APPENDIX A: "CHAPTER 0". BASIC LOGIC AND SET THEORY

WAYNE AITKEN AND LINDA HOLT

Summer 2019 Edition

This appendix reviews basic logic and some of the other set theoretical background needed for the course.

1. The logical basis

The purpose of this course is to systematically develop the number systems commonly used in mathematics. A second purpose is to illustrate the axiomatic method through the development of these number systems. In the spirit of the axiomatic method, our development of the number systems will be rigorous and self-contained: we will give careful proofs for our results. There are, however, two exceptions where we will allow results without proof:

(1) *Axioms.* These are fundamental statements that are accepted without the need for formal justification. Sometimes they are presented as "self-evident", but technically they do not need to be obvious. They are, however, accepted as true for the purpose of proving further results.

In this course the only axioms are the Dedekind-Peano axioms and the iteration axiom. In an optional section near the end of Chapter 1 the iteration axiom will be shown to be a consequence of the other axioms, so the only axioms that are necessary for this course are the Dedekind-Peano axioms.[1] In more advanced mathematics, the axiom of choice, and certain advanced set theoretic axioms are also sometimes needed.

(2) *Principles of logic and elementary set theory.* From the axioms, we will derive other results using logic. So we will take as given the knowledge of classical deductive logic. This logic can be used freely

[1]These axioms, coupled with some basic set theory, suffice for a large part of mathematics. Even geometry can be developed from these axioms. For example, once you have developed the real numbers $\mathbb{R}$, you can define the plane to be $\mathbb{R}^2$ and three-dimensional space to be $\mathbb{R}^3$. In this approach you develop all the theorems of Euclidean geometry using the coordinate point of view and no new geometric axioms are needed. This is in contrast to Euclid's original approach, updated by Hilbert, which developes geometry using geometric axioms that do not rely on the real numbers.

> to derive new results. For example, we assume the basic principles related to connectives ($\wedge$, $\vee$, $\implies$, $\neg$, $\iff$) quantifiers ($\forall$, $\exists$, $\exists!$), and equality ($=$). The principles of classical first-order logic will be reviewed below from the point of view of Gentzen-style natural deduction.
>
> We will regard elementary set theory as part of our logical background and toolkit. These includes concepts, rules, and facts that are in common use in modern mathematics. Included under the heading of set theory are principles concerning ordered pairs, functions, and relations as well as sets (because ordered pairs, functions, and relations can be modeled as certain types of sets). One purpose of this chapter is to outline these core principles of set theory. These principles can be developed axiomatically from a small set of axioms, but we will not do so here. We simply take them as given.[2]

Aside from the axioms, and the basic facts of logic and set theory, every statement we wish to establish or use must be proved. Even something as simple as the commutative law of addition, or even the equation $1 + 1 = 2$, will be proved.

Likewise, every *concept* not occurring in the axioms, logic, or elementary set theory must be defined before it can be used. Such a definition must use only set theoretical and logical concepts as well as previously established concepts. For example, we will define addition and multiplication using the concept of function from basic set theory. Similarly, will provide definitions for all the number systems except the natural numbers using various set-theoretical ideas applied to previously established number systems. The set of natural numbers is an exception; it will not be defined. Since the set $\mathbb{N}$ is part of the axioms, it does not actually need to be defined. In general, terms used in the axioms do not need to be defined, and such undefined terms are called *primitive terms*.

2. Proofs

We can view a proof as sequence of assertions, called "steps", each of which can be justified by appealing to previously established results, rules, previous steps of the current proof, and assumptions. The final step is the result you are trying to prove, or something that immediately yields the result.

[2]The best known axiomatic development of set theory uses the Zermelo-Fraenkel axioms including the axiom of choice. This is a very powerful axiom system and is overkill for what we do here. The principles discussed in this chapter can be proved in a weaker axiom system akin to Zermelo's original system without the axioms of infinity, choice, replacement, or foundation. The axiom of infinity is not needed in the background set theory since the existence of infinite sets is a consequence of the Dedekind-Peano axioms introduced in Chapter 1.

A typical step is justified by two things (1) one or more established statements or assumptions that support the current claim, and (2) a rule of logic, set theory, or theorem of mathematics that connects these statements to the current claim. The established statements (1) can included previously proved results, prior definitions, axioms and principles taken as given, hypotheses from the statement of the theorem currently being proved, previous steps from the current proof (that are valid in the current context), and local assumptions made to specify the current context. Some of the rules of inference used for (2) will be discussed later in this chapter.

In practice, (1) the supporting facts, and (2) the rule used to justify a step are not always specifically mentioned if they are obvious from context. For example, expressions such as 'thus' or 'from this it follows' are used to indicate that the previous step or series of steps is being used as supporting facts. Often the rule (2) is clear from the claim itself. However, beginners should err on the side of supplying more details than is necessary rather than too few details. When every detail is not written down, it only be because supplying the missing details is easy to both the author of the proof and the careful reader.

