in the range of f. ∎

Exercise 17. Show that the open interval $(0, 1)$ is equivalent to $\mathbf{R}$.

The argument above is the celebrated *diagonal argument* of Cantor. It is so named because it constucts x by examining the "diagonal" of the "matrix" of d's.) A natural question to ask now is whether there are any other sizes of infinite set besides the two we've already encountered. Cantor proved two further theorems, which together show that infinite sets can be incomprehensibly large. The first of these, Theorem 5, is proved by a generalization of the diagonal argument.

Theorem 5. *If X is any nonempty set, then $X \prec P(X)$.*

Proof. There is an obvious injection $g: X \to P(X)$, namely $g(x) = \{x\}$; so the proof must show that every function $f: X \to P(X)$ fails to be a surjection. So, let f be any function from X into $P(X)$. I will construct a subset T of X that is not in the range of f.

The construction is as follows. For any given $x \in X$, I will ensure that subset T is different from subset $f(x)$, by making them differ on whether they include the element x itself: that is, if $x \in f(x)$, then x will be excluded from T, and if $x \notin f(x)$, then x will be included in T. After I do this for every $x \in X$, I wind up with

$$T := \{x \in X : x \notin f(x)\},$$

and T is by construction different from $f(x)$ for each $x \in X$—that is, T is not in the range of f. ∎

Exercise 18. Let T be a denumerable subset of a nondenumerable set S. Show that $(S - T) \approx S$.

VII. Trichotomy.

Some properties of $\approx$ for finite sets are preserved for infinite sets (*e.g.* transitivity), while others are not preserved (*e.g.* the fact that no finite set is equivalent to a proper subset of itself). A very important property of $\approx$ on finite sets is the so-called *Trichotomy Principle*: given any two nonempty finite sets X and Y, exactly one of $\left\{\begin{matrix} X \prec Y \\ X \approx Y \\ X \succ Y \end{matrix}\right\}$ must hold. It turns out that this is true for infinite sets as well, although it is extremely hard to prove. The proof that at most one holds is the the content of the **Schröder-Bernstein Theorem**, which says that if there is an injection $f: X \to Y$ and an injection $g: Y \to X$, then there is a bijection $h: X \leftrightarrow Y$. This is not an easy theorem! Think for a minute about how you might go about constructing h from f and g.[8] The fact that at *least* one holds is even harder. The difficulty here is that you start out by knowing *nothing at all* about X and Y, except that they are nonempty sets. Where do you start?? You need to have proved quite a bit about the structure of all sets in general before you even have a starting place.

[8] I will append to this handout a particularly accessible proof of this theorem, which I found in an article by Diestel and Thomassen (*American Mathematical Monthly*, **113**(2006), 161–165.)

The Schröder-Bernstein Theorem is very useful. For example, consider the open interval $(0, 1)$ and the half-closed interval $[0, 1)$. Since the first interval is a subset of the second, the so-called "inclusion map" $x \mapsto x$ is an injection $(0, 1) \to [0, 1)$;[9] and it is easy to see that the map $f(x) = \frac{1}{2}x + \frac{1}{4}$ is an injection $[0, 1) \to \left[\frac{1}{4}, \frac{3}{4}\right) \subseteq (0, 1)$. By Schröder-Bernstein, then, $(0, 1) \approx [0, 1)$.

Of course, it is not too hard to come up with a direct bijection between $(0, 1)$ and $[0, 1)$; after all, these sets only differ by one element. But here, in outline, is a harder example; another appears in Exercise **22.** Consider the set "$C[0, 1]$," the set of all continuous real-valued functions on the closed interval $[0, 1]$. There are injections from $\mathbf{R}$ into $C[0, 1]$—for example, the one that maps each real number c to the constant function $f_c(x) \equiv c$, so

$$R \preceq C[0, 1];$$

and, because a continuous function is completely determined by its values on the rational numbers, there turns out to be an injection from $C[0, 1]$ into a set known to be equivalent to $\mathbf{R}$, so that

$$C[0, 1] \preceq \mathbf{R}.$$

By Schröder-Bernstein, then, $C[0, 1] \approx \mathbf{R}$; but it is by no means easy to find an explicit bijection between *these* two sets.

