Set theory basic notes

Randall Holmes

January 14, 2025

Contents

1 Introduction

This document discusses simple set theory foundations for undergraduate math. It is meant to be read. Not everything in it is test material; if something scares you, ask whether it might be on a test.

I do believe that some axioms and definitions are appropriate, and I give some. But I try to restrict them to what is actually useful for discrete math at an undergraduate level.

Something interesting happens (which is part of the reason I am writing this). The sum of what we teach in a discrete math course does add up to an axiomatic set theory, close to Zermelo set theory though not identical to it, and adequate for the foundations of classical mathematics.

I originally wrote this for Math 287 but adapting it to Math 305, spring 2025.

2 Sets

2.1 Primitive notions and equality of sets

I shall begin with the simplest assumptions about sets.

primitive properties and relations: Some objects in our world are *sets*.

Primitive notion 1 (sets and individuals): Some objects are *sets*. Objects which are not sets we call *individuals*[1].

Primitive notion 2 (membership): There is a relation of *membership*, written $\in$, which holds between general objects and sets they "belong to".

Primitive notion 3 (equality): We presume familiarity with the general notion of equality, and with its basic logical properties ($x = x$ [equality is reflexive], and "if $x = y$ and $P[x]$ then $P[y]$" [substitution of equals for equals]).)

We say "x is an element of a" or "x is a member of a", "x belongs to A" or "x is contained in a" to mean $x \in a$. "x is included in a" has another meaning for us. We write $x \notin A$ for "x is not an element of A".

We try to avoid saying "x is in A", because this is ambiguous: it can be confused with the subset relation.

identity criterion: With any data type, we want to be able to tell when two objects of that type are the same.

If A and B are sets, $A = B$ holds if and only if A and B have the same elements, that is, for any x, $x \in A$ holds exactly when $x \in B$ holds. Another way of putting this is, There is no element of A which does not belong to B, and there is no element of B which does not belong to A. This avoids vacuous quantification.

This is summarized in

[1] what I call "individuals" have also been called "atoms" or "urelements"

Axiom 1 (extensionality): If A and B are sets, and no element of B is not an element of A, and no element of A is not an element of B, then $A = B$.

Equivalently, if A and B are sets, $((\forall x \in A : x \in B) \wedge (\forall x \in B : x \in A) \rightarrow A = B$.

I could also say $(\forall x : x \in A \leftrightarrow x \in B) \rightarrow A = B$, for any sets A, B, but I prefer to avoid quantification over absolutely everything.

individuals and empty set: From the introduction of the membership relation, we extract this statement:

Axiom 1b (axiom of sethood): If $x \in A$, then A is a set. Equivalently, if A is an individual and x is any object, $x \not\in A$ (in which form we would call it the axiom of individuals).

We call this axiom 1b because it is a footnote to extensionality, and also because usual treatments of set theory do not allow for individuals at all (but this sometimes seems unnatural in undergraduate math classes).

A special case of the identity criterion is that if A and B are sets, and both have no elements at all, then they are equal. There is at most one set with no elements (we will introduce an axiom that says there is one in a moment). In addition, any objects in our world that are not sets have no elements. These objects can be called by various names: we call them individuals.

An important point in mathematics pedagogy is that official foundations of mathematics usually say that everything is a set, but common sense allows for individuals. Moreover, familiar mathematical objects (such as the natural numbers) have implementations as sets, but there is nothing inevitable about these implementations, and it is often natural to treat items not explicitly given as sets as individuals (natural and real numbers, for example).

2.2 Axioms which allow us to construct sets; notation for sets

properties define sets: If we have a set A given and a statement $P[x]$ about objects x in general, there is a set $\{x \in A : P[x]\}$ for which this is true: for any object a, $a \in \{x \in A : P[x]\}$ if and only if $a \in A$ and $P[a]$.

If I'm being paranoid about quantification over the entire universe, note that I can say

$$(\forall a \in \{x \in A : P[x]\} : x \in A) \wedge (\forall a \in A : a \in \{x \in A : P[x]\} \leftrightarrow P[a]).$$

The notation "$\{x \in A : P[x]\}$" is called set builder notation: it is important, and it has variations which are important.

This says: Given a set A and a property P, we can extract the collection of all elements of A which have the property P as a new set.

Notice that for any A, $\{x \in A : x \neq x\}$ is an empty set, and there is only one, for which we adopt the notation $\emptyset$.

We codify this as an axiom.

Axiom 2 (separation): For any sentence $P[x]$ about general objects x, and any set A, we have an object $\{x \in A : P[x]\}$, which is a set, and the axiom "for any a, $a \in \{x \in A : P[x]\}$ if and only if $a \in A$ and $P[a]$".

No universal set, so logical relations are not always implemented as set relations: We prove a theorem (we do not want to give credence to there being a "paradox" here, as people thought during a crisis of foundations at the beginning of the last century).

Let A be a set. Define $\texttt{Russell}(A)$ as $\{x \in A : x \not\in x\}$.

We prove that $\texttt{Russell}(A) \not\in A$.

We prove this by contradiction. Suppose $\texttt{Russell}(A) \in A$.

Now consider the status of the sentence $\texttt{Russell}(A) \in \texttt{Russell}(A)$.