There is another way to justify a step. A step can be justified by including a whole subproof for the step. For example, a step of the form $\neg P$ is often justified by including a subproof with the extra assumption P that ends with a contradiction. A step of the form $P \implies Q$ can be justified by including a subproof that starts with assumption P and ends with Q. A proof of the statement $\forall x \in A, P(x)$ can be proved by a subproof that starts with the assumption of $a \in A$, where a is an arbitrary but fixed element of A, and ends with $P(a)$. Another example is proof by cases: one can justify a step with a proof by cases by appealling to multiple subproofs, each involving a separate case. Subproofs can themselves have subproofs. (Warning: statements established in a subproof cannot be regarded as valid outside the subproof since they are typically proved in the context of additional assumptions that are not in force outside the subproof).

The use of subproofs in a proof is the main thing that sets proofs in real mathematics apart from the simple two column proofs of traditional high school geometry courses. Care must be used in writing proofs to signal to the reader where the subproof ends and the main proof resumes. In other words, the reader needs to be alerted to any context change. There are several styles used to present a subproof. A subproof can be put before or after the step it justified. One might write 'Claim: P' followed by a subproof of P. A subproof can be removed from the main body of the proof and be proved separately as the proof of a *lemma*.[3] Such lemmas can be put before or after the main proof. If a lemma is put after a proof, care should be taken

[3]A *lemma* is a type of theorem that is not necessarily of independent interest, but is useful for establishing another result, or other results. In this class, I will use the term *claim* for a statement whose proof is embedded as a subproof in a larger proof, and the term *lemma* for a statement whose proof is separated from a larger proof.

to make sure that the lemma is independent of the proof it is supposed to support.

3. Formal proofs and the axiomatic method

As discussed above, each step in a proof should be justified (sometimes with a short justification, sometimes with a long subproof). In an *informal proof* the justification is fairly open-ended. It can involve any fact or rule that is mutually accepted by the writer and intended readers. It can involve facts and rules learned in prior math courses, or, in geometry or topology for instance, facts that are obvious from one's intuition. Conclusions that the intended reader can justify without too much work on their own are often written without full justification. Many informal proofs are really proof outlines.

In a *formal proof*, on the other hand, only facts and rules that are explicitly established can be used. Appeals to prior math courses, intuition, or details for the reader to work out are not allowed. Formal proofs are particularly suited to illustrating the axiomatic method and so will be the main type of proof in this course, but from time to time informal proofs and arguments will be allowed. You can use informal proofs in your scratch work to help you develop ideas and work out examples, and in exercises that are clearly labeled as "informal". Writing a formal proof should only be done after you have a strong understanding of the statement and how it can be justified.

Tom Hales explains the distinction between informal proof and formal proof:

> Traditional mathematical proofs are written in a way to make them easily understood by mathematicians. Routine logical steps are omitted. ... Proofs, especially in topology and geometry, rely on intuitive arguments ...
>
> A formal proof is a proof in which every logical inference has been checked all the way back to the fundamental axioms of mathematics. All the intermediate logical steps are supplied, without exception. No appeal is made to intuition, even if the translation from intuition to logic is routine.[4]

Some go further and require that formal proofs be presented in a purely symbolic formal language, or insist that they be written in a way that a suitable computer proof-checking program could check each step in a mechanical manner. We will not go that far, but will adhere to fairly strict standards, especially in the early part of the course.

The *axiomatic method* is the technique of carefully developing a body of results from a small set of axioms. The development of an axiomatic theory ideally should be rigorous and self-contained. The historic inspiration for the axiomatic method was Euclid's *Elements of Geometry*. In this course we

---

[4]Notices of the AMS, December 2008, page 1371.

will illustrate the axiomatic method in the development of the basic number systems.

There is a major psychological difficulty in using the axiomatic method: one starts by proving facts that are already known or obvious. In our development of the number systems, we need to pretend ignorance of anything about the numbers except for what has been established in this course. This is hard to do since facts about number systems have been ingrained into our minds from such an early age.

When developing a formal proof of a results it helps to adopt a hyper-skeptical attitude. Do not accept a step until you can see the justification for the step. When constructing a proof, you want to be both creative and critical (in the good sense) until you are completely satisfied that you have a tight, rigorous proof. *One of the best skills you can develop is to know when you do and do not have a valid proof of a result.* Ideally you should be able to assess the quality of your work independent of an external reviewer.

In addition to learning to be a skeptic, in order to succeed in this course, you need to develop a strong attention to detail. Cultivate a habit of careful, slow reading. It is all right, and often advisable, to read a section quickly to get the main ideas, as long as you follow it up with a second and third careful reading. If you do so, you will develop a thorough and lasting understanding of the material, and you will find it much easier to correctly complete the exercises.

4. Basic rules of logic

As mentioned above, a step in a proof is typically justified by appealing to a rule of inference applied to previously established statements. Many of the rules of inference come from logic, some come from set theory, and some rules of inference will be based on theorems proved in the course. We now present some of the logical rules of inference and logical identities that are commonly used in proofs. The reader is assumed to be already familiar with most of these, and these are stated mainly for reference. Many of these rules are taken from Genzen's natural deduction approach to proof where, for each logical operator, rules will be given for establishing a statement of a certain form (often called 'introduction rules') and other rules will be given for using a statement of that form (often called 'elimination rules') to prove other statements.

