Introduction to set theory

J. D. Monk

January 18, 2007

These notes are intended to give an introduction to set theory sufficient for further work in abstract mathematics at the undergraduate or graduate level. As background we assume a basic knowledge of elementary mathematics, including some elementary set theory, familiarity with the notion of a function, natural numbers and their properties, and in general, mathematical "maturity". Although we will begin with elementary set theory and the natural numbers, the material may be hard to follow without having an intuitive grasp of the notions already.

One bit of background that is very important is the use in mathematics of elementary logic. We begin with some comments about this. Mathematicians use ordinary logical language such as "not", "implies", etc., in a rather special way. Here are some salient points that may cause difficulties:

• Two truth values. Basic to the logical framework is the assumption that every mathematical statement is either true or false. There is no middle ground like "sometimes true, sometimes false". Somewhat formally, this means that for every mathematical statement φ (whether true or false) we can assert "φ or not φ but not both". This does not mean that we expect to be able to prove from our set theoretical axioms either "φ" or "not φ"; in fact, it is a deep result of logic that this is not in general possible (Gödel's incompleteness theorem).

• Implications. When we say "φ implies ψ", we mean merely that either φ is false, or ψ is true. There does not have to be any kind of connection between the two statements. If we know for sure that φ is false and then assert "φ implies ψ", we sometimes say that "φ implies ψ" is *vacuously true*. For example $0 = 1$ implies that the Riemann hypothesis holds. While somewhat ridiculous, this statement will be recognized by any mathematician as true. Another example of this strange use of "implies" is that any statement whatsoever implies something proved to be a theorem in mathematics.

• Arguments by contradiction. Frequently we want to prove a statement S, and do so by assuming that S is false and deriving a contradiction from this assumption. Then we are justified in saying the S is true after all. Thus at a given point in an argument, we have a collection Σ of assumptions, from which we want to derive S. We argue for a while from the assumptions $\Sigma \cup \{\text{not}(S)\}$, and obtain a contradiction, for example $0 = 1$. Then a rule of logic says that S is a consequence of Σ.

• Implications and converses. When we prove that S implies T, we have not said anything about the *converse*, that T implies S. Frequently one of the first things a mathematician does after having proven an implication is to ask the question, whether the converse is also true.

• Contrapositive. If we have proved that S implies T, then we can say on the basis of logic alone that T being false implies that S is false; the second implication here is called the *contrapositive* of the first. This contrapositive is actually logically equivalent to the given implication, and it is sometimes easier to prove a contrapositive than to prove the implication itself.

• "or". "φ or ψ" means that either φ is true, or ψ is true, or both are true. As in the case of implication, there does not have to be any relationship between the statements. Thus, for example "the empty set has members, or Fermat's last theorem is true" is a true statement of mathematics, since Fermat's last theorem was recently proved, although the empty set does not have any members.

• Equivalences. When we say that S is equivalent to T, we are really stating two implications: that S implies T, and that T implies S. Usually we have to prove each implication separately. We use "iff " to abbreviate "if and only if". So "S iff T" means that S and T are equivalent.

• Existence. When we claim that some mathematical object exists, we generally do not claim that we have a construction of it. We may prove existence by contradiction, assuming that the object does not exist and deriving a contradiction. To say that there is an x such that a statement S holds is the same thing as saying that it is not the case that for every x S fails to hold.

• Equality. When we write $x = y$ we mean that x and y denote the same object. Some properties of equality are purely logical. For example, $x = x$; if $x = y$ then $y = x$ and if $x = y$ and $y = z$ then $x = z$. Furthermore we can "substitute equals for equals": if we know that a statement $S(x)$ holds for x and we also know that $x = y$, then we can infer that $S(y)$ holds.

• "For all". "For all $x \ldots$" means just "for each and every object $x, \ldots$".