This expands to $\texttt{Russell}(A) \in \{x \in A : x \not\in x\}$

which is equivalent to $\texttt{Russell}(A) \in A$ and $\texttt{Russell}(A) \not\in \texttt{Russell}(A)$

which is equivalent to $\texttt{Russell}(A) \not\in \texttt{Russell}(A)$ if (as we have assumed) $\texttt{Russell}(A) \in A$ is true.

But this is absurd. So we have shown that $\texttt{Russell}(A) \in A$ cannot be true (and so that $\texttt{Russell}(A) \not\in \texttt{Russell}(A)$), which has the more general consequence that there is no set V such that every object x is a member of V.

It follows that there can be no set relation implementing equality, membership or the subset relation, as the domain of any such relation would have to be V, the nonexistent universal set.

It also follows that for no set A can there be a set of all x not belonging to A (a true complement of A). If such a set existed, its union with A would be V. If a working universe U is understood in a particular context, $U \setminus A$ will play the role of the complement of $A \subseteq U$; but it doesn't contain everything that is not in A.

I extend remarks I made earlier. If I ever write $\{x : P[x]\}$, this does not mean the set of *all* x such that $P[x]$, but rather it means $\{x \in U : P[x]\}$ where U is some universal set which can be understood from context. For example, if we are talking about sets of integers, we might write $\{x \in \mathbb{Z} : P[x]\}$ as $\{x : P[x]\}$, particularly if we have to write many set builder notations and want to save typing. Many textbook authors (who certainly know better) defines $\overline{A}$ as the set of all things not in A. We define it as $\{x : x \not\in A\}$, that is, as $\{x \in U : x \not\in A\}$ where U is a "universe" understood from context. We discourage this notation, preferring $U - A$. So for example if E is the set of even integers, it is pretty clear that what I mean by $\overline{E}$ should be the set of odd integers, that is $\mathbb{Z} - E$, but there are plenty of things not in E which are also not in this set. In some weird context I might discuss the set of *real numbers* which are not even integers, and $\mathbb{R} - E$ is probably preferable to $\overline{E}$ for this unless it is *very* clear that we are in a discussion of sets of real numbers.

subset relation and power set: We define $A \subseteq B$, read "A is a subset of B" as "A is a set and B is a set and for any x, if x is an element of A, x is an element of B" **or** "A is a set, B is a set, and anything which is not an element of B is not an element of A". The contrapositive formula has value because the implicit quantifier is never vacuous: there is always something not in B, and it's clear from this definition that $\emptyset \subseteq B$: anything not in B is not in $\emptyset$.

Equivalently, we can define $A \subseteq B$ as $(\forall x \in A : x \in B)$.

Note that $\emptyset \subseteq A$ and $A \subseteq A$ always hold.

We assert as a basic assumption that

Axiom 3 (power set): for any set A, there is a set $\mathcal{P}(A)$, called the power set of A, whose elements are exactly the subsets of A.

Notice that the axiom of extensionality actually says that if $A \subseteq B$ and $B \subseteq A$, then $A = B$. That is exactly what it says.

list notation for finite sets: The notation $\{x\}$ denotes the set whose only element is x.

The notation $\{x, y\}$ denotes the set whose only elements are x and y.

The notation $\{x, y, z\}$ denotes the set whose only elements are x, y, z.

And so forth. List notation $\{x_1, x_2, \ldots, x_n\}$ defines a set whose only elements are the x_i's. The general definition of this (familiar) list notation is technically exacting to state (we might do it eventually).

Notice that order and repetitions of items do not make any difference in the reference of this notation. $\{x, x\}$ is the same set as $\{x\}$. Notice that this means that you cannot tell that a set written $\{x, y\}$ has two elements unless you are given the information that x and y are distinct.

$\{x, y\}, \{y, x\}, \{x, y, y\}$ etc. are all the same set.

assumptions behind list notation: Behind the ability to write this notation, there are two basic assumptions.

Axiom 4 (singletons): For any object a, there is a set $\{a\}$ such that for any x, $x \in \{a\}$ exactly if $x = a$.

It is fun to notice that if a is a set, $\{a\} = \{x \in \mathcal{P}(a) : x = a\}$ is already given by axioms previously stated. But we are not presuming that every object is a set, so we need the axiom of singletons. [2]

[2] An inverse operation of sorts to the singleton operation is definite description. It

Axiom 5 (binary union): For any sets A and B, there is a set $A \cup B$ such that for any x, $x \in A \cup B$ if and only if either $x \in A$, or $x \in B$, or both. This set is called the union of A and B.

Now $\{x, y\} = \{x\} \cup \{y\}$ and more generally $\{x_1, x_2, \ldots, x_n\} = \{x_1\} \cup \{x_2\} \cup \ldots \cup \{x_n\}$.

Relations of part and whole on sets?: There is a temptation to say that a set is a whole made up of its elements. A classic textbook written by a man who certainly knew better gave packs of wolves and bunches of grapes as examples of sets.