4.1. **Statements of the form $P \wedge Q$.** ("Conjunctions"). Statements of the form "P and Q" (written symbolically as $P \wedge Q$) are very well-behaved. To establish $P \wedge Q$ one can proceed by first establishing P then establishing Q. In other words, you can justify $P \wedge Q$ by citing the earlier result P and the earlier result Q. This rule is represented schematically as follows:

$$\frac{\begin{array}{c}P\\Q\end{array}}{P \wedge Q} \quad (\wedge \text{ introduction rule})$$

You can use a prior result of the form $P \wedge Q$ to justify P or justify Q:

$$\frac{P \wedge Q}{P} \qquad \frac{P \wedge Q}{Q} \qquad (\wedge \text{ elimination rules})$$

These rules extend in the obvious way to conjuncts of three or more statements:

$$P_1 \wedge P_2 \wedge P_3, \quad \text{et cetera.}$$

4.2. **Statements of the form $P \vee Q$.** (“Disjunction” or “inclusive or”). The simplest way to establish “P or Q” (symbolically $P \vee Q$) is to first establish P or, alternatively, first establish Q, This is not the only way of doing so, but it is conceptually the simplest. These inference rules are written schematically as follows:

$$\frac{P}{P \vee Q} \qquad \frac{Q}{P \vee Q} \qquad (\vee \text{ introduction rules})$$

In practice when you want to prove $P \vee Q$ it might not be possible to prove P or to prove Q directly. But there are other techniques. Another strategy is to assume one of the two is false, and deduce the other is true. For example, suppose you *assume* P is false, and you deduce Q from this assumption (in a subproof). This does not prove Q by itself, but it does prove $P \vee Q$. The reason why is that there are two cases: either P is true, and you are done, or P is false, in which case you showed that Q is true. In either case one of the two is true. This technique is indicated schematically as follows:

$$\frac{\neg P \Rightarrow Q}{P \vee Q} \qquad \frac{\neg Q \Rightarrow P}{P \vee Q}$$

An established statement of the form $P \vee Q$ can be used to justify a later result R via proof by cases. In such a proof you (1) prove R in a subproof (called a “case”) where P is assumed to be true, and (2) prove R in a subproof where Q is assumed. Using $P \vee Q$ together with the two subproofs, you can then conclude R. In other words, from $P \vee Q$ and $P \Rightarrow R$ and $Q \Rightarrow R$, you can conclude R:

$$\frac{\begin{array}{c} P \vee Q \\ P \Rightarrow R \\ Q \Rightarrow R \end{array}}{R} \qquad (\vee \text{ elimination “proof by cases”})$$

These rules extend in the obvious way to disjuncts of three or more statements:

$$P_1 \vee P_2 \vee P_3, \quad \text{et cetera.}$$

4.3. **Statements of the form $P \Rightarrow Q$.** ("Conditionals"). A common way to establish a claim of the form "if P then Q" (symbolically $P \Rightarrow Q$) is to supply a subproof. One assumes P and derives Q in the subproof. From the existence of this subproof one is entitled to assert $P \Rightarrow Q$. This is not the only way to prove $P \Rightarrow Q$. There are other rules such as the transitive rule for $\Rightarrow$ that we will discuss below, but it is the most basic way.

You can later use a result of the form $P \Rightarrow Q$ by applying it to an established statement P to derive a statement Q. This rule is called *modus ponens*, and can be schematically indicated as follows.

$$\frac{\begin{array}{c} P \Rightarrow Q \\ P \end{array}}{Q} \quad \text{(modus ponens)}$$

(This rule is also called $\Rightarrow$ elimination).

4.4. **Statements of the form $\neg P$.** ("Negations"). The negation of P can be justified by showing $P \Rightarrow \mathcal{F}$ (typically with a subproof) where $\mathcal{F}$ is any contradiction of a previously established result, or any obviously false statement (for example the statement $1 \neq 1$).[5] This rule is represented as follows:

$$\frac{P \implies \mathcal{F}}{\neg P} \quad (\neg \text{ introduction rule})$$

Once you have $\neg P$, you can use it to eliminate a case. For example, if you have $P \vee Q$ and you also have $\neg P$, you can conclude Q.

$$\frac{\begin{array}{c} P \vee Q \\ \neg P \end{array}}{Q} \quad \text{(elimination of case)}$$

In classical logic, we automatically accept any statement of the form

$$P \vee \neg P,$$

but we automatically reject

$$P \wedge \neg P.$$

(Such a conjunction is considered obviously false, and is sometimes called a "contradiction").

4.5. **Statements of the form $P \iff Q$.** ("Biconditionals"). Statements of the form "P if and only if Q" can be established with the following rule:

$$\frac{\begin{array}{c} P \implies Q \\ Q \implies P \end{array}}{P \iff Q} \quad (\iff \text{ introduction rule})$$

[5]For example, if Q has been established in the current context, then $\mathcal{F}$ can be $\neg Q$, or if $\neg Q$ has been established then $\mathcal{F}$ can be Q.

