Set Theory (MATH 6730)

The Axioms of Set Theory and Some Basic Consequences

We will use the axiom system *Zermelo–Fraenkel set theory with Choice*, abbreviated ZFC. The axioms are formulas in the language $\mathcal{L}$ (with no function symbols, no constant symbols, and a single relation symbol: the 2-place symbol $\in$). In the 'intended model', the objects are sets (hence every set is a family of sets), which will be reflected in the terminology we introduce.

1. Extensionality: Sets are determined by their members; more precisely, two sets are equal iff they have the same members. Formally:

$$\mathsf{Ext}^\sharp \qquad \forall x\,\forall y\,\big(x = y \leftrightarrow \forall z\,(z \in x \leftrightarrow z \in y)\big).$$

Notice, however, that the forward direction of the first $\leftrightarrow$ is a consequence of our logical axioms, therefore we only include the following formula in ZFC:

$$\mathsf{Ext} \qquad \forall x\,\forall y\,\big(\forall z\,(z \in x \leftrightarrow z \in y) \to x = y\big).$$

Remark 1. Why do we choose Ext to be an axiom rather than $\mathsf{Ext}^\sharp$? $\mathsf{Ext}^\sharp$ expresses better (more fully) what we want to require by the extensionality axiom; however, Ext is simpler and still equivalent to $\mathsf{Ext}^\sharp$. When we will discuss models of ZFC (e.g., to construct a model in which CH fails), and have to check that the axioms of ZFC hold, it will be useful to have as simple axioms as possible. This is the main reason for preferring Ext to $\mathsf{Ext}^\sharp$.

We will follow similar practice with some of the other axioms as well, keeping in mind that the simpler, weaker versions remain equivalent to the 'intended meaning' — at least in the presence of the other axioms of ZFC.

The informal statement that "Ext together with our logical axioms implies $\mathsf{Ext}^\sharp$" is meant to say rigorously that $\mathsf{Ext} \vdash \mathsf{Ext}^\sharp$. Here is a sketch of how to prove this in our formal proof system:[1]

1. Since Ext is a sentence, by the Generalization Theorem (3.11(iv)) we see that it suffices to prove $\mathsf{Ext} \vdash x = y \leftrightarrow \forall z\,(z \in x \leftrightarrow z \in y)$.
2. Since $(\varphi \to \psi) \to \big((\psi \to \varphi) \to (\varphi \leftrightarrow \psi)\big)$ is a tautology for any formulas φ, ψ, we can combine Metatheorem 3.11(ii) with MP to get that it suffices to prove
   (a) $\mathsf{Ext} \vdash x = y \to \forall z\,(z \in x \leftrightarrow z \in y)$ and
   (b) $\mathsf{Ext} \vdash \forall z\,(z \in x \leftrightarrow z \in y) \to x = y$.
3. There is a short formal proof establishing (b), which uses logical axioms of the form $\forall v\,\varphi \to \varphi$ from (Ax2).
4. For (a), we use first the Deduction Theorem (3.11(iii)), and then the Generalization Theorem (noting z is not free in Ext or $x = y$) to obtain that it suffices to verify
   (a)$'$ $\{\mathsf{Ext}, x = y\} \vdash z \in x \leftrightarrow z \in y$.
   To establish (a)$'$, one can write down a formal proof which uses the logical axiom $x = y \to (z \in x \to z \in y)$. (Ideas similar to those described in step 2 will be useful to deal with $\leftrightarrow$, and you might also need to show that $x = y \vdash y = x$.)

Later on we will usually be satisfied with informal proofs, but we will be careful to make sure that we only use informal proofs which can be formalized (if necessary) in our formal proof system.

[1] We will use the Metatheorems 3.11(i)–(viii) from the handout 'Background in Logic'.

2. Comprehension: Given a set x and a property φ (of sets), there is a set whose elements are exactly those elements z of x which satisfy φ. By extensionality, there is exactly one such set. There are two important facts to note here.