Theorem 5 gives us an infinite heirarchy of infinite sets, each one larger than the one before it:

$$\mathbf{N} \prec P(\mathbf{N}) \prec P(P(\mathbf{N})) \prec \cdots, \qquad (**)$$

the second of which, $P(\mathbf{N})$, turns out to be equivalent to $\mathbf{R}$. It gets increasingly difficult to imagine these sets, but Theorem 5 guarantees their existence.

So can we hope at least that the infinite heirarchy $(**)$ is able to span the possible sizes, so that every set X is $\preceq$ one of the sets in $(**)$? This turns out not to be the case. Cantor's second way of exhibiting an infinite set that is larger than those under consideration—Theorem 6—completely puts paid to this hope (or indeed to any hope of ever getting a handle on the complete range of possible sizes of infinite sets).

Theorem 6. *Say you have a collection $\mathcal{C}$ of sets with the property that $\mathcal{C}$ contains no largest set; that is, for each $X \in \mathcal{C}$ there is a set $Y \in \mathcal{C}$ such that $X \prec Y$. Then for each $X \in \mathcal{C}$,*

$$X \prec \bigcup_{Y \in \mathcal{C}} Y.$$

Proof. Suppose that the conclusion of the theorem were false. Then by the Trichotomy Principle, there would be some $X_0 \in \mathcal{C}$ such that $\bigcup_{Y \in \mathcal{C}} Y \preceq X_0$. By hypothesis, there also exists a set $Y_0 \in \mathcal{C}$ such that $X_0 \prec Y_0$. But since

$$X_0 \prec Y_0 \preceq \bigcup_{Y \in \mathcal{C}} Y \preceq X_0,$$

by transitivity of $\preceq$ we have both $X_0 \preceq Y_0$ and $Y_0 \preceq X_0$, so that by Schröder–Bernstein, $X_0 \approx Y_0$, contradiction. ∎

VIII. Objections and Paradoxes.

There is much more to Cantorian set theory, but perhaps you now have enough of it in front of you to begin to begin to mistrust it. Theorems 5 and 6 establish the "existence," in some sense, of sets that completely beggar the imagination. Since these sets are so hard to envision, and since mathematical objects find their existence in our thoughts, there is merit to the question: In exactly what sense do these large sets "exist" at all? Moreover, since the more-or-less tame beginnings of naïve theory lead by inexorable logic to its weird outer reaches, some mathematicians rejected the whole theory, saying that any infinite set, as a completed

[9] The inclusion map pins down the intuitive fact that whenever $X \subseteq Y$, then $X \preceq Y$.

whole, is meaningless (and hence nonexistent), because no one can hold it in his/her head in its entirety. Moreover, even those mathematicians who wanted to continue to develop and apply naïve set theory came to realize that at least some modifications were unavoidable: reasoning within the theory by the accepted rules of logic, they began to discover that in its outer reaches, the theory was plagued with internal inconsistencies. Two of the principal inconsistencies are Cantor's Paradox and Russell's Paradox, both of which arise from reasoning about the set of all sets.

Cantor's Paradox. Let $\mathcal{S}$ be the set of all sets. By Theorem 5, we have $\mathcal{S} \prec P(\mathcal{S})$; but on the other hand, since the elements of $P(\mathcal{S})$ are sets, $P(\mathcal{S}) \subseteq \mathcal{S}$, so that $P(\mathcal{S}) \preceq \mathcal{S}$. This contradicts the Trichotomy Principle.