These statements of the form $P \Leftrightarrow Q$ can be used to justify other statements with the following inference rules:

$$\frac{P \iff Q}{P \implies Q} \qquad \frac{P \iff Q}{Q \implies P} \qquad (\iff \text{ elimination rules})$$

The connective $\Leftrightarrow$ satisfies the reflexive, symmetric, and transitive laws:

$$P \iff P \qquad \frac{P \iff Q}{Q \iff P} \qquad \frac{\begin{array}{c} P \iff Q \\ Q \iff R \end{array}}{P \iff R}$$

It also satisfies a *substitution law*: if $P \iff Q$ then you can replace any occurrence of P with Q in a larger compound statement and the result with be equivalent. This is sometimes written as follows: assume $\varphi(P)$ is a compound statement in which P occurs and assume $\varphi(Q)$ is the same statement but where one or more occurrences of P have been replaced by Q:

$$\frac{P \iff Q}{\varphi(P) \iff \varphi(Q)} \qquad (\iff \text{ substitution rule})$$

4.6. **Contradictions and cases.** From a contradiction $(Q \land \neg Q)$ or any other result that is known to be false (written $\mathcal{F}$) one can derive anything you want. This is written as follows:

$$\frac{\mathcal{F}}{P} \quad \text{(contradiction rule)}$$

This is a rather strange rule at first glance, but it is useful in proofs by cases. Suppose you want to justify a step R and you decide to prove it by cases based on a previous result $P_1 \lor \ldots \lor P_n$. In other words, you give a subproof for each case P_i. Your goal is to prove R in each of these cases. Some of these cases might turn out to be impossible. For example, P_i might imply an absurdity $\mathcal{F}$. The above rule will then allow you to conclude R in that case. In other words, if a case leads to a contradiction, you can automatically move on to the next case since everything is true in a contradictory case.

4.7. **Other useful rules.** There are several other rules that are useful to know. (These rules can be derived from the rules we have already considered).

$$\frac{\begin{array}{c} P \implies Q \\ \neg Q \end{array}}{\neg P} \qquad \frac{\begin{array}{c} P \implies Q \\ Q \implies R \end{array}}{P \implies R} \qquad \frac{\begin{array}{c} P \iff Q \\ P \end{array}}{Q} \qquad \frac{\begin{array}{c} P \iff Q \\ Q \end{array}}{P} \qquad \frac{\begin{array}{c} P \iff Q \\ \neg Q \end{array}}{\neg P}$$

$$\frac{P \implies Q}{P \land R \implies Q \land R} \qquad \frac{P \implies Q}{P \lor R \implies Q \lor R} \qquad \frac{P}{Q \implies P}$$

4.8. **Useful identities.** The following are important logical identities including commutative and associative laws. Each line below represents a type of statement that can be accepted as automatically true.[6]

$$
\begin{array}{rcl}
P \wedge Q & \iff & Q \wedge P \\
(P \wedge Q) \wedge R & \iff & P \wedge (Q \wedge R) \\
P \vee Q & \iff & Q \vee P \\
(P \vee Q) \vee R & \iff & P \vee (Q \vee R) \\
P \wedge P & \iff & P \\
P \vee P & \iff & P \\
\neg\neg P & \iff & P \\
(P \implies Q) & \iff & \neg P \ \vee \ Q \\
(P \implies Q) & \iff & (\neg Q \implies \neg P) \\
P \vee Q & \iff & (\neg Q \implies P) \\
P \vee Q & \iff & (\neg P \implies Q)
\end{array}
$$

There are two distributive laws

$$
\begin{array}{rcl}
(P \vee Q) \wedge R & \iff & (P \wedge R) \vee (Q \wedge R) \\
(P \wedge Q) \vee R & \iff & (P \vee R) \wedge (Q \vee R)
\end{array}
$$

and two De Morgan laws.

$$
\begin{array}{rcl}
\neg(P \vee Q) & \iff & (\neg P) \ \wedge \ (\neg Q) \\
\neg(P \wedge Q) & \iff & (\neg P) \ \vee \ (\neg Q)
\end{array}
$$

5. Quantifiers

The above illustrates "propositional logic". Logic becomes more sophisticated and powerful when we introduce *quantifiers*. This results in "predicate logic" or "quantificational logic".

We begin with the universal quantifier $\forall$. The most direct way to justify the assertion $\forall x,\ Px$ is through a subproof where a is taken to be an arbitrary but fixed object and where Pa is proved. Here Px is a predicate with variable x. (Here a should be a new term that is not being used for any other purpose in the current context).

If you already have $\forall x,\ Px$ you can use it to justify special cases of the predicate Px using the following rule which is valid for any desired a:

$$
\frac{\forall x,\ Px}{Pa} \qquad \forall \text{ elimination rule}
$$

The other type of quantifier is the existential quantifier $\exists$. The most direct way to justify the assertion $\exists x, Px$ is to appeal to a statement of the form Pa. Here a can be any term making the predicate Px true, it does not

[6]Such identities are sometimes called "tautologies" in logic textbooks.

have to be arbitrary in any sense. This introduction rule can be represented schematically as follows:

$$\frac{Pa}{\exists x, Px} \qquad \exists \text{ introduction rule}$$

If you have $\exists x, Px$ already, you can use it to define a new constant a representing a choice of object such that Pa is true.[7]