This temptation should be firmly resisted. If A is part of B and B is part of C, then A is part of C, for any objects A, B, C on any reasonable understanding of the relation of part to whole. But if a, b are any two distinct objects, $a \in \{a, b\}, b \in \{a, b\}$, and $\{a, b\} \in \{\{a, b\}\}$. If membership were transitive as the relation of part to whole is, it would follow that $a \in \{\{a, b\}\}$ and $b \in \{\{a, b\}\}$. But $\{\{a, b\}\}$ has only one element, the set $\{a, b\}$: a and b are not both in it (even if one of them were weirdly the same as $\{a, b\}$)

Related temptations are the common desire to say that $\emptyset$ belongs to every set, or to write $\emptyset$ as $\{\emptyset\}$ (the latter is a set with one element while $\emptyset$ has no elements).

It is important to notice that none of this has to do with famous worries about infinite sets: these issues arise from the simplest construction of sets by finite listing. They do have to do essentially with allowing sets to be elements of sets, which is an important move for the uses of sets intended in mathematics.

The natural relation of part to whole on sets is the subset relation. We say "A is included in B" for $A \subseteq B$, not $A \in B$. Note that if $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.

It may seem peculiar that every set has the empty set as a part: a part of a set could be defined as a *nonempty* subset of the set, which would

would be handy to have an operator θ such that $\theta(\{x\}) = x$ and $\theta(u) = \emptyset$ if x is not a singleton set. Then $\theta(\{x \in A : P[x]\}$ would represent the unique x in A such that $P[x]$ if there is one. We note the desirability of this without postulating it. We would read $\theta(A)$ as "the unique element of A", and use of this notation would usually presume that A has exactly one element.

preserve the idea that disjoint sets have no common part (though they do have the common subset $\emptyset$)

Notice that $x \in A$ is equivalent to $\{x\} \subseteq A$: the elements of A correspond to but are not identical with atomic parts (smallest possible nonempty subsets) of A.

Notice that the examples given in the famous textbook are objects with disjoint parts of an understood kind: a pack of wolves does have a natural association with a set of wolves, and a bunch of grapes with a set of grapes. But the mass of all human cells is roughly speaking the same as the mass of all human beings, while the set of human cells is a lot larger then the set of human beings.

Other interesting binary operations on sets: We define $A \cap B$ as

$$\{x \in A : x \in B\}.$$

This is called the intersection of A and B.

Exercise: prove that $\{x \in A : x \in B\} = \{x \in B : x \in A\}$

We define $A - B$ (also sometimes written $A \setminus B$) as $\{x \in A : x \notin B\}$. Here, $x \notin B$ simply means "x is not an element of B". This is called the set difference of A and B or the complement of B relative to A.

These, along with union, are the basic operations for the parlor game of Venn diagrams, which we will play (which does have its uses, mostly to illustrate very simple properties of two or three sets).

Notice that the two operations introduced here require no new axioms, because the set defined is included in a set already given. Also notice that this was not true of unions of two sets, which is why we needed an axiom for that.

3 Pairs, lists, relations, functions

The abstraction of ordered lists: In addition to considering sets, which are not ordered ($\{x,y\}$ is the same set as $\{y,x\}$) we want to consider ordered pairs (x,y) or ordered lists $[x_1,x_2,\ldots,x_n]$, which are the same exactly if they have the same number of items appearing in the same order (repetitions being significant). Our reasons for using different delimiters for lists will be revealed later.

This is an independent, if related idea. It is interesting that it can be implemented entirely in terms of the set concept. But please notice that there is nothing inevitable or unique about this implementation.

Concrete examples of the general use of the ordered list concept are not hard to come by (permutations versus combinations).

Basic properties of the ordered pair: The basic properties of the ordered pair concept are...

for any objects x,y there is a pair (x,y)

$(x,y)=(z,w)$ if and only if $x=z$ and $y=w$.

A definition of (x,y) as a set which works, and the one which is now almost exclusively used, is $(x,y)=\{\{x\},\{x,y\}\}$. For this to work as an ordered pair definition, we need to be able to construct it for any x and y (our axioms of singletons and union let us do that) and we need to be able to extract x and y from (x,y) (that is, given an ordered pair we need to be able to identify its first and second components). I tell you how to do this, though this proof and the exact definition of the ordered pair are not examinable content: x is the only object which belongs to all elements of $(x,y)=\{\{x\},\{x,y\}\}$, and y is the only object which belongs to exactly one element of $(x,y)=\{\{x\},\{x,y\}\}$. A reason that proofs about this ordered pair notion can be tricky (and a good reminder of the perils of set notation) is that one cannot assume that $\{\{x\},\{x,y\}\}$ has two elements: if $x=y$, $\{\{x\},\{x,y\}\}=\{\{x\}\}$.

What we require of the general list concept is

for any $x_1,\ldots,x_n$ there is a list $[x_1,\ldots,x_n]$

$[x_1,\ldots,x_n]=[y_1,\ldots,y_n]$ iff $x_i=y_i$ where $1\leq i\leq n$.

We actually define lists as functions:

$$[x_1, \ldots, x_n] = \{(i, x_i) : 1 \leq i \leq n\} :$$

our official discussion of functions is below.