Russell's Paradox is very similar to this one.[10] It goes as follows. For any given set X, either X is an element of itself ($X \in X$) or X is not an element of itself ($X \notin X$). For example, the set T of all conceivable thoughts is itself a conceivable thought, so $T \in T$; but the set $\mathbf{N}$ of natural numbers is not itself a natural number, so $\mathbf{N} \notin \mathbf{N}$. So, said Russell, let $\mathcal{N}$ (for "normal") be the set of all sets that are not elements of themselves; that is,

$$\mathcal{N} := \{X \in \mathcal{S}: X \notin X\}\,. \tag{††}$$

Then consider the question:

Is $\mathcal{N}$ an element of $\mathcal{N}$?

Suppose first that $\mathcal{N} \in \mathcal{N}$. Then $\mathcal{N}$ satisfies the entrance criterion "$X \notin X$" mentioned in definition (††) of $\mathcal{N}$; in other words, $\mathcal{N} \notin \mathcal{N}$. On the other hand, suppose that $\mathcal{N} \notin \mathcal{N}$. Then $\mathcal{N}$ is one of those sets X such that $X \notin X$; and since $\mathcal{N}$ is the set of *all* such X, we must have $\mathcal{N} \in \mathcal{N}$.

Neither of the above paradoxes reflects an error of reasoning. They show that naïve set theory is internally inconsistent. This certainly added fuel to the fire of the objectors. You will will learn this semester how classical mathematicians attempted to fix the problems and fend off the objectors, and you will be able to judge for yourself how successful these attempts were.

Exercise 19. Show that it follows from The Schröder-Bernstein theorem that for nonempty sets X and Y, $\left\{\begin{matrix} X \prec Y \\ X \succ Y \end{matrix}\right\}$ cannot both hold.

Exercise 20. Let $\mathcal{SQ}$ denote the set of all infinite sequences of 0's an 1's. Show that $\mathbf{R} \sim \mathcal{SQ}$. (*Hints*: **(1),** think of the elements of $\mathcal{SQ}$ as binary decimals; **(2),** Exercises **17** and **18** are relevant.)

Exercise 21. Show that $\mathbf{R} \sim P(\mathbf{N})$. (*Hint*: Start by finding a one-to-one correspondence between elements of $\mathcal{SQ}$ and subsets of $\mathbf{N}$.)

Exercise 22. Show that $\mathbf{R} \sim (\mathbf{R} \times \mathbf{R})$. (*Hint*: Start by showing that $\mathcal{SQ} \sim (\mathcal{SQ} \times \mathcal{SQ})$.)

Exercise 23. Let $\mathbf{R}^{\mathbf{R}}$ be the set of all functions $f: \mathbf{R} \to \mathbf{R}$.

[**a**]: Show that $P(\mathbf{R}) \preceq \mathbf{R}^{\mathbf{R}}$.

[**b**]**:** Show that $\mathbf{R}^{\mathbf{R}} \preceq P(\mathbf{R} \times \mathbf{R})$. (*Hint*: Identify each $f: \mathbf{R} \to \mathbf{R}$ with its graph $\{(t, f(t): t \in \mathbf{R}\}$.)

[**c**]**:** Show that $\mathbf{R}^{\mathbf{R}} \sim P(\mathbf{R})$. (*Hint*: You need the first two parts of this exercise, Exercise **22**, and the Schröder-Bernstein Theorem.)

Exercise 24 (another proof that $(\mathbf{N} \times \mathbf{N}) \approx \mathbf{N}$). This exercise develops a completely different bijection between $(\mathbf{N} \times \mathbf{N})$ and $\mathbf{N}$. Prove that the function

$$f : (\mathbf{N} \times \mathbf{N}) \longrightarrow \mathbf{N}$$

given by

$$z = f(n,k) := 2^n(2k+1) - 1$$

is one-to-one and onto. (*Hint*: Rewrite each $z + 1 \in \mathbf{N}^+$ by factoring out as many 2's as you can.)

[10] It is possible to come up with Russell's Paradox by analyzing exactly what is going wrong in Cantor's Paradox, and my guess is that that is how Russell came up with it. But I don't know this for a fact.

Appendix. An Accessible Proof of the Schröder-Bernstein Theorem.

The proof below does presuppose some familiarity with some of the terminology of graph theory, but the argument is very easy to follow.