In addition to the two basic quantifiers $\forall$ and $\exists$, we have a third quantifier $\exists!$ ("there exists a unique") which can be defined in terms of the other two. Here are two (equivalent) definitions:

$$\begin{aligned} \exists!\, x,\ Px \quad &\overset{\text{def}}{\iff} \quad \exists x \Big(Px \ \wedge\ \forall y\, (\, Py\ \Rightarrow y = x)\Big) \\ &\overset{\text{def}}{\iff} \quad \Big(\, \exists x,\ Px \Big) \wedge \Big(\forall y\, \forall z\, \big(Py \wedge Pz\ \Rightarrow y = z\big)\Big) \end{aligned}$$

Note: '!' stands for "unique" here, but it only means "unique" when it is used after '$\exists$'. Also, be careful of the term "unique": an object cannot be unique by itself. Uniqueness only makes sense in the context of a predicate Px that such an object satisifies.

Here are some identities involving quantifiers:

$$\begin{aligned} \neg\big(\,\exists x,\ Px\,\big) \quad &\iff \quad \forall x,\ \neg Px \\ \neg\big(\forall x,\ Px\,\big) \quad &\iff \quad \exists x,\ \neg Px \\ \forall x\, \forall y, P(x,y) \quad &\iff \quad \forall y\, \forall x, P(x,y) \\ \exists x\, \exists y, P(x,y) \quad &\iff \quad \exists y\, \exists x, P(x,y) \\ \forall x\big(Px\ \wedge\ Qy\big) \quad &\iff \quad \big(\forall x,\ Px\big) \wedge \big(\forall x,\ Qx\big) \\ \exists x\ \big(Px\ \vee\ Qy\big) \quad &\iff \quad \big(\exists x,\ Px\big) \vee \big(\exists x,\ Qx\big) \end{aligned}$$

Warning: $\forall x\, \exists y, P(x,y)$ is not logically equivalent to $\exists y\, \forall x, P(x,y)$. The order of the quantifiers makes a big difference in meaning in this case.

6. Equality

Equality = satisfies the reflexive, symmetric, and transitive laws:

$$a = a \qquad \frac{a = b}{b = a} \qquad \frac{\begin{matrix} a = b \\ b = c \end{matrix}}{a = c}$$

The symmetry law is also written

$$a = b \iff b = a.$$

Equality also satisfies a *substitution law*: if $a = b$ then you can replace any occurrence of a with b in a larger compound term to form an equivalent

[7]Related to this is an elimination rule that allows you to justify a statement R as follows: if you have $\exists x, Px$ and if you know that $Pa \implies R$ for arbitrary a (for example if you have a subproof with assumption Pa that proves R where a is arbitrary), then you can conclude R. Here R is a statement that does not involve a.

term.[8] This is sometimes written as follows: assume $\tau(a)$ is a term in which a occurs and assume $\tau(b)$ is the same term but where one or more occurrences of a have been replaced by b:

$$\frac{a = b}{\tau(a) = \tau(b)} \qquad (= \text{substitution rule})$$

There is also a second substitution rule[9] for statements: assume $\varphi(a)$ is a statement in which a occurs and assume $\varphi(b)$ is the same statement but where one or more occurrences of a have been replaced by b:

$$\frac{a = b}{\varphi(a) \iff \varphi(b)} \qquad (= \text{substitution rule 2})$$

7. Elementary set theory

The basics concepts, rules, and facts of set theory will be used extensively in this course.

7.1. **Equality and inclusion.**

Two sets are equal if and only if they have the same elements:

$$A = B \iff \forall x\ (x \in A \iff x \in B)$$

The set A is a subset of B if and only if every element of A is in B:

$$A \subseteq B \iff \forall x\ (x \in A \implies x \in B)$$

Two sets are equal if and only if each is a subset of the other. This gives rise to the following rules:

$$\frac{\begin{array}{c} A \subseteq B \\ B \subseteq A \end{array}}{A = B} \qquad \frac{A = B}{A \subseteq B} \qquad \frac{A = B}{B \subseteq A}$$

We also have the following:

$$\frac{\begin{array}{c} A \subseteq B \\ x \in A \end{array}}{x \in B} \qquad \frac{\begin{array}{c} A \subseteq B \\ B \subseteq C \end{array}}{A \subseteq C} \qquad A \subseteq A$$

7.2. **The empty set.**

The empty set $\varnothing$ is the set with no elements:

$$\neg\big(\exists\, x,\ x \in \varnothing\big)$$

The empty set is a subset of all sets:

$$\varnothing \subseteq A$$

[8] A *term* is an expression denoting an object.

[9] Warning: in these substitution rules we assume that the bound variables are distinct from the variables occurring in a and b.

Here are rules to show a set A is empty or nonempty:

$$\frac{\neg(\exists\, x,\ x \in A)}{A = \varnothing} \qquad \frac{\exists\, x,\ x \in A}{A \neq \varnothing}$$

7.3. **Small sets.** Small sets can be denoted by denoting all the elements. For example, $\{0, 1, 4, 9, 16, 25, 36, 49, 64, 81\}$ denotes the square integers less than 100.