A definition of the ordered pair as a set (historical, easier than the usual one (?)): The first definition of the ordered pair (x, y) given by Norbert Wiener in 1914 (our official one was given by Kuratowski in 1920) was $(x, y) = \{\{\{x\}, \emptyset\}, \{\{y\}\}\}$. If you like solving logic puzzles, you might have fun figuring out why it is easier to extract x and y from this "pair". Hint: this set definitely has two elements, and it has one element with one element and one element with two, whether $x = y$ or not. You are in no way responsible for this. But it is worth noticing that definitions of this kind of concept can take different forms: all we need of the ordered pair is that (x, y) exists for any x and y, and that given (x, y), we can identify its first projection x and its second projection y.

The Cartesian product: Given sets A, B, we define the Cartesian product of A and B, which we write $A \times B$, as the collection of all ordered pairs (a, b) with $a \in A$ and $b \in B$.

The existence of the Cartesian product is a consequence of the axioms we have already given.

We define $\mathcal{P}^2(A)$ as $\mathcal{P}(\mathcal{P}(A))$, and more generally $\mathcal{P}^{n+1}(A)$ as $\mathcal{P}(\mathcal{P}^n(A))$ for $n \geq 2$.

It is then straightforward to observe that for any $a \in A$, $b \in B$, both $\{a\}$ and $\{a, b\}$ belong to $\mathcal{P}(A \cup B)$ so $\{\{a\}, \{a, b\}\}$ belongs to $\mathcal{P}^2((A \cup B)$, and so we can define $A \times B$ as the set of all x in $\mathcal{P}^2((A \cup B)$ such that for some $a \in A, b \in B$, $x = \{\{a\}, \{a, b\}\}$:

$$A \times B = \{x \in \mathcal{P}^2(A \cup B) : (\exists a \in A : \exists b \in B : x = (a, b)\}$$

Here we have written the set builder notation very carefully so that you can see that axiom 2 provides us with this set. This might more often be written

$$A \times B = \{(a, b) : a \in A \wedge b \in B\}$$

The expansion above can be taken as a general hint of how to deal with complicated expressions left of the colon in set builder notation (where axiom 2 formally allows only a variable letter and the bounding set) and the bounding set on the left of the colon can be deduced from information given on the right of the colon (clearly $(a, b) \in A \times B$ – once we know that Cartesian products exist in general).

The proofs that Cartesian products exist should not be important directly to you; the mere assertion that they exist should be enough, as you really shouldn't work directly with the details of pairs as sets very much (what you do with them should actually be quite independent of their implementation). What should have some interest is that an implementation is possible!

The definition of a relation: Let A, B be sets. A relation from A to B is a triple $R = (A, B, G)$ where G is a subset of $A \times B$, and where we define (x, y, z) (for this specific purpose) as $((x, y), z)$.

We call A the domain of R ($\texttt{dom}(R)$), B the codomain of R ($\texttt{cod}(R)$) and G the graph of R ($\texttt{graph}(R)$). There is another school of thought (which by temperament I prefer) which identifies a relation with its graph, but there are technical problems with this, because in general the domain and codomain cannot be determined from the graph, and it is common to speak of the domain and codomain as features of the relation.

We define $x\,R\,y$ as the assertion $(x, y) \in \texttt{graph}(R)$.

We define the image or range of R as the set of y in the codomain of R such that there is x such that $x\,R\,y$.

We define the preimage of R as the set of x in the domain of R such that there is y such that $x\,R\,y$.

Many but not all transitive verbs in mathematics can be read as relations. The most general ones, such as $x = y$, $x \in y$, $x \subseteq y$ cannot, because there are no sets large enough to serve as domain or codomain of these "logical relations".

the definition of functions: A function is a relation $F = (A, B, G)$ with the property that for each $x \in A$, there is exactly one $y \in B$ such that $x\,F\,y$.

We write $f : A \to B$ to mean "f is a function with domain A and codomain B" or, less formally, f is a function from A to B.

This is more concrete than the common notion in textbooks that a function is a rule. As we will see, the set of ordered pairs G codes the rule.

We expand the language a little: it is equivalent and perhaps easier to follow to say that for each $x \in A$, there is $y \in B$ such that $x\,F\,y$ (so the preimage of F is the domain) and for any x, y, z, if $x\,F\,y$ and $x\,F\,z$, then $y = z$.

If F is a function and $x \in \mathtt{dom}(F)$, we define $F(x)$ as the unique y such that $x\,F\,y$. [3]

It is worth noting that there are "logical functions" which are not implementable as sets. For example, there can be no function F such that $F(x) = \mathcal{P}(x)$ for all x, since the domain of such a function would be the collection of all sets, which can be shown not to exist by the Russell argument.

You learned a quite different definition of the function concept in high school and in college calculus. We illustrate that our formalization is adequate to support that informal definition.

We set out to define the function $y = 2x + 5$ from real numbers to real numbers. We suppose that we have the set $\mathbb{R}$ of real numbers handy.

We then have the graph of f as the set G of all u in $\mathbb{R} \times \mathbb{R}$ such that there is $x \in \mathbb{R}$ such that $u = (x, 2x + 5)$.

