SET THEORY HANDOUT FOR MATH 113

ALLEN KNUTSON

1. The axioms of set theory

Mathematics has not always been as it is regarded today – a rock-solid collection of work, built on an axiomatic foundation. No, only in the beginning of the last century did people really worry about the underpinnings. There have been a couple of different approaches to provide first-principles mathematics, but the only one most people learn is **set theory**.

Euclid's plane geometry begins with the notion of "point" and "distance" without really defining them.[1] We will begin with **set** and a binary relation $\in$, or **"is an element of"**, with which to compare two sets a, b; either $a \in b$ or $a \notin b$. So there is nothing else we can talk about – any object we ever discuss is some kind of set, or a property of sets, and the only *fundamental* properties relating two sets are that they might be equal, or that one might be $\in$ the other.

Axiom (of equality)**.** *Given two sets A, B, if $\forall t, t \in A \iff t \in B$, then $A = B$.*

This is a mathematical formulation of the idea you may have that sets are determined by their elements – there's no order, no repetition, either t is in there or it's not. In particular, if you want to prove two sets are equal, you will typically use the axiom of equality.

We define A to be a **subset** of B, and write $A \subseteq B$, if we only have one direction above: $\forall t, t \in A \implies t \in B$. (This is a property relating two sets that is not fundamental – it's built out of the fundamental one, $\in$.)

Many of the other standard axioms say that the class of sets defined by a certain property P form a set – in other words, there exists a set A such that $\forall t, P(t)$ is true $\iff t \in A$.

Define a **union** $\bigcup A$ of A by $\forall t, t \in \bigcup A \iff \exists S \in U$ such that $t \in S$. In words, we're thinking of A as a set of sets (it must be – everything's a set!), and taking the union of those.[2]

Why do we say "a" union? You're surely more used to hearing "the" union. Note that the axiom of equality says that if we had two unions, they'd be equal (since they have the same elements).

Axiom (of union)**.** *For any set A, there exists a union $\cup A$ (necessarily unique).*

This is a satisfying axiom – something we have a handle on, that we know would be unique, actually exists.

Date: September 11, 2000.

[1] "A point is that which has no parts." Poetic, but basically Euclid defines other things in terms of points, and leaves points as the basic, undefinable beginning concept. Which is fine.

[2] This may not sound like the union of two sets you're more used to, $A \cup B$. That's a special case of this, namely $\bigcup\{A, B\}$. Given the union-of-two-sets construction, we can define the union of finitely many sets inductively; the definition we're using here is better in that it lets us take the union of any number of sets.

1.1. **Why bother?** Before presenting more axioms like the Axiom of Union, that say various properties are "set-forming," let's check why we're bothering. Why can't we just say "the set of all such-and-such satisfying a certain property," and continue?

Here's a famous pitfall, Russell's paradox. For a set A, let $SS(A)$ be true if $A \in A$, false if $A \notin A$, where "SS" stands for "self-swallowing". Now assume that there exists a set R (for Bertrand Russell) such that

$$\forall t,\ t \in R \iff \sim SS(t)$$

where $\sim SS(t)$ means "$SS(t)$ false".

Now that we have such a set R in our clutches, we can ask: is $SS(R)$ true, or false? Think about it!

The point of this example is not "wow, mathematics is weird," true as that might be. The point is rather that not every property we might think of is set-forming – there isn't always a *set* matching a given description. As such we should be very grateful for axioms that say "if your description is of this special approved type... yes, there *is* a set matching that description."

Axiom (of separation)**.** *For any set A, and any property P of sets, there exists a set S such that $\forall t, t \in S \iff (t \in A \text{ and } P(t))$.*

Of course we usually write this as $S = \{t \in A : P(t)\}$ (and again, it is unique by the Axiom of Equality). The point is that it's easy to make sets as subsets of some other set. For example, we can now make the intersection $\cap A$ of a nonempty family of sets; it's a certain subset of the union.

$$\bigcap A := \{t \in \bigcup A : \forall S \in A, t \in S\}$$

To know that this exists *as a set*, not just a description like "the set of all non-self-swallowing sets", we're using the axiom of union and then the axiom of separation. (Obviously nobody ever stops to mention this when they're taking intersections in a proof somewhere.)