Here are equivalences related to small sets:

$$\begin{array}{rcl} x \in \{a\} & \Longleftrightarrow & x = a \\ x \in \{a, b\} & \Longleftrightarrow & (x = a) \vee (x = b) \\ x \in \{a, b, c\} & \Longleftrightarrow & (x = a) \vee (x = b) \vee (x = c) \\ & \text{etc.} & \end{array}$$

Here are some equalities:

$$\begin{array}{rcl} \{a, b\} & = & \{b, a\} \\ \{a, a\} & = & \{a\} \end{array}$$

7.4. **Intersections, unions, and differences.** These are governed by the following equivalences:

$$\begin{array}{rcl} x \in A \cap B & \Longleftrightarrow & (x \in A) \wedge (x \in B) \\ x \in A \cup B & \Longleftrightarrow & (x \in A) \vee (x \in B) \\ x \in A - B & \Longleftrightarrow & (x \in A) \wedge (x \notin B) \end{array}$$

They satisfy the following rules

$$\frac{\begin{array}{c} A \subseteq C \\ B \subseteq C \end{array}}{A \cup B \subseteq C} \qquad \frac{\begin{array}{c} C \subseteq A \\ C \subseteq B \end{array}}{C \subseteq A \cap B} \qquad \frac{\begin{array}{c} C \subseteq A \\ C \cap B = \emptyset \end{array}}{C \subseteq A - B}$$

They satisfy the following inclusions:

$$\begin{array}{rcl} A & \subseteq & A \cup B \\ B & \subseteq & A \cup B \\ A \cap B & \subseteq & A \\ A \cap B & \subseteq & B \\ A - B & \subseteq & A \end{array}$$

And they satisfy the following equalities:

$$\begin{aligned}
A \cap B &= B \cap A \\
A \cup B &= B \cup A \\
(A \cap B) \cap C &= A \cap (B \cap C) \\
(A \cup B) \cup C &= A \cup (B \cup C) \\
A \cap A &= A \\
A \cup A &= A \\
A \cap \varnothing &= \varnothing \\
A \cup \varnothing &= A \\
(A \cup B) \cap C &= (A \cap C) \cup (B \cap C) \\
(A \cap B) \cup C &= (A \cup C) \cap (B \cup C) \\
(A - B) \cup B &= A \cup B \\
(A - B) \cap B &= \varnothing
\end{aligned}$$

7.5. **Quantification over a set.** The quantifier $\forall\, x \in A$ is defined by the following

$$\forall\, x \in A,\ Px \ \overset{\text{def}}{\iff}\ \forall\, x\ \big(x \in A \implies Px)$$

The quantifier $\exists\, x \in A$ is defined by the following

$$\exists\, x \in A,\ Px \ \overset{\text{def}}{\iff}\ \exists\, x\ \big(x \in A \wedge Px)$$

The most direct way to justify the assertion $(\forall\ x \in A,\ Px)$ is through a subproof where a is taken to be an arbitrary but fixed element of A and where Pa is proved.

We have the following elimination rule:

$$\frac{\begin{array}{c}\forall\, x \in A,\ Px \\ a \in A\end{array}}{Pa}$$

Observe that we can use this type of quantifier to show the subset relation:

$$A \subseteq B \iff \forall\, x \in A, (x \in B)$$

To justify $(\exists\, x \in A,\ Px)$ we have the following introduction rule:

$$\frac{\begin{array}{c}Pa \\ a \in A\end{array}}{\exists\, x \in A,\ Px}$$

If you have $(\exists\ x \in A,\ Px)$ already, you can use it to define a new constant a representing a choice of element of A such that Pa.[10]

[10]Related to this is an elimination rule that allows you to justify a statement R as follows: if you have $\exists\, x \in A,\ Px$ and if you know that $Pa \implies R$ for arbitrary $a \in A$, then you can conclude R. Here R is a statement that does not involve a.

We have a third quantifier $\exists!\ x \in A$. Here are two (equivalent) definitions:[11]

$$\begin{aligned}\exists!\ x \in A,\ Px \quad &\overset{\text{def}}{\Longleftrightarrow} \quad \exists\, x \in A, \Big(Px\ \wedge\ \forall\, y \in A\ (\,Py\ \Rightarrow y = x)\Big) \\ &\overset{\text{def}}{\Longleftrightarrow} \quad \Big(\exists\, x \in A,\ Px\Big) \wedge \Big(\forall\, y, z \in A,\ \big(Py \wedge Pz\ \Rightarrow y = z\big)\Big)\end{aligned}$$

Here are some rules associated to these concepts:

$$\frac{\exists\, x \in A,\ Px \qquad A \subseteq B}{\exists\, x \in B,\ Px} \qquad \frac{\forall\, x \in B,\ Px \qquad A \subseteq B}{\forall\, x \in A,\ Px}$$