- This axiom does not allow 'unrestricted comprehension', i.e., it does not imply that there is a set whose members are all sets with property φ.[2]
- Since we work in first-order logic, φ must be an $\mathcal{L}$-formula (in informal discussions, a property expressible by an $\mathcal{L}$-formula).[3]

Comprehension is an axiom scheme, one axiom for every formula φ, which may have other free variables than z. If φ has its free variables among $z, w_1, \ldots, w_n$, we may write $\varphi(z, w_1, \ldots, w_n)$ instead of φ. So formally, the comprehension axioms are the formulas

$$\mathsf{Cmpr} \qquad \forall x\, \forall w_1 \ldots \forall w_n\, \exists y\, \forall z\, \big(z \in y \leftrightarrow (z \in x \wedge \varphi(z, w_1, \ldots, w_n))\big).$$

The set y asserted to exist (and unique by extensionality) is denoted by $\{z \in x : \varphi(z, w_1, \ldots, w_n)\}$ or $\{z \in x : \varphi\}$.

Remark 2. This notation (and other similar notation introduced later, e.g., $\emptyset$, $\bigcup A$, $\mathcal{P}(A)$) is introduced for our convenience; it is not part of our formal language. However, the use of this notation can be eliminated if necessary as follows: $\{z \in x : \varphi(z, w_1, \ldots, w_n)\}$ is the unique set $y = y(x, w_1, \ldots, w_n)$ (depending on $x, w_1, \ldots, w_n$) such that

$$\text{(1)} \qquad \mathsf{ZFC} \vdash \forall z\, \big(z \in y \leftrightarrow (z \in x \wedge \varphi(z, w_1, \ldots, w_n))\big).$$

The first application of the comprehension axioms is

Claim 3. *There exists a unique set with no members.*

This is called *the empty set*, and is denoted by $\emptyset$.

Proof of Claim 3. We need that there exists a set x; this follows by logic (namely by the fact that $\vdash \exists x\, x = x$). Then the set $y = \{z \in x : z \neq z\}$ exists by comprehension, and has no elements. The uniqueness follows by extensionality. $\square$

---

[2]This will avoid Russell's paradox, where φ would be $z \notin z$.

[3]This will avoid paradoxes like Berry's Paradox: The set of all natural numbers not definable in fewer than 12 words.

3. Pairing: For any (not necessarily distinct) sets x, y there exists a set whose members are exactly x and y. In a formula:

$$\mathsf{Pair}^\sharp \qquad \forall x\, \forall y\, \exists p\, \forall u\, \big(u \in p \leftrightarrow (u = x \lor u = y)\big).$$

The set p asserted to exist (and unique by extensionality) is denoted by $\{x, y\}$. The set $\{x, y\}$ is sometimes referred to as an *unordered pair*. The unordered pair $\{x, x\} = \{x\}$ is called a *singleton*.

By comprehension, such a set p exists if there is some set z which has x and y as members (but may have other members); indeed, from z we can get $p = \{x, y\}$ by comprehension as the set $\{u \in z : u = x \lor u = y\}$. Therefore — for the reason discussed in Remark 1 — we include only the following weaker statement in ZFC:

$$\mathsf{Pair} \qquad \forall x\, \forall y\, \exists z\, (x \in z \;\land\; y \in z).$$

Formally, the relationship between Pair and $\mathsf{Pair}^\sharp$ can be expressed as follows:

$$\mathsf{Pair}^\sharp \vdash \mathsf{Pair} \qquad \text{and} \qquad \{\mathsf{Cmpr}, \mathsf{Pair}\} \vdash \mathsf{Pair}^\sharp.$$

Definition 4. For any sets x and y we define *the ordered pair* (x, y) to be the set $\{\{x\}, \{x, y\}\}$.

By the pairing axiom (applied three times), $\{\{x\}, \{x, y\}\}$ is indeed a set. The following claim, which is straightforward to prove, justifies the name ‘ordered pair’.

Claim 5. *For arbitrary sets x, y, u, v we have*

$$(x, y) = (u, v) \qquad \textit{iff} \qquad x = u \;\; \textit{and} \;\; y = v.$$

4. Union: For any set (i.e., family of sets) $\mathcal{A}$, there is a set whose elements are exactly those sets which are members of members of $\mathcal{A}$. Formally:

$$\mathsf{Uni}^\sharp \qquad \forall \mathcal{A}\, \exists U\, \forall x\, \big(x \in U \leftrightarrow \exists A\, (x \in A \;\land\; A \in \mathcal{A})\big).$$

The set U asserted to exist (and unique by extensionality) is denoted by $\bigcup \mathcal{A}$, and is called the *union* of (the family) $\mathcal{A}$.