Axiom (of power set)**.** *For any set A, there exists a set $P(A)$ such that $\forall t$,*

$$t \in P(A) \iff t \subseteq A.$$

This set $P(A)$ is the **power set** of A, so named because the number of elements $|P(A)|$ in $P(A)$ is $2^{|A|}$, if A finite. (And again, it's necessarily unique.)

So far these axioms have gotten us from one set to another – but how do we even get started? Maybe there are *no* sets at all! If there weren't, all these axioms would be vacuously true, so we'll need something else.

Axiom (of the empty set)**.** *There exists a set $\emptyset$ such that $\forall t, t \notin \emptyset$.*

So the empty set $\emptyset = \{\}$ exists. From there we can build its power set, $\{\{\}\}$, and its power set, $\{\{\}, \{\{\}\}\}$, and so on, and start picking subsets out of those... lots of sets, all finite. (We will talk about infinite sets in section 3.)

Note that there's a shorter form of this axiom: "There exists *a* set." From any set A, we can construct the empty set by taking the subset of all (for example) self-swallowing and also not self-swallowing elements of A. But this version is somehow more concrete because, like the others, it names a necessarily unique set. (Once again, any two empty sets are equal by the axiom of equality.) This is just a matter of taste really.

2. FUNCTIONS

Given two sets A and B, define the **set difference** $A \setminus B$ by $A \setminus B := \{t \in A : t \notin B\}$ (using the axiom of separation). Note that B doesn't have to be a subset of A to talk about this: $A \setminus B = A \setminus (A \cap B)$.

It is frequently useful to be able to make *ordered* pairs (a, b) of sets – for example, we need them to define relations and functions. The set $\{a, b\}$ doesn't cut it, because by the axiom of equality $\{a, b\} = \{b, a\}$. There is a standard trick: define

$$(a, b) := \{\{a\}, \{a, b\}\}.$$

Then a is the unique element of $\cap(a, b)$, and b the unique element of $\cup(a, b) \setminus \cap(a, b)$.

Now that you know the definition $(a, b) := \{\{a\}, \{a, b\}\}$, you can forget it – everybody writes (a, b) and talks about the first and second elements, because they can (using the above rules). The only reason I include it is to emphasize that everything we ever touch is a set or a property of sets.

Actually, we can't make ordered pairs yet – we need another axiom:

Axiom (of pairs)**.** *Given two sets A, B, there exists a set S with $A \in S$ and $B \in S$.*

(Such an S could have extra unwanted elements beyond $\{A, B\}$, but we can just Separate them out.) Note that if $A = B$ then this says that $\{A\}$ is a (one-element) set.

With ordered pairs, we can define relations and functions. With functions around, we can actually define "finite" – a set A is **finite** if every one-to-one map $f : A \to A$ is also onto. More about this later.

We're used to defining a function from A to B as a subset of $A \times B$. (Exercise: what axioms are you using to know that $A \times B$ is a set?) More generally, define a *function* f as a property $f\big((a, b)\big)$ of ordered pairs (a, b), with the condition

$$\forall a, \exists! b \text{ such that } f\big((a, b)\big)$$

(of course Leibniz would write this as $f(a) = b$ but that doesn't fit with our notation for properties).

Axiom (of replacement)**.** *Given a set A, and a function f, there exists a set $f(A)$ such that*

$$\forall t, t \in f(A) \iff \Big(\exists a \in A, \exists b, \text{ such that } f\big((a, b)\big)\Big).$$

This wouldn't be anything new if we assumed f was a function from A to B, namely a subset of $A \times B$ – we'd just define $f(A)$ as the set of second elements of the elements of f (which are ordered pairs). (Exercise: figure out how to build that using the previous axioms mentioned.)

This is a very powerful axiom – too powerful for some people's tastes, who adopt only limited versions of it in their set theories. But they are definitely out of the mathematical mainstream. Feel free to Replace all you want in this class.

3. NUMBERS

An **ordinal** is a set with the following funny condition: every element of it is also a subset!

There are two important examples:

- the empty set
- if O is an ordinal, then $O \bigcup \{O\}$ is an ordinal.