In general, if we are given a definition $y = f(x)$ of a function by a rule, as is usually done in calculus, and we understand a domain A and codomain B of f from context, then

$$f = (A, B, \{(x, y) \in A \times B : y = f(x)\}).$$

A notation for this might be $(x \in A \mapsto f(x) \in B)$. The sets can be omitted if they are understood from context (it is very odd to explicitly provide B as I do here, but I am making a point).

[3] If we had the definite description operator, we could write

$$F(x) = \theta(\{y \in \mathtt{cod}(F) : x\,F\,y\}).$$

So the function above could be written as $(x \in \mathbb{R} \mapsto 2x+5 \in \mathbb{R})$, which could just be written $(x \in \mathbb{R} \mapsto 2x+5)$ [certainly] or as $(x \mapsto 2x+5)$ if you are confident that the domain is understood [$(x \in \mathbb{Z} : 2x+5)$ is not the same function!]

A weird point which I should mention but not make too much of is that $(x \in \mathbb{R} : x^2 \in \mathbb{R})$ and $(x \in \mathbb{R} : x^2 \in \mathbb{R}^+ \cup \{0\})$ are functions with the same values at the same inputs, but they are not the same function because they have different codomains. They have the same graph: on the alternative view identifying functions with their graphs, they would be the same function.

Images and inverse images: If $f : A \to B$ and $C \subseteq A$, we define $f[A]$ as $\{f(x) : x \in C\}$. which can also be written in the form $\{b \in B : (\exists c \in C : f(c) = b)\}$. We provide the second form to make it clear that this set exists by axiom 2, and also to provide practice in reading more complicated forms of set builder notation.

If $f : A \to B$ and $C \subseteq B$, we define $f^{-1}[A]$ as $\{a \in A : f(a) \in C\}$. Notice that this involves no mention of the inverse function of f or even any assumption that f has an inverse function (which is fortunate, since we don't mention inverse functions until the next section!)

Many authors use parentheses in these notations. This is traditional (you will see this usage in books) but a bad idea: if there is an object C in the domain of f which is also a subset of the domain of f, which is not at all impossible, then $f(C)$ could mean two different things, and similarly if f has an inverse function and C is an element of the codomain which is also a subset of the codomain, $f^{-1}(C)$ would have two possible interpretations.

Levin also allows $f^{-1}(x)$ to be used to mean $f^{-1}(\{x\})$, and I do not. If f^{-1} exists as an inverse function, this is really bad notation.

Some kinds of function which are commonly considered: A function f is an injection, or one-to-one, if for any $x, y \in \texttt{dom}(f)$, if $f(x) = f(y)$ then $x = y$.

A function f is a surjection, or onto, if for any $y \in \texttt{cod}(f)$, there is x such that $y = f(x)$. Notice that in our weird example at the end of the last subsection, the first function is not surjective, and the second is.

A function f is a bijection iff it is one-to-one and onto (i.e., an injection and a surjection).

Sizes of sets: We say that two sets A, B are of the same cardinality, or the same size, or have the same number of elements, which we write $A \sim B$, iff there is a bijection from A to B. We associate with each set A an object called its cardinality, written $|A|$, in such a way that $|A| = |B|$ if and only if $A \sim B$. We will usually talk about this only in the case where A is a finite set and $|A|$ is a non-negative integer. It would take us a bit far afield to actually show a definition of $|A|$ which works for all sets. If we discuss only subsets of a fixed set X, we can define $|A|_X$ as $\{B \in \mathcal{P}(X) : B \sim A\}$.

Inverse relations and functions: For any relation $R = (A, B, G)$, there is an inverse relation $R^{-1} = (B, A, \{(y, x) : (x, y) \in G\})$.

A function f is a relation, so f^{-1} defined as above certainly exists. We say that f has an inverse function exactly if f^{-1} is itself a function. This is true exactly if f is both injective and surjective, that is, iff f is a bijection. So the condition on functions of being a bijection is exactly the same as the condition of having an inverse function.

Our official definition of ordered lists: We decouple lists from ordered pairs. We use the notation $[x_1, x_2, \ldots, x_n]$ to make this clear.

We define an ordered list as a graph of a function whose domain is an interval in the integers (assuming familiarity with the integers; our treatment below will at least suggest how the integers themselves could fit into our scheme). This allows variations in how they are indexed. If x is an ordered list, x_i is then simply defined as $x(i)$. The list $[x, y, z]$ is for us the set $\{(x, 1), (y, 2), (z, 3)\}$ (or it might be $\{(x, 0), (y, 1), (z, 2)\}$ if the context tells you our indexing starts at 0).

Note that this definition supports lists of length 0 and 1 (and lists $[x, y]$ of length 2 are not the same as ordered pairs (x, y)).

Traditional names for sets you already know about: The name $\mathbb{Z}$ is traditional for the set of integers (positive and negative whole numbers and zero), $\mathbb{Q}$ for rational numbers, and $\mathbb{R}$ for real numbers. The name $\mathbb{Z}^+$ is traditional for the set of positive integers, $\mathbb{Q}^+$ for positive rational numbers, and $\mathbb{R}^+$ for positive real numbers. The name $\mathbb{N}$ for the set of natural numbers suffers from an ambiguity as whether 0 is a natural number or not. We usually view $\mathbb{N}$ as referring to $\{n \in \mathbb{Z} : n \geq 0\}$, which includes 0, the set of nonnegative integers, which seems best because we already have a generally accepted name for the positive integers: but the Art of Proof takes the other tack. Notice that the set $\mathbb{N}$ as defined here is the set of sizes of finite sets.