By comprehension, we can obtain the set $U = \bigcup \mathcal{A}$ from any set B which has all members of members of $\mathcal{A}$ among its elements (but may have other elements); indeed, then $\bigcup \mathcal{A} = \{x \in B : \exists A\, (x \in A \;\land\; A \in \mathcal{A})\}$. Therefore it suffices to include only the following weaker statement in ZFC:

$$\mathsf{Uni} \qquad \forall \mathcal{A}\, \exists B\, \forall x\, \big(\exists A\, (x \in A \;\land\; A \in \mathcal{A}) \to x \in B\big).$$

As before, the relationship between Uni and $\mathsf{Uni}^\sharp$ is the following:

$$\mathsf{Uni}^\sharp \vdash \mathsf{Uni} \qquad \text{and} \qquad \{\mathsf{Cmpr}, \mathsf{Uni}\} \vdash \mathsf{Uni}^\sharp.$$

Before stating the power set axiom, we need the definition of a subset of a set.

Definition 6. For any two sets x, y we say that x is a *subset* of y, and write $x \subseteq y$, if every member of x is a member of y. If $x \subseteq y$ and $x \neq y$, we say that x is a *proper subset* of y, and write $x \subsetneq y$ (or just $x \subset y$).

Formally, $x \subseteq y$ and $x \subsetneq y$ may be viewed as abbreviations for the formulas

$$\forall z\,(z \in x \rightarrow z \in y) \quad \text{and} \quad \neg x = y \wedge \forall z\,(z \in x \rightarrow z \in y),$$

respectively.

5. Power set: For any set A there is a set whose elements are exactly the subsets of A. In a formula:

Pset$^\sharp$ $$\forall A\,\exists P\,\forall x\,(x \in P \leftrightarrow x \subseteq A).$$

The set P asserted to exist here (and unique by extensionality) is denoted by $\mathcal{P}(A)$, and is called *the power set* of A. As before, comprehension implies that we can get the existence of the set $\mathcal{P}(A)$ if there is a set Z that has all subsets of A as members (but may have other members); namely, $\mathcal{P}(A) = \{x \in Z : x \subseteq A\}$. Therefore it suffices to include in ZFC the following weaker statement:

Pset $$\forall A\,\exists Z\,\forall x\,(x \subseteq A \rightarrow x \in Z).$$

As before, the relationship between Pset and $\mathsf{Pset}^\sharp$ is the following:

$$\mathsf{Pset}^\sharp \vdash \mathsf{Pset} \qquad \text{and} \qquad \{\mathsf{Cmpr}, \mathsf{Pset}\} \vdash \mathsf{Pset}^\sharp.$$

The axioms discussed so far are sufficient to introduce many of the set constructions used in elementary set theory (and to show that the constructions, when applied to sets, yield sets).

The axioms discussed so far imply the following statements.

Claim 7. *Let $A, B, \mathcal{F}$ be arbitrary sets.*

(i) *If $\mathcal{F} \neq \emptyset$, then there exists a set S whose members are exactly the sets that are elements of all members of $\mathcal{F}$. Formally:*

$$\forall \mathcal{F} \left(\exists F_0\, F_0 \in \mathcal{F} \to \exists S\, \forall x\, (x \in S \leftrightarrow \forall F\, (F \in \mathcal{F} \to x \in F))\right).$$

(ii) *There exists a set C whose members are exactly the ordered pairs (x, y) with $x \in A$ and $y \in B$. Formally,*

$$\forall A\, \forall B\, \exists C\, \forall u\, (u \in C \leftrightarrow \exists x\, \exists y\, (x \in A \,\wedge\, y \in B \,\wedge\, u = (x, y)).$$

($u = (x, y)$ is an abbreviation for a formula!)

The sets asserted to exist in (i)–(ii) *are unique.*

Idea of Proof. (i) Let $F_0 \in \mathcal{F}$, and argue that

$$S = \{x \in F_0 : \forall F\, (F \in \mathcal{F} \to x \in F)\}$$

satisfies the requirements.