So we can make lots of ordinals, starting from the empty set $\{\}$:

$$\{\{\}\}$$

$$\{\{\},\{\{\}\}\}$$

$$\{\{\},\{\{\}\},\{\{\},\{\{\}\}\}\}$$

$$\{\{\},\{\{\}\},\{\{\},\{\{\}\}\},\{\{\},\{\{\}\},\{\{\},\{\{\}\}\}\}\}$$

$$\{\{\},\{\{\}\},\{\{\},\{\{\}\}\},\{\{\{\},\{\{\}\},\{\{\},\{\{\}\}\}\}\},\{\{\},\{\{\}\},\{\{\},\{\{\}\}\},\{\{\{\},\{\{\}\},\{\{\},\{\{\}\}\}\}\}\}\}$$

etc.

There are some amazing theorems about ordinals (that we won't prove):

Theorem.
- *Every element of an ordinal is an ordinal.*
- *If A,B are two ordinals, either $A = B$, $A \in B$, or $B \in A$.*
- *If $S \subseteq A$ is a nonempty subset of an ordinal, $\cap S$ is an element of S.*

The first two say that we should regard $\in$ as an ordering on ordinals, much like $<$ on the natural numbers. The third says that given any list of ordinals less than an ordinal, that list has a least element (and constructs it: just take the intersection). Remember that this was the crucial property that made induction work, so we can do induction on any ordinal! (We are unlikely to make use of that in this class.)

So where are numbers? If we think of natural numbers as "sizes of finite sets" (including the empty set), then we need to understand what makes two sets "the same size". Say that two sets A, B are **the same size** if there exists a one-to-one and onto function $f : A \to B$. (Actually, set theorists usually call these "the same cardinality", but the name isn't important.) Note that this is an equivalence relation on sets; the identity function $f(a) = a$ shows A has the same size as A, one-to-one and onto functions can be inverted and composed, showing symmetry and transitivity.

Then we *could* think of 3 as something like an equivalence of sets – "the set of all sets with three elements". But this is not a set (use Union on it to make the "set" of all sets; use Separation on that to pick out the non-self-swallowing ones, and we're back in Russell's paradox). Instead, we use the following theorem (we won't prove):

Theorem 1. *For any finite set A, there is a unique ordinal O of the same size.*

So set theorists then *define* the natural numbers as finite ordinals. In other words, "3" is defined to be $\{\{\},\{\{\}\},\{\{\},\{\{\}\}\}\}$, the ordinal with 3 elements.

So far we know of no ways to make infinite sets; all the axioms from before that construct new sets from old ones, only make bigger finite sets. We're ready to state a handy axiom:

Axiom (of infinity)**.** *There exists a set $\mathbb{N}$ such that $n \in \mathbb{N} \iff n$ is a finite ordinal.*

It is easy to show that

- $\mathbb{N}$ is an ordinal

- if $n \in \mathbb{N}$, $n \cup \{n\}$ is too

This latter gives us a function $f : \mathbb{N} \to \mathbb{N}$ that is 1:1 but not onto (it's the add-one function and misses $0 := \{\}$). So $\mathbb{N}$ is not a finite set.

Again, there is a terrible alternate way of stating the axiom of infinity: "There exists an infinite set." Given one, one can (with some work) construct $\mathbb{N}$ from it.

4. The axiom of choice

This is not a big deal in this class – for the finite sort of stuff we'll be studying (mostly), we'll never need the axiom of choice. So if you just skip what's here don't worry much about it. But it is about the only really different standard axiom of set theory, so it's worth seeing it if only to appreciate the others.

We usually take the Cartesian product $P \times Q$ of two sets, and define it as the set of ordered pairs $\{(p, q) : p \in P, q \in Q\}$. From there one can define $P \times Q \times R$ as $(P \times Q) \times R$, and so on,[3] for any finite number. But what about an infinite product?