Proof strategies for set notions: If $P[x]$ is a sentence of our language including the variable x, and T is a possibly complicated expression (it doesn't have to be a variable) we write $P[T/x]$ for the result of substituting T for x. We will usually just write $P[T]$ for this but the more explicit form can sometimes be useful. So if $P[x]$ is the sentence $x > 3$, $P[a + b/x]$ or just $P[a + b]$ is $a + b > 3$.

To prove $T \in \{x \in A : P[x]\}$ we have the rule of *set introduction*

$$\frac{\begin{array}{c} T \in A \\ P[T/x] \end{array}}{T \in \{x \in A : P[x]\}}$$

and to use a statement $T \in \{x \in A : P[x]\}$ which we have assumed or proved, we have the rule of *set domain*

$$\frac{T \in \{x \in A : P[x]\}}{T \in A}$$

and the rule of *set elimination*

$$\frac{T \in \{x \in A : P[x]\}}{P[T/x]}$$

These rules simply formally expand what we mean by set builder notation.

Then we have rules for the subset relation.

This we will call simply *inclusion*, the rule for using a subset statement:

$$\frac{\begin{array}{c} T \in A \\ A \subseteq B \end{array}}{T \in B}$$

The next rule we will give the more grandiose name *subset introduction*. It involves an assumption and a block of statements.

Goal: $A \subseteq B$, where A and B are sets (if we were being really really formal we would want lines saying that A and B are sets above in the proof, but we won't clutter our summary with this):

Assume (line m): $x \in A$ [x must be a variable, and must be a new variable which does not appear elsewhere in the proof, except maybe in blocks already closed]

$\vdots$

Proof lines

(line n) $x \in B$

(line $n+1$): $A \subseteq B$ subset introduction m-n [and as usual the indented block is closed and one cannot use the lines in it again, since they involve the arbitrary object x and assumption $x \in A$ which were introduced only to prove this subset statement]

Analogous to biconditional introduction in a way is a strategy for proving $A = B$ where A, B are sets:

We could define quantifiers using set builder notation: We could define the sentence $(\forall x \in A : P(x))$ as $A \subseteq \{x \in A : P(x)\}$. This is read, for all $x \in A$, $P(x)$.

We could define the sentence $(\exists x \in A : P(x))$ as $\{x \in A : P(x)\} \neq \emptyset$. This is read, for some $x \in A$, $P(x)$, or there exists $x \in A$ such that $P(x)$.

[4]

Defining quantifiers in terms of sets might be taken as an odd maneuver. It is equally odd in discrete math texts (I think odder) that quantifiers are often introduced before sets are introduced...but with set bounds as here, so they depend on informal understanding of sets anyway.

[4] Further, $(\forall x \in A : P(x) \rightarrow Q(x))$ is definable as $\{x \in A : P(x)\} \subseteq \{x \in A : Q(x)\}$ ($P \rightarrow Q$ can be defined as $(\forall x \in A : P \rightarrow Q)$, x not occurring in P, Q and A nonempty). This is not as absurd as it looks: quantified implication was defined first in the actual history! Similar maneuvers can define the other propositional connectives: we content ourselves with observing that $(\forall x : \neg P(x))$ is definable as $\{x \in A : P(x)\} = \emptyset$ and that all the propositional connectives can be defined in terms of negation and implication.

We can then define $\{(x,y) \in A \times B : P(x,y)\}$ as

$$\{u \in A \times B : (\exists x \in A : (\exists y \in B : u = (x,y) \wedge P(x,y)))\}.$$

This can be used as a general model for how to treat complicated expressions appearing left of the colon in set builder notation.

And then we can say in general that a function definition $y = F[x]$ where $F[x]$ stands in for some complicated expression in x is equivalent to

$$f = (A, B, \{(x, F[x]) \in A \times B : x \in A\}),$$

where A is the intended domain (often implicit in a definition of this kind), B is the intended codomain, and the definition only succeeds if for every $x \in A$ it is the case that $F[x] \in B$ (though this can be qualified: in calculus you often work with partial functions, which may be undefined at some elements of the implicitly understood domain: the calculus definition of domain is more analogous to what we call preimage above).

The notation $(x \in A \mapsto F[x] \in B)$ is convenient for this. The mention of the domain A is often omitted, and mentioning the codomain as I do here would be strange in practice.

Expansion of set builder notation: Where $f : A \rightarrow B$, define

$$\{f(x) : x \in A \wedge P[x]\}$$

as $\{y \in B : (\exists x \in A : y = f(x) \wedge P[x])\}$. The $P[x]$ component is optional. If $f : X \rightarrow C$ is a function whose domain X includes $A \times B$, $\{f(x,y) : x \in A \wedge y \in B\}$ can be read

$$\{z \in C : (\exists xy : z = f(x,y) \wedge x \in A \wedge y \in B\}.$$

This isn't exhaustive, but might give an idea of how to proceed in reading complex set builder notation. The expressions $f(x)$ and $f(x,y)$ can be expanded out: one does have to be able to identify a bounding set which will contain all values of the expression to the left of the colon with inputs from given sets.