Theorem (Schröder-Bernstein). *If there exist injections $f: X \to Y$ and $g: Y \to X$, then there exists a bijection $h: X \leftrightarrow Y$.*

Proof. Without loss of generality, we can picture any element $x \in X \cap Y$ as two distinct elements, one in X and the other in Y (so that X and Y become disjoint sets). Make the (bipartite) graph whose vertex set is $X \cup Y$ and whose edge set is

$$\{\{x, f(x)\}: x \in X\} \cup \{\{y, g(y)\}: y \in Y\}\,.$$

A minute's thought will convince you that each vertex in this graph is either of degree one (if it is not in the range of one of $\{f, g\}$) or of degree two (if it is in the range of one of $\{f, g\}$). It follows that each connected component of this graph must have one of three forms: (a), it can be a cycle, which necessarily has an even number of edges (f-edges alternating with g-edges); (b), it can be a one-sided infinite path, (f-edges again alternating with g-edges); or (c), it can be a two-sided infinite path (f-edges again alternating with g-edges.)[11]

Here, then, is how to make h: simply delete alternate edges in each component. (The only restriction on how this may be done is that in a type (b) components, one must start with the *second* edge). The edges that remain will clearly put the elements of X and Y into one-to-one correspondence. ∎

[11] Thus, this unimaginably huge graph has small (finite or denumerable) connected components. That's what makes this proof work.

List of All the Exercises

Exercise 1 (page 1). For each integer $n \geq 2$, let C_n be the closed interval $\left[0, \frac{1}{n}\right]$, and let O_n be the open interval $\left(0, \frac{1}{n}\right)$. Find $\bigcup_{n=2}^{\infty} C_n$; $\bigcup_{n=2}^{\infty} O_n$; $\bigcap_{n=2}^{\infty} C_n$; and $\bigcap_{n=2}^{\infty} O_n$.

Exercise 2 (page 2). Show that if a set B has n elements, then B has exactly 2^n subsets.

Exercise 3 (page 3). Show that R has all three properties [**a**], [**b**], and [**c**].

Exercise 4 (page 3). The properties [**a**], [**b**], and [**c**] are logically independent (in the sense that no two of them imply the third). Show this by demonstrating that each of the following three binary relations on **Z** is missing exactly one of the three properties and that no two of them are missing the same property.

$xR_1y \iff |x - y| \leq 2$

$xR_2y \iff x \leq y$

$xR_3y \iff x \leq 2 \;\&\; y \leq 2$

Exercise 5 (page 3). Find the flaw in the following "proof" that any binary relation that is symmetric and transitive must also be reflexive.

> For any $a \in S$, I claim that aRa. Proof: Take any b such that aRb. Since R is symmetric, also bRa. But then since R is transitive, aRa. ∎

Exercise 6 (page 4). The equivalence relation of exercise **3** generates two equivalence classes. What are they?

Exercise 7 (page 5). Let $g\colon A \to B$ and $f\colon B \to C$. Show that if f and g are both injective, then so is $f \circ g$.

Exercise 8 (page 5). Let $g\colon A \to B$ and $f\colon B \to C$. Show that if f and g are both surjective, then so is $f \circ g$.

Exercise 9 (page 5). Let $g\colon A \to B$ and $f\colon B \to C$. Show that if f and g are both bijective, then so is $f \circ g$.

Exercise 10 (page 6). Show that $\mathbf{N} \approx \mathbf{Z}$.

Exercise 11 (page 6). Show that $\approx$ is reflexive, symmetric and transitive. That is, for nonempty sets X, Y, and Z: show that $X \approx X$; show that if $X \approx Y$, then $Y \approx X$; and show that if $X \approx Y$ and $Y \approx Z$, then $X \approx Z$.

Exercise 12 (page 6). Let Y be a denumerable set, and let X be an infinite subset of Y. Show that X is denumerable. (*Hint*: Since Y is denumerable, we can list the elements $(y_0, y_1, y_2, \ldots)$. Go along the list and look for the elements of X.)