(ii) Show that

$$C = \{u \in \mathcal{P}(\mathcal{P}(\bigcup\{A, B\})) : \exists x\, \exists y\, (x \in A \,\wedge\, y \in B \,\wedge\, u = (x, y)\}$$

satisfies the requirements. $\square$

Definition 8. Let $A, B, \mathcal{F}$ be arbitrary sets.

- If $\mathcal{F} \neq \emptyset$, the (unique) set S from Claim 7(i) is called the *intersection* of (the family) $\mathcal{F}$, and is denoted by $\bigcap \mathcal{F}$.
- $\bigcap\{A, B\}$ is called the intersection of A and B, and is denoted by $A \cap B$. Similarly, $\bigcup\{A, B\}$ is called the union of A and B, and is denoted by $A \cup B$.
- A and B are called *disjoint* if $A \cap B = \emptyset$.
- The set $\{x \in A : \neg\, x \in B\}$ (indeed a set by comprehension!) is called the *difference* of A and B, or the *relative complement* of B in A, and is denoted by $A \setminus B$.
- The set C from Claim 7(ii) is called the *cartesian product* of A and B, and is denoted by $A \times B$.

Informally, we may write

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

instead of the more precise description of $C = A \times B$ in the proof of Claim 7(ii).

Claim 9. *If R is a set having an ordered pair (x, y) as a member, then $x, y \in \bigcup\bigcup R$.*

Definition 10. A *relation* is a set of ordered pairs. For any set R,

- the *domain* of R is the set $\mathrm{dmn}(R) = \{x \in \bigcup\bigcup R : \exists y\, (x, y) \in R\}$;
- the *range* of R is the set $\mathrm{rng}(R) = \{y \in \bigcup\bigcup R : \exists x\, (x, y) \in R\}$;
- the *inverse* or *converse* of R is the relation

$$R^{-1} = \{(y, x) \in \mathrm{rng}(R) \times \mathrm{dmn}(R) : (x, y) \in R\}.$$

Note that by the comprehension and union axioms and by Claim 7(ii), $\mathrm{dmn}(R)$, $\mathrm{rng}(R)$, R^{-1} are indeed sets whenever R is a set.

Definition 11. A *function* is a relation f such that for every $x \in \mathrm{dmn}(f)$ there is a unique $y \in \mathrm{rng}(f)$ such that $(x, y) \in f$. The unique y with $(x, y) \in f$ is denoted by $f(x)$, and called the *value of f at x*.

- By $f \colon A \to B$ we mean that f is a function, and A, B are sets satisfying $\mathrm{dmn}(f) = A$ and $\mathrm{rng}(f) \subseteq B$.[4]
- If $f \colon A \to B$ and $C \subseteq A$, then the *restriction of f to C*, is the function $f{\restriction}C = f \cap (C \times B)$; clearly, $f{\restriction}C \colon C \to B$.
- If $f \colon A \to B$ and $C \subseteq A$, then the *image of C* under f is the set

$$f[C] = \mathrm{rng}(f{\restriction}C) \; \big(= \{y \in B : \exists x\, (x \in C \wedge f(x) = y)\} = \{f(x) : x \in C\}\big).$$

- If $f \colon A \to B$ and $D \subseteq B$, then the *preimage of D* under f is the set

$$f^{-1}[D] = \{x \in A : f(x) \in D\}.$$

Note that under the assumptions in Definition 11, $f{\restriction}C$ is indeed a function, and $f[C]$ and $f^{-1}[D]$ are sets by comprehension.

Remark 12. For a function $f \colon A \to B$, the distinction between $f(x)$ $(x \in A)$ and $f[X]$ $(X \subseteq A)$ is important, because for an element $a \in A$ which is also a subset of A, the sets $f(a)$ and $f[a]$ may be different.

The definitions: *composition* of functions, *injective* (or *one-to-one*) function, *surjective* (or *onto*) function, and *bijective* function are as usual; see pp. 72–73 of [1].

An alternative notation we might use for a function f with domain I is $\langle f(i) : i \in I\rangle$ or $\langle f_i : i \in I\rangle$. With this notation, if A is a function with domain I, i.e., A is a function $\langle A_i : i \in I\rangle$, then we define

$$\bigcup_{i\in I} A_i = \bigcup \mathrm{rng}(A) \quad (I \text{ arbitrary}),$$

$$\bigcap_{i\in I} A_i = \bigcap \mathrm{rng}(A) \quad (I \neq \emptyset). \tag{2}$$

[4] Thus, in the notation $f \colon A \to B$, the arrow $\to$ is <u>not</u> the logical connective for 'if … then'!