• Logic, rigorously. To put mathematics on a completely rigorous foundation, one should develop logical notions such as the above in a rigorous fashion. We are not going to do this, since it would take too much time. The logical underpinnings are subtle, and take some getting use to. The idea in a nutshell is to create a formal language (for logic and set theory), define some logical axioms, define some rules of inference for going from known results to new ones, and define a proof as a finite sequence built up according to these axioms and rules. It is a good idea to keep this idea of logic in mind when going through this introduction to set theory. (Even in a very rigorous development of the basics of set theory it is not necessary to know more advanced things in logic like the completeness and incompleteness theorems.)

1. The axiomatic framework, and simple operations on sets

Although we are not going to make a formal development of set theory, in the interest of rigor we will base the development on some axioms. These axioms are informal, but they can be made completely precise. Altogether there will be 9 axioms, but we are not going to introduce them all at once. They are all listed together in Appendix A for ease of reference.

The basic notion of a set is left without a definition. We have to start from some undefined notions, and set is the notion that is basic for set theory. We think of a set as some theoretical construct which gathers together various objects into a whole. This theoretical construct is undefined; one can argue whether or not it has a real existence in any sense. Thus a set is some object which has *members*, and we write $a \in b$ to indicate that a is a member of b. We also write $a \notin b$ to indicate that a is not a member of b. What are the members? They are also sets. In fact, it greatly simplifies the whole framework to assume that everything that we talk about is a set. So a set a is a member of a more complicated set b, which in turn may be a member of another set c, etc. Any reasonable method of forming new sets from old is allowed, intuitively. Our axioms set precise limits on this, however. As an instance of the principle of substitution of equals for equals mentioned in the introduction we have the following logical principle:

If $x = y$ and $u = v$, then $x \in u$ implies that $y \in v$.

How can we imagine that everything is a set, and still be able to develop mathematics, with all of its concepts like functions, fractions, planes, integrals, etc.? Well, we can start everything out from the empty set, a set which has no members, and somehow build larger sets, even infinite sets, by some well-defined processes from the empty set. For example, we can form in succession

$\emptyset$	the empty set
$\{\emptyset\}$	the set whose only member is the empty set
$\{\emptyset, \{\emptyset\}\}$	the indicated set (it has two members)
$\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$	a set with three members
etc, etc.	

We imagine somehow continuing this forever, even beyond the famous three dots $\ldots$; much later in these notes we will come back to this point. With many sets available to us intuitively, we can define ordinary notions of mathematics in terms of these sets. That will be the main purpose of these notes.

Now even without any axioms we can make the following important definition. A set A is a *subset* of a set B, in symbols $A \subseteq B$, iff any element of A is an element of B. We also say that A is *contained* in B, or *included* in B. We call A a *proper subset* of B if it is a subset of B and is different from B; then we write $A \subset B$. We also write $B \supseteq A$ and $B \supset A$ in place of $A \subseteq B$ and $A \subset B$, respectively. Note that $A \subseteq B$ and $A \subset B$ mean entirely different things from $A \in B$.

Warning: many books use $\subset$ where we use $\subseteq$, and $\subsetneqq$ where we use $\subset$.

Now we start in on our axioms.

Axiom 0. There is at least one set.

We label this as Axiom 0 because it is really an axiom of logic.

Axiom 1. (*Extensionality*) **If a and b have the same members, then $a = b$.**

According to this axiom, no matter how a and b are defined, if it turns out that they have the same members, then they are equal. It may be difficult to prove that two sets have the same members. For example let $A = \{0\}$ (the set whose only member is the integer 0), and let $B = \{0\}$ if Riemann's hypothesis holds, and $B = \{1\}$ otherwise. Then it is not yet known whether A and B have the same members. Here we are using some simple notions, $\{a\}$ and integers 0 and 1, which have not yet been introduced, but whose meaning should be clear.

Axiom 2. (*Pairing*) **For any sets a, b there is a set c which has exactly a and b as members.**

According to the extensionality axiom, the set asserted to exist in this axiom is uniquely determined by a and b. We denote it by $\{a, b\}$. We call $\{a, b\}$ the *unordered pair* a, b. Note that in this axiom, although we use the different letters a, b, we do not assume that they denote different things. This is another commonplace mathematical usage. For any sets a, b, x, we have $x \in \{a, b\}$ iff $x = a$ or $x = b$ (or both, if $a = b$).