Given a set $A = \{a, b, c, \ldots\}$ of sets, define the **Cartesian product** $\times A$ of A as the set of functions

$$\times A := \left\{ f : A \to \bigcup A \quad : \quad \forall a \in A, f(a) \in a \right\}$$

This is a very mysterious definition at first. But imagine that $A = \{P, Q\}$. Then this is the set of functions from the two-element set A to $P \cup Q$, such that the element $P \in A$ is mapped to an element of P, and the element $Q \in A$ is mapped to an element of Q – in short, a function giving a choice of an element of P and an element of Q. Which is what the Cartesian product is supposed to do.[4]

In the case A a finite set and each element a finite set, it's easy to show that the size $|A|$ of A is the product of the sizes of A's elements.

Sometimes it's easy to think of elements of a product $\times A$. Let A be the set of all pairs of shoes; each element of A is a two-element set (the two shoes). Then "the left shoe" is a function picking a shoe from each pair.

Sometimes it's obviously impossible: if one of the elements of A is the empty set, $A = \{\ldots, \{\}, \ldots\}$, then there can be no functions from A to $\bigcup A$ with the required property – they would have to take $\{\}$ to an element of $\{\}$, and there are none. Thinking about the product rule for sizes mentioned above, this is the statement that when you multiply a bunch of numbers and one of them is zero, the product is zero.

Axiom (of choice). *If A is a set of nonempty sets, then $\times A$ is nonempty. In other words, there exists a function $A \to \bigcup A$ choosing an element from each set $P \in A$.*

On the face of it, this is very intuitive, or rather, the inverse statement is very unintuitive – how could there *not* be a way to simultaneously choose an element from each set (since by assumption, they each do have elements)? In particular, one might hope to prove it from the other axioms (rather than separately assume it).

[3]Annoyance: this is not the *same* set as $P \times (Q \times R)$. All we have instead is an obvious 1:1 and onto function from one to the other.

[4]Again, this set $\times\{P, Q\}$ is not actually *equal* to the set of ordered pairs $\{(p, q)\}$ – there's just a natural one-to-one and onto function from one to the other.

Note that this axiom is very different from the others: they typically asserted that a set with a certain exact description actually existed, and then the axiom of equality said that it was unique. Here the assertion is that $\exists t, t \in \prod A$, but if there's one there's automatically lots of them (unless all the elements of A are of size 1).

Most theorems that one can only prove with the axiom of choice have a "there exists" in the statement, and one can never pin down an example of the thing said to exist. If one could, one wouldn't need to use the axiom of choice.

If A is the set of all pairs of *socks*, then we don't have a concrete way to pick an element from each pair (put them through a dryer?), and we need the axiom of choice to assert that there is *some* way.

Theorem (The Banach-Tarski "paradox", doubling the sphere)**.** *There exists a way to cut the sphere in $\mathbb{R}^3$ into eight subsets $A_1, A_2, \ldots, A_8$, such that $A_1 - -A_4$ can be rotated and put back together to form a sphere of the same size, and likewise $A_5 - -A_8$.*

Don't try hard to visualize this – each of these subsets are bad, like the rationals as a subset of the reals, only much much worse. And this is not a paradox, it's a theorem – proven using the axiom of choice. So if you don't like it, fine, don't assume the axiom of choice.

Theorem. *The following two axioms are equivalent:*

- *the axiom of choice*
- *for any two sets A and B, either there exists a 1:1 function from $A \to B$, or there exists a 1:1 function from $B \to A$*

If you like to use "exists a 1:1 function from $A \to B$" as the definition of "A's size is smaller than or equal to B's size" (and set theorists *do* like this definition), then this second axiom says that one can always compare sizes. Many versions of the axiom of choice are intuitive, and one can read about many others in e.g. [J].

Theorem. *1. [Gödel] There exists a model of set theory in which all the other usual axioms are true, and the axiom of choice is true.*
2. [Cohen] There exists a model of set theory in which all the other usual axioms are true, but the axiom of choice is false.

So unless the other axioms let us prove false things (and maybe they do! nobody knows, but nobody thinks it very likely), we'll never be able to prove or disprove the axiom of choice. Basically, most people are happy to use it all the time; the 7th floor of Evans is one of the only places in the world where people worry about it much.

References

[J] P.T. Johnstone, Notes on logic and set theory, Cambridge University Press (1987).

E-mail address: `allenk@math.berkeley.edu`