Before returning to the remaining axioms of ZFC, we will discuss

- further abbreviations to simplify formulas, and
- *classes*, which are collections of sets that are 'too big' to be sets.

Restricted quantifiers and the quantifier $\exists!$. For every $\mathcal{L}_C$-formula φ and for arbitrary variables x, y we use the following abbreviations, called *restricted quantifiers*:

$$\forall x \in y\ \varphi \quad \text{abbreviates} \quad \forall x\,(x \in y \rightarrow \varphi), \quad \text{and}$$
$$\exists x \in y\ \varphi \quad \text{abbreviates} \quad \exists x\,(x \in y\ \wedge\ \varphi).$$

The quantifier $\exists!$, which stands for 'there exists a unique ...' is defined as follows: for every $\mathcal{L}_C$-formula φ and for every variable x,

$$\exists!\, x\, \varphi \quad \text{abbreviates} \quad \exists x\,\big(\varphi\ \wedge\ \forall y\,(\mathrm{Subf}_y^x(\varphi) \rightarrow x = y)\big)$$

where y is the first variable such that $y \neq x$ and y does not occur in φ.

Classes, class relations, and class functions. We will soon see that the axioms of ZFC imply that $\{x : x = x\}$ is not a set; that is, there is no set which contains all sets as members. Nevertheless, it is intuitively indispensable to allow ourselves to talk about this collection, the collection of all sets. More generally, for every formula $\varphi \equiv \varphi(x)$, the collection of all sets x satisfying φ will be called *the class of sets defined by* φ. In particular, if $\varphi \equiv x = x$, then we get the class $\mathbf{V}$ of all sets. A *class* is a collection of sets defined by a formula. A *proper class* is a class that is not a set.

For any formula $\psi \equiv \psi(x, y)$, the collection of all ordered pairs (x, y) of sets satisfying ψ is called *the class relation defined by* ψ. For example, $\subseteq$ and $\subsetneq$ are class relations defined by the formulas displayed after Definition 6. A *class relation* is a collection of ordered pairs of sets defined by a formula $\psi \equiv \psi(x, y)$. If $\mathbf{A}$ is a class (defined by a formula $\varphi \equiv \varphi(x)$), then the class relation $\mathbf{R}$ defined by ψ is a *class relation on* $\mathbf{A}$ if $\mathbf{R} \subseteq \mathbf{A} \times \mathbf{A}$, that is, $\mathsf{ZFC} \vdash \forall x\, \forall y\, \big(\psi(x, y) \rightarrow (\varphi(x)\ \wedge\ \varphi(y))\big)$.

If $\mathbf{A}$ is a class defined by a formula $\varphi \equiv \varphi(x)$, then a formula $\psi \equiv \psi(x, y)$ defines a *class function* on $\mathbf{A}$ if the class relation $\mathbf{F}$ defined by $\psi(x, y)$ has the property that for every set $x \in \mathbf{A}$ there is a unique set y such that $(x, y) \in \mathbf{F}$. More formally — but allowing classes in restricted quantifiers —, this property can be expressed as follows:

$$\mathsf{ZFC} \vdash \forall x \in \mathbf{A}\, \exists! y\, \psi(x, y). \tag{3}$$

Since $x \in \mathbf{A}$ means "x satisfies φ", the formally correct version is the following:

$$\mathsf{ZFC} \vdash \forall x\, \big(\varphi(x) \rightarrow \exists! y\, \psi(x, y)\big).$$

The unique y asserted to exist for each $x \in \mathbf{A}$ in (3) is denoted by $\mathbf{F}(x)$.