We also define $\{a\}$ to be $\{a, a\}$. Thus for any sets x and a, $x \in \{a\}$ iff $x = a$. We call $\{a\}$ *singleton* a.

Axiom 3. (*Union*) **For any set A there is a set B such that for all x, $x \in B$ iff $x \in a$ for some $a \in A$.**

Thus here we think of A as being a whole collection of various sets a, and B is the union of all those sets. Since the set B is unique by the extensionality axiom, we can define $\bigcup A$ to be that set. We call $\bigcup A$ the *union of* A. Thus for any x, $x \in \bigcup A$ iff x is a member of at least one member of A. We also define $a \cup b = \bigcup\{a, b\}$, and call it the *union of* a *and* b. So for any x, $x \in a \cup b$ iff $x \in a$ or $x \in b$.

We can define other sets easily now; for example,

$$\begin{aligned} \{a, b, c\} &= \{a, b\} \cup \{c\}; \\ \{a, b, c, d\} &= \{a, b, c\} \cup \{d\}; \quad \text{etc.} \end{aligned}$$

We can extend this notation indefinitely, but without the notion of natural number (yet) we cannot make a general definition here.

Of course we would now like to define the other elementary operations on sets: intersections and complements. To do this it is useful to introduce a much more general axiom, as follows.

Axiom 4. (*Comprehension*) **If A is a given set and $P(x)$ is a property of sets x, then there is a set B whose elements are exactly those members x of A for which the property $P(x)$ holds.**

Note here that we have many axioms, depending on the choice of A and of $P(x)$. These axioms come close to our feeling, expressed when first mentioning the concept of a set above, that we can form a set out of any property that sets might have. But we have restricted this general principle just to the elements of a given set A. The reason for this restriction is Russell's paradox. If we allow unrestricted comprehension, then we could form the set B whose elements are all the sets x such that x is not a member of x. But if $B \in B$, then $B \notin B$, and if $B \notin B$, then $B \in B$, a contradiction. This famous paradox is one of the main reasons that an axiomatic development of set theory is necessary. The paradox has many colloqual forms that might help in remembering it. For example, there is the barber paradox: a barber in some town who shaves only those men who do not shave themselves; does he shave himself? The liar paradox: a Greek who says "I am lying". Is she? We will run into useful mathematical forms of these paradoxes in the Cantor diagonal argument later.

Note also the notion of "property", which we have not made precise. It is a notion that can be made completely precise in a formal development of the logical basis of set theory. For our purposes, we take any "reasonable" statement.

The comprehension axioms and the extensionality axiom justify the most common way of defining sets. If A is a given set and $P(x)$ is a property of sets, then we define

$$\{x \in A : P(x)\} = \text{the unique set } B \text{ whose elements are exactly those} \\ \text{members } x \text{ of } A \text{ such that } P(x) \text{ holds.}$$

Informally we will also use the notation $\{x : P(x)\}$ to mean something like "$\{x \in A : P(x)\}$, where A is a set which can be shown to exist by our axioms". That is, $P(x)$ implies that x is in such a suitable set A. Then by the extensionality axiom, $\{x : P(x)\}$ does not depend on the particular set A of this sort. We can say that a set x belongs to $\{x : P(x)\}$ iff $P(x)$ holds. Notice that this is not a back-door way of allowing $\{x : x \notin x\}$, since the associated set A cannot be proved to exist from our axioms (we hope!). This notation will be explained more thoroughly the first few times we use it.