Exercise 13 (page 6). Let Y be any infinite set. Show that Y has a subset W such that $\mathbf{N} \approx W$. (*Hint*: Argue that it is possible to pull a sequence of distinct elements $(y_0, y_1, y_2, \ldots)$ out of Y; then let W be the set of elements in the sequence.)

Exercise 14 (page 6). Show that if X and Y are both denumerable, then $X \bigcup Y$ is also denumerable. (*Hint*: Obviously $X \bigcup Y$ is not finite. First: assume that $X \bigcap Y = \emptyset$ (that is, that there are no elements in common). Show that $(X \bigcup Y) \approx \mathbf{Z}$, and apply exercises **10** and **11**. Next: show that if $(X \bigcap Y) \neq \emptyset$, then $(X \bigcup Y)$ is equivalent to an infinite subset of **Z**, and apply exercises **10**, **11** and **12**.)

Exercise 15 (page 6). Let X be in infinite set. Show that X has a *proper* subset $\widehat{X}$ such that $X \approx \widehat{X}$. (*Hint*: The result you proved in Exercise **13** tells you where to start.)

Exercise 16 (page 8). Show that **Q** is denumerable.

Exercise 17 (page 9). Show that the open interval $(0, 1)$ is equivalent to **R**.

Exercise 18 (page 9). Let T be a denumerable subset of a nondenumerable set S. Show that $(S - T) \approx S$.

Exercise 19 (page 11). Show that it follows from The Schröder-Bernstein theorem that for nonempty sets X and Y, $\left\{\begin{matrix} X \prec Y \\ X \succ Y \end{matrix}\right\}$ cannot both hold.

Exercise 20 (page 11). Let $\mathcal{SQ}$ denote the set of all infinite sequences of 0's an 1's. Show that $\mathbf{R} \sim \mathcal{SQ}$. (*Hints*: **(1),** think of the elements of $\mathcal{SQ}$ as binary decimals; **(2),** Exercises **17** and **18** are relevant.)

Exercise 21 (page 11). Show that $\mathbf{R} \sim P(\mathbf{N})$. (*Hint*: Start by finding a one-to-one correspondence between elements of $\mathcal{SQ}$ and subsets of $\mathbf{N}$.)

Exercise 22 (page 11). Show that $\mathbf{R} \sim (\mathbf{R} \times \mathbf{R})$. (*Hint*: Start by showing that $\mathcal{SQ} \sim (\mathcal{SQ} \times \mathcal{SQ})$.)

Exercise 23 (page 11). Let $\mathbf{R}^{\mathbf{R}}$ be the set of all functions $f\colon \mathbf{R} \to \mathbf{R}$.

[**a**]: Show that $P(\mathbf{R}) \preceq \mathbf{R}^{\mathbf{R}}$.

[**b**]**:** Show that $\mathbf{R}^{\mathbf{R}} \preceq P(\mathbf{R} \times \mathbf{R})$. (*Hint*: Identify each $f\colon \mathbf{R} \to \mathbf{R}$ with its graph $\{(t, f(t)\colon t \in \mathbf{R}\}$.)

[**c**]**:** Show that $\mathbf{R}^{\mathbf{R}} \sim P(\mathbf{R})$. (*Hint*: You need the first two parts of this exercise, Exercise **22**, and the Schröder-Bernstein Theorem.)

Exercise 24 (page 11). (Another proof that $(\mathbf{N} \times \mathbf{N}) \approx \mathbf{N}$).) This exercise develops a completely different bijection between $(\mathbf{N} \times \mathbf{N})$ and $\mathbf{N}$. Prove that the function

$$f : (\mathbf{N} \times \mathbf{N}) \longrightarrow \mathbf{N}$$

given by

$$z = f(n, k) := 2^n(2k + 1) - 1$$

is one-to-one and onto. (*Hint*: Rewrite each $z + 1 \in \mathbf{N}^+$ by factoring out as many 2's as you can.)