We have seen several class functions on $\mathbf{V}$; for example, the class function $\bigcup\colon \mathcal{A} \mapsto \bigcup \mathcal{A}$ is defined by the formula

$$\upsilon(x, y) \equiv \forall z \left(z \in y \leftrightarrow \exists u \, (z \in u \wedge u \in x)\right),$$

while the class function $\mathcal{P}\colon A \mapsto \mathcal{P}(A)$ is defined by the formula

$$\pi(x, y) \equiv \forall z \, (z \in y \leftrightarrow z \subseteq x).$$

$\bigcap\colon \mathcal{F} \mapsto \bigcap \mathcal{F}$, as we have defined it, is a class function on $\mathbf{V} \setminus \{\emptyset\}$. To extend it to a class function on $\mathbf{V}$, we need to define $\bigcap \mathcal{F}$ for $\mathcal{F} = \emptyset$. In [1] this is done by stipulating that $\bigcap \emptyset = \emptyset$. (If this definition is adopted, then the equality in (2) also holds for $I = \emptyset$.)

6. Infinity: There exists a set which has $\emptyset$ as its member, and is closed under the class function $x \mapsto x \cup \{x\}$ defined by the formula

$$\psi(x, y) \equiv \forall z \left(z \in y \leftrightarrow (z \in x \vee z = x)\right);$$

our intuition is that this set is 'infinite'. The formula of the axiom (containing some abbreviations) is the following:

Inf $$\exists u \, (\emptyset \in u \wedge \forall x \in u \, \forall y \left(\psi(x, y) \rightarrow y \in u\right).$$

7. Replacement: If the domain of a (class) function is a set, then its range is also a set. As we saw with some of the earlier axioms, comprehension implies this if we only assume the following weaker statement as our axiom: If the domain of a (class) function is a set, then there is a set which contains its range as a subset.

Replacement is an axiom scheme, just as comprehension; we have an axiom for every formula φ that defines a (class) function on a set:

Repl $$\forall A \, \forall w_1 \, \ldots \, \forall w_n \left(\forall x \in A \, \exists! y \, \varphi(x, y, A, w_1, \ldots, w_n) \right. \\ \left. \rightarrow \exists Y \, \forall x \in A \, \exists y \in Y \, \varphi(x, y, A, w_1, \ldots, w_n)\right).$$

8. Foundation: Every nonempty set is disjoint from at least one of its elements. Formally,

$$\forall x\,(\neg\, x = \emptyset \to \exists y \in x \; x \cap y = \emptyset),$$

or without abbreviations (except for using $\wedge$ and $\exists$),

$$\mathsf{Fnd} \qquad \forall x\,\big(\exists z\, z \in x \to \exists y\,(y \in x \;\wedge\; \neg\exists z\,(z \in x \;\wedge\; z \in y))\big).$$

Claim 13. (i) *No set A satisfies $A \in A$.* (ii) *No sets A, B satisfy $A \in B \in A$.*

Proof of (i). Assume that there is a set A with $A \in A$, and let $x = \{A\}$. Then $x \neq \emptyset$, so by the axiom of foundation, x has an element disjoint from x. But the only element of x is A, and $x \cap A \neq \emptyset$, as $A \in x$ and $A \in A$. This contradiction proves (i). □

Corollary 14. $\mathbf{V}$ *is not a set; that is, there is no set that has all sets as elements.*

Proof. If $\mathbf{V}$ was a set, we would have $\mathbf{V} \in \mathbf{V}$, contradicting Claim 13. Another argument, which does not rely on the axiom of foundation, would be this: If $\mathbf{V}$ is a set, then by comprehension, so is $\mathbf{S} = \{A \in \mathbf{V} : A \notin A\}$ from Russell's Paradox, which we saw is impossible. □

9. Choice: For any family $\mathcal{A}$ of pairwise disjoint, nonempty sets there is a set C which has exactly one element from every set in $\mathcal{A}$. We can think of C — or more precisely, $C \cap \bigcup \mathcal{A}$ — as describing a way to simultaneously choose one element from every set in $\mathcal{A}$. Formally:

$$\begin{aligned}\mathsf{AC} \quad \forall \mathcal{A}\big((\forall x \in \mathcal{A}\, x \neq \emptyset \;\wedge\; \forall x \in \mathcal{A}\, \forall y \in \mathcal{A}\,(x \neq y \to x \cap y = \emptyset))& \\ \to \exists C\, \forall x \in \mathcal{A}\, \exists! y \in x\, y \in C\big).&\end{aligned}$$

There are many equivalent formulations of the Axiom of Choice; the one above is among the easiest to state. Later in the course we will discuss several equivalent formulations of the Axiom of Choice and other statements equivalent to this axiom.