Now we can introduce several concepts of elementary set theory:

$$\begin{aligned}
\emptyset &= \{x : x \neq x\}; \\
a \cap b &= \{x \in a : x \in b\}; \\
a \backslash b &= \{x \in a : x \notin b\}; \\
\bigcap A &= \begin{cases} \emptyset & \text{if } A = \emptyset, \\ \{x : \forall a \in A (x \in a)\} & \text{if } A \neq \emptyset; \end{cases} \\
a \cap b &= \bigcap \{a, b\}.
\end{aligned}$$

Here the definitions of $a \cap b$ and $a \backslash b$ are clear applications of the above notation. For $\emptyset$, take any set A (which exists by Axiom 0), and let $B = \{x \in A : x \neq x\}$. Then B has no elements, and by the extensionality axiom it is unique. This is an instance of the extended notation $\{x : P(x)\}$ mentioned above. For $\bigcap$, note that if $A \neq \emptyset$, then we can choose $b \in A$ and let $B = \{x \in b : x \in a \text{ for all } a \in A\}$; then B is as indicated in the second case of the definition. Almost always, when we use the notation $\bigcap A$ we want A to be nonempty, and usually prove that it is.

Summarizing these new definitions, we have:

$\emptyset$ is the set with no members. (The *empty set.*)
$a \cap b$ is the set whose members are those sets which are in a and b. (The *intersection* of a and b.)
$a \backslash b$ is the set whose members are those sets which are in a but not in b. (The *complement of* b *in* a.)

$\bigcap A$ is the *intersection* of A. Note that for $A \neq \emptyset$, this consists of all elements which are in every member of A.
$a \cap b$ is the *intersection* of a and b.

We also call sets A, B *disjoint* iff $A \cap B = \emptyset$.

In the following proposition we give some simple properties of these notions. More are found in the exercises. All of this material should be, or become, familiar to the reader, and will be used often without reference to this part of the notes.

Proposition 1.1. *(i)* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
(ii) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
(iii) $A \backslash (B \cup C) = (A \backslash B) \cap (A \backslash C)$.
(iv) $A \backslash (B \cap C) = (A \backslash B) \cup (A \backslash C)$.
(v) If $B \subseteq A$, *then* $C = A \backslash B$ *iff* $C \cap B = \emptyset$ *and* $C \cup B = A$.
(vi) If $A \subseteq B$, *then* $\bigcup A \subseteq \bigcup B$.
(vii) If $\emptyset \neq A \subseteq B$, *then* $\bigcap B \subseteq \bigcap A$.
(viii) $\bigcup A \subseteq B$ *iff for all* $a \in A$, $a \subseteq B$.
(ix) If $B \neq \emptyset$, *then* $A \subseteq \bigcap B$ *iff for all* $b \in B$, $A \subseteq b$.

Proof. We prove only (ii) and (vii), and leave the rest as an exercise. Suppose that $x \in A \cup (B \cap C)$. Then there are two possibilities.

Case 1. $x \in A$. Then $x \in A \cup B$ and $x \in A \cup C$, so $x \in (A \cup B) \cap (A \cup C)$.

Case 2. $x \in B \cap C$. Then $x \in B$ and $x \in C$, so $x \in A \cup B$ and $x \in A \cup C$. So $x \in (A \cup B) \cap (A \cup C)$.

This proves $\subseteq$ in (ii). For $\supseteq$, suppose that $x \in (A \cup B) \cap (A \cup C)$. Thus $x \in A \cup B$ and $x \in A \cup C$. Suppose that $x \notin A$. Then $x \in B$, since $x \in A \cup B$. Similarly, $x \in C$. So $x \in B \cap C$. Thus we have shown that $x \notin A$ implies that $x \in B \cap C$. By logic, $x \in A$ or $x \in (B \cap C)$, so $x \in A \cup (B \cap C)$.

For (vii), suppose that $\emptyset \neq A \subseteq B$ and $x \in \bigcap B$. To show that $x \in \bigcap A$, take any $a \in A$. Then $a \in B$, and so $x \in a$. Hence $x \in \bigcap A$. $\square$

A useful heuristic to work out some of the elementary operations on sets are the Venn diagrams, which are given here for several of our basic notions; see the next page.