Here are some equivalences:

$$\begin{aligned}\neg\big(\,\exists\, x \in A,\ Px\,\big) \quad &\Longleftrightarrow \quad \forall\, x \in A,\ \ \neg Px \\ \neg\big(\,\forall\, x \in A,\ Px\,\big) \quad &\Longleftrightarrow \quad \exists\, x \in A,\ \ \neg Px \\ \forall\, x \in A,\ \ \forall\, y \in A,\ \ P(x,y) \quad &\Longleftrightarrow \quad \forall\, y \in A,\ \ \forall\, x \in A,\ \ P(x,y) \\ \exists\, x \in A,\ \ \exists\, y \in A,\ \ P(x,y) \quad &\Longleftrightarrow \quad \exists\, y \in A,\ \ \exists\, x \in A,\ \ P(x,y) \\ \forall\, x \in A, \big(Px\ \wedge\ Qy\big) \quad &\Longleftrightarrow \quad \big(\forall\, x \in A,\ Px\big) \wedge \big(\forall\, x \in A,\ Qx\big) \\ \exists\, x \in A, \big(Px\ \vee\ Qy\big) \quad &\Longleftrightarrow \quad \big(\exists\, x \in A,\ Px\big) \vee \big(\exists\, x \in A,\ Qx\big)\end{aligned}$$

Warning: as we see above, the negation of $\exists\, x \in A, Px$ is $\forall\, x \in A,\ \neg Px$ not $\forall\ x \notin A,\ \neg Px$. Remember that the negation should be a statement about elements of A, not about elements outside of A.

7.6. **General unions and intersections.** Let Z be a set of sets (for intersections we require that Z is nonempty). Then we have the following types of unions and intersections:

$$\bigcup Z\ =\ \bigcup_{X \in Z} X\ =\ \big\{u\ \mid\ \exists\, X \in Z,\ u \in X\big\}$$

$$\bigcap Z\ =\ \bigcap_{X \in Z} X\ =\ \big\{u\ \mid\ \forall\, X \in Z,\ u \in X\big\}$$

We have the following special cases:

$$\bigcup\{A\} = A, \quad \bigcup\{A, B\} = A\ \cup\ B, \quad \bigcup\{A, B, C\} = A\ \cup\ B \cup\ C, \quad \text{etc.}$$

$$\bigcap\{A\} = A \quad \bigcap\{A, B\} = A \cap\ B, \quad \bigcap\{A, B, C\} = A\ \cap\ B \cap\ C, \quad \text{etc.}$$

The general union and intersection are especially useful for cases where Z is an infinite set of sets.

They satisfy the following rules

$$\frac{\forall\, X \in Z,\ X \subseteq C}{\bigcup Z \subseteq C} \qquad \frac{\forall\, X \in Z,\ C \subseteq X}{C \subseteq \bigcap Z}$$

$$\frac{X \in Z}{X \subseteq \bigcup Z} \qquad \frac{X \in Z}{\bigcap Z \subseteq X}$$

[11]The notation $\forall y, x \in A$ is short for $\forall y \in A, \forall x \in A$.

8. Ordered pairs

An *unordered pair* is a set $\{a, b\}$. Here $\{a, b\} = \{b, a\}$. When we we want the order to be significant for equality, we use *ordered pairs*.[12]

We use (a, b) to denote the ordered pair with first coordinate a and second coordinate b. We have the following:

$$(a, b) = (c, d) \iff (a = c) \wedge (b = d)$$

The *Cartesian product* $A \times B$ of sets A and B is the set of ordered pairs with first coordinate in A and second coordinate in B:

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

We sometimes write $A \times A$ as A^2.

9. Functions

Modern set theory does not concern itself only with basic properties and operations on sets, but it also concerns itself with functions and their properties. In fact, functions are considered to be a special type of set: a type of set of ordered pairs. Each function has a *domain* and *codomain*.

If A and B are sets, then we write $f : A \to B$ to indicate that f is a function with domain A and codomain B. Such a function f maps each element $a \in A$ to an element $fa \in B$. We sometimes write fa as $f(a)$, especially when grouping needs to be indicated. Schematically we have the following:

$$\frac{\begin{array}{c} f : A \to B \\ a \in A \end{array}}{fa \in B}$$

We call fa the *value*, or the *image of* a. (Warning: there may be elements of B that are not of the form fa. However, if f is surjective then every element of B is indeed of the form fa.)

When are two functions equal? If $f\colon A \to B$ and $g\colon A \to B$ are functions with matching domain and codomain then

$$f = g \iff \forall\, x \in A,\ (fx = gx).$$

We will use a notation that distinguishes between two kinds of arrows, written as $\to$ and $\mapsto$. We use $\to$ as above to indicate domain and codomain. We use $\mapsto$ to illustrate a definition or description of a function. More specifically, if we want to define $f : A \to B$ by a rule, we sometimes indicate the rule by writing an expression of the form $x \mapsto \varphi(x)$. Here x stands for an arbitrary element of the domain, and $\varphi(x)$ is an expression for the value of the function.

[12] In some set theory books, ordered pairs are defined in terms of unordered pairs. Sometimes (a, b) is defined as $\{\{a\}, \{a, b\}\}$, but sometimes it is defined differently. You do not need to worry about how ordered pairs are defined, but instead concentrate on the key identity $(a, b) = (c, d) \iff (a = c) \wedge (b = d)$.