Infinite unions: If I is an index set and $A : I \rightarrow \mathcal{P}(B)$ is a function, where by convention we can write $A(i)$ as A_i, we define $\bigcup_{i \in I} A_i$ as

$\{x \in B : (\exists i \in I : x \in A_i)\}$: we can take unions of infinite collections of subsets of a fixed set. Similarly, we define $\bigcap_{i \in I} A_i$ as $\{x \in B : (\forall i \in I : x \in A_i)\}$.

The book discusses notation $\bigcap_{i=1}^{k} A_i$ equivalent to $\bigcap_{i \in \{1,\ldots,k\}} A_i$, with recursive definition which I gave on the board and may ask you to write out as an exercise (and similarly for unions).

4 Infinite sets, natural numbers, and recursion

We describe the usual implementation of the non-negative integers as sets. 0 is defined as $\emptyset$. 1 is defined as $\{0\}$. 2 is defined as $\{0, 1\}$. And so on. In this way each natural number n (in the inclusive sense which allows 0) is defined as a set with n elements in the usual sense, namely $n = \{0, \ldots, n-1\}$.

First of all, we are not claiming that we have revealed that 3 actually is $\{0, 1, 2\}$. This is an implementation, not a revelation of the true nature of the non-negative integers.

Secondly, it would be good to be rather suspicious as to whether a definition has actually been given! We proceed to show that this definition can indeed be given formally, but it requires another axiom.

Definition: For any set x, $x^+ = x \cup \{x\}$.

Definition: We say that a set I is *inductive* iff $\emptyset \in I$ and $(\forall x \in I : x^+ \in I)$.

Axiom of Infinity: There is an inductive set $\mathcal{I}$.

Definition: We define $\mathbb{N}_0$ as the set of all $n \in \mathcal{I}$ such that n belongs to every inductive set. (we define $\mathbb{N}$ as $\mathbb{N} \setminus \{\emptyset\}$).

Theorem: $\mathbb{N}_0$ is inductive.

Every inductive set (including $\mathcal{I}$) contains $\emptyset$ as a member (definition of inductive) so $\emptyset \in \mathbb{N}_0$. Now suppose that $x \in \mathbb{N}_0$: so x belongs to every inductive set. Let $\mathcal{J}$ be an arbitrary inductive set: x belongs to $\mathcal{J}$, so x^+ belongs to $\mathcal{J}$, because $\mathcal{J}$ is inductive. Thus x^+ belongs to every inductive set (including $\mathcal{I}$), and so $x^+ \in \mathbb{N}_0$.

Definition: We define $\sigma(x)$ [the successor of x, which we may also write $x + 1$] as x^+, for each $x \in \mathbb{N}_0$.

Definition: We define 0 as $\emptyset$, 1 as 0^+, 2 as 1^+, 3 as 2^+, 4 as 3^+, 5 as 4^+, 6 as 5^+, 7 as 6^+, 8 as 7^+, 9 as 8^+, and 10 as 9^+.

Theorem (mathematical induction): For any sentence $P[x]$, if we can prove $P[0]$ and we can prove $(\forall k \in \mathbb{N}_0 : P[k] \to P[k+1])$, we can prove $(\forall n \in \mathbb{N}_0 : P[n])$.

The proof is straightforward and almost too slick to follow. We claim that $S = \{x \in \mathbb{N}_0 : P[x]\}$ is inductive. $0 \in S$ because $0 \in \mathbb{N}_0$ and $P[0]$ by the basis step. If $x \in S$ then $x \in \mathbb{N}_0$ and $P[x]$. We then have $x + 1 \in \mathbb{N}_0$ because $\mathbb{N}_0$ is inductive and we have $P[x + 1]$ by the induction step, so $x + 1 \in S$. Thus S is inductive, so for every $n \in \mathbb{N}_0$, $n \in S$, so for every $n \in \mathbb{N}_0$, we have $P[n]$.

The collection of natural numbers was defined exactly so as to make mathematical induction work!

Definition: We define a *sequence* as a function with domain $\mathbb{N}$, and we will write $s(i)$ as s_i when s is a sequence.

Iteration theorem: For any set A, function $f : A \to A$, and $a \in A$, we define an f, a-inductive set as a set which contains $(0, a)$ and for any $n \in \mathbb{N}$ contains $(n + 1, f(x))$ if it contains (n, x). The collection $G_{f,a}$ of all pairs $(x, y) \in \mathbb{N} \times A$ which belong to all f, a-inductive sets is the graph of a sequence $b^{f,a} = ((\mathbb{N}, A), G_{f,a})$: we define $f^n(a)$ as $b^{f,a}_n$.

Notice that we are formally articulating the recursive definition $b_0 = a; b_{n+1} = f(b_n)$: the Iteration Theorem is the beginning of a formal justification of recursive definitions in general.

Note also that the Theorem requires proof. I may insert it here, but it is quite tricky and not examinable. But it is useful to see that we can in fact demonstrate the validity of recursive definitions in the set theory framework.

It is immediate that $b^{f,a}$ is a relation from $\mathbb{N}$ to A. What takes a wee bit more work is showing that it is a function, and so a sequence.