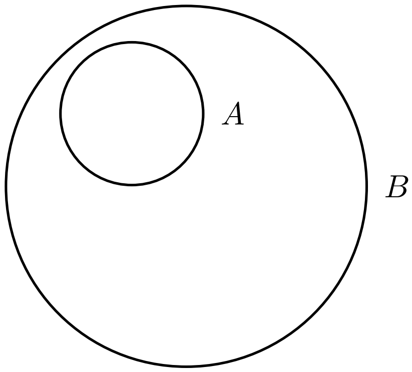
Exercises, chapter 1

1. Prove the following:
   (a) If $\{x, y\} = \{x, z\}$, then $y = z$.
   (b) If $u \neq v$ and $\{x, u\} = \{y, v\}$, then $x = v$ and $y = u$.

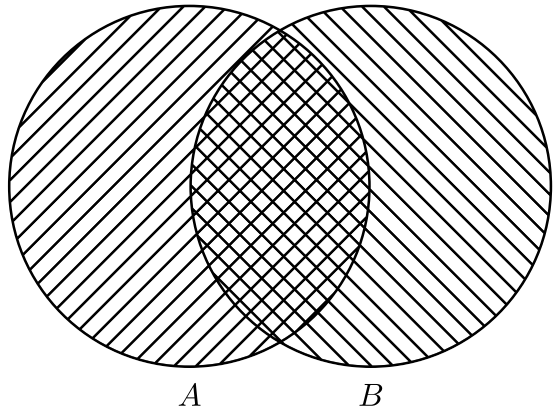
2. Finish the proof of Proposition 1.1.

3. For any sets a, b let the *symmetric difference* of a and b be defined as $a \triangle b = (a \backslash b) \cup (b \backslash a)$. Show that for any x, $x \in a \triangle (b \triangle c)$ iff x is in exactly one or exactly three of a, b, c. **Note that "show" means "prove".**)

4. Infer from exercise 3 that for any sets a, b, c one has $a \triangle (b \triangle c) = (a \triangle b) \triangle c$.



$A \subseteq B$



In the figure on the right, $A \cup B$ is the entire lined region. $A \cap B$ is the cross-hatched region. $A \backslash B$ is the single lined region on the left.

---

5. For this exercise we assume a knowledge of natural numbers and induction on the set of natural numbers. Suppose that $a_0, \ldots, a_{m-1}$ is a system of sets. Prove that a set x is in $a_0 \triangle a_1 \triangle \cdots \triangle a_{m-1}$ iff $\{i < m : x \in a_i\}$ has an odd number of elements.

6. Prove that $a = b$ iff $a \triangle b = \emptyset$.

7. Let x and y be sets. Assume that for every set a, if $x \in a$ then $y \in a$. Prove that $x = y$.

8. Suppose that a, b, c are sets, with $a \subseteq c$ and $b \subseteq c$. Show that the following conditions are equivalent:
   (a) $a \subseteq b$.
   (b) $c \backslash b \subseteq c \backslash a$.
   (c) $a \backslash b = \emptyset$.
   (d) $(c \backslash a) \cup b = c$.

9. Suppose that a, b, v are sets, with $a \cup b \subseteq v$. For brevity let $x' = v \backslash x$ for every $x \subseteq v$. Prove that $[(a' \cup b)' \cup b]' \cup (a \cup b')' \cup a = v$.

10. Under the same assumptions as in problem 9, but also with $c, d \subseteq v$, show that

$$([(b' \cup c)' \cup a' \cup c] \cap d') \cup (a \cap b') \cup d = v.$$

11. Show that there do not exist sets a, b, c such that $a \cap b \neq \emptyset$, $a \cap c = \emptyset$, and $(a \cap b) \backslash c = \emptyset$.

12. Show that for any sets a, b, c the following conditions are equivalent:
   (a) $c \subseteq a$.
   (b) $(a \cap b) \cup c = a \cap (b \cup c)$.

13. Given sets $a \subseteq b \subseteq c$, find a set x such that $b \cap x = a$ and $b \cup x = c$.