First, we show that there is exactly one y (namely a) such that $(0, y) \in G_{f,a}$. To see this, verify that $G_{f,a} \setminus \{(0, y) : y \in A \wedge y \neq a\}$ is f, a-inductive: it contains $(0, a)$ and if it contains (n, x), it contains $(n +$

$1, f(x))$ because this is in $G_{f,a}$ and $n+1$ is not equal to 0, because $n+1$ is nonempty and 0 is empty as a set. So this set includes $G_{f,a}$ and since it is obviously included in it, it is equal to $G_{f,a}$, and the only pair $(0, y)$ that it contains is $(0, a)$.

We continue the proof by mathematical induction. Suppose that for a fixed $k \in \mathbb{N}_0$, there is only one y such that $(k, y) \in G_{f,a}$. Consider the set $G_{f,a} \setminus \{(k+1, z) : z \in A \wedge z \neq f(y)\}$. This set is again fairly easy to show f, a-inductive, but we need a Lemma: $(0, a)$ is obviously in this set because $0 \neq k+1$. Suppose (n, x) is in this set. If $n = k$ then $x = y$ by ind hyp, so $(n+1, f(x)) = (n+1, f(y))$ is in this set. if $n \neq k$ we claim that $(n+1, f(x))$ is in the set because it is in $G_{f,a}$ and $n+1 \neq k+1$. This requires a Lemma which we now prove:

Lemma: If $n, k \in \mathbb{N}_0$ and $n+1 = k+1$, then $n = k$.

Proof: If $n+1 = k+1$, but $n \neq k$ then we have $n \cup \{n\} = k \cup \{k\}$ with $n \neq k$, from which it follows that $n \in k$ and $k \in n$.

We *could* quite reasonably handle this by adding a last

Axiom of Foundation (lite): For all x, y, either $x \notin y$ or $y \notin x$: no set is an element of any element of itself.

But we don't strictly speaking need to. It can be proved, it's just messy.

So it is enough to prove by mathematical induction that for all $n \in \mathbb{N}_0$, for all $k \in n$, $n \notin k$. Basis step: for all $k \in 0$, $0 \neq k$, because 0 has no elements at all! Induction step: suppose for a fixed $m \in \mathbb{N}_0$ that there is no $n \in m$ such that $m \in n$. How do we show that there is no $n \in m+1$ such that $m+1 \in n$? There doesn't seem to be an immediate way to do this.

But we can, by strengthening the statement of the theorem to force the induction to work. We instead prove that for all $n \in \mathbb{N}_0$, for all $k \in n$, for all $p \in \mathbb{N}_0$ with $n \subseteq p$, $p \notin k$. The induction step is the same: this is true for 0 because 0 has no elements. Now suppose for a fixed m that for all $k \in m$, for all natural numbers p with $m \subseteq p$, $p \notin k$. It follows that for all $k \in m$, for all $q \supseteq m+1$, $q \notin k$, because $m \subseteq m+1 = m \cup \{m\} \subseteq q$. Further $q \notin m$, because m is an element of $m+1$ and so of q and by inductive hypothesis cannot be an element of an element of m. So we have shown that $q \supseteq m+1$ cannot belong to any element of $m+1$, since any such element is either an element of m or equal to m.

This completes the rather mind-boggling proof of the Lemma, and the proof of the Iteration Theorem.

Definitions of addition and multiplication: We can then define $m+n$ as $\sigma^n(m)$ and $m \cdot n$ as $(\sigma^m)^n(0)$. It takes work to prove the known properties of addition and multiplication of non-negative natural numbers in this very restricted framework, but it can be done (mathematical induction is the key). We can define $x \leq y$ as $(\exists k \in \mathbb{N}_0 : x+k=y)$. In this context, it is easier to define $x < y$ in terms of $x \leq y$ than vice versa.

More natural definitions of finite cardinality, addition, and multiplication: We can define $|A|$ as the unique $n \in \mathbb{N}_0$ such that $A \sim n$, then define $m+n$ as $|(m \times 0) \cup (n \times 1)|$ and $m \cdot n$ as $|m \times n|$, for $m, n \in \mathbb{N}_0$.

One would have to prove some theorems at this point, which we will not do, though they are not terribly hard (if one can breathe the air at this level of abstraction). It's useful to prove $m \sim n \rightarrow m = n$ for all $m, n \in \mathbb{N}_0$, to make sure that the definition of finite cardinality makes sense. It is also useful to prove that $m+n$ and $m \times n$ actually belong to $\mathbb{N}_0$. It is rather easier to prove familiar algebraic properties of the arithmetic operations using these definitions.

Definition of decimal notation: We can define $\Delta(n,d)$ as $d + 10 \cdot n$ for any $d \in \{0,1,2,3,4,5,6,7,8,9\}$ and $n \in \mathbb{N}_0$. Of course it is usual to write $\Delta(\Delta(1,2),3)$ as 123.

Summary of our intentions: This section is really a teaser: we want to give some indication of how it is that foundations of the usual mathematical subjects can be given entirely in terms of set theory. An immediate puzzle is, how do we then define the integers? I leave you to think about that, for now.