9.1. **Composition.** Suppose $f\colon A \to B$ and $g\colon B \to C$ are functions such that the codomain of f is equal to the domain of g. We define the composition $g \circ f : A \to C$ to be the function given by the rule

$$x \mapsto g(fx).$$

In other words, if $x \in A$ then $(g \circ f)(x) = g(f(x))$. Schematically:

$$\frac{\begin{array}{c} f : A \to B \\ g : B \to C \\ a \in A \end{array}}{(g \circ f)(a) = g(f(a))} \qquad \frac{\begin{array}{c} f : A \to B \\ g : B \to C \end{array}}{(g \circ f) : A \to C}$$

Composition satisfies the associative law:

$$\frac{\begin{array}{c} f : A \to B \\ g : B \to C \\ h : C \to D \end{array}}{h \circ (g \circ f) = (h \circ g) \circ f}$$

9.2. **Images and inverse images of sets.** Suppose S is a subset of the domain of $f : A \to B$. Then

$$f[S] \stackrel{\text{def}}{=} \{fx \mid x \in S\}$$

so

$$y \in f[S] \iff \exists\, x \in S,\ y = fx$$

and

$$f[S] \subseteq B$$

The set $f[S]$ is called the *image* of $A \subseteq S$. Warning: the term *image* is ambiguous: it can refer to elements $fa \in B$ or subsets $f[S] \subseteq B$. It is usually clear what is meant based on context, but if there is a chance of confusion, we use the phrase "image of the *set* A" to indicate that we mean a subset of the codomain B and not an element of B.

The *image of the function* $f : A \to B$ is the image of the whole domain A. Thus the image of f is $f[A]$.

Suppose S is a subset of the codomain of $f : A \to B$. Then

$$f^{-1}[S] \stackrel{\text{def}}{=} \{x \in A \mid fx \in S\}.$$

This set, called the *inverse image* or *preimage*, is defined even if the inverse function f^{-1} is not defined. We have

$$x \in f^{-1}[S] \iff fx \in S$$

and

$$f^{-1}[S] \subseteq A.$$

Also

$$f^{-1}[B] = A.$$

9.3. **Identity functions.** If A is a set, then the identity function

$$id_A : A \to A$$

is the function defined by the rule $x \mapsto x$. Thus we have the simple law

$$\frac{a \in A}{id_A(a) = a}$$

We also have composition laws:

$$\frac{f : A \to B}{f \circ id_A = f} \qquad \frac{g : B \to A}{id_A \circ g = g}$$

9.4. **Injective and surjective functions.** There are two types of functions that arise often in mathematics: injective functions (also called one-to-one functions) and surjective functions (also called onto functions). These are important in this course, so *students are advised to review these concepts until they have a good understanding of these concepts.*

An *injective function* or *injection* $f : A \to B$ is a function that sends distinct elements of the domain to distinct elements of the codomain. More formally:

$$f : A \to B \text{ is injective} \iff \forall x, y \in A,\ (x \neq y \implies fx \neq fy).$$

This is more commonly expressed in the following equivalent form:

$$f : A \to B \text{ is injective} \iff \forall x, y \in A,\ (fx = fy \implies x = y).$$

So, to prove a function is injective you will often check that the equality $fx = fy$ implies the equality $x = y$ for all x, y in the domain A, and then use the following rule:

$$\frac{\begin{array}{c} f : A \to B \\ \forall x, y \in A,\ (fx = fy \implies x = y) \end{array}}{f : A \to B \text{ is injective}}$$

An *surjective function* or *surjection* $f : A \to B$ is a function whose image is equal to the codomain:

$$f : A \to B \text{ surjective} \iff f[A] = B.$$

In other words,

$$f : A \to B \text{ surjective} \iff \forall\, b \in B,\ \ \exists\, a \in A,\ \ fa = b.$$

So one common way to check that a function is surjective is to take an arbitrary element b in the codomain B, and show that you can find an element a in the domain A that maps to B.

Another way to show a function is injective or surjective is to find an inverse function. (It is enough to have a left inverse for injective, and a right inverse for surjective. We will discuss inverses in more detail below):

$$\frac{f : A \to B \quad g : B \to A \quad g \circ f = id_A}{f \text{ injective}} \qquad \frac{f : A \to B \quad g : B \to A \quad g \circ f = id_A}{g \text{ surjective}}$$

Composition of functions behaves well:

$$\frac{f : A \to B \text{ injective} \quad g : B \to C \text{ injective}}{g \circ f : A \to C \text{ injective}} \qquad \frac{f : A \to B \text{ surjective} \quad g : B \to C \text{ surjective}}{g \circ f : A \to C \text{ surjective}}$$

9.5. **Bijective functions.** A *bijective function* or *bijection* $f : A \to B$ is a function that is both injective and surjective. We have

$$f : A \to B \text{ is bijective} \iff \forall\, b \in B,\ \exists!\, a \in A,\ f a = b$$

and

Identity maps are bijections.

We also have the following laws:

$$\frac{f : A \to B \quad g : B \to A \quad \forall\, a \in A.\ g(f a) = a \quad \forall\, b \in B.\ f(g\, b) = b}{f \text{ and } g \text{ bijective}} \qquad \frac{f : A \to B \quad g : B \to A \quad g \circ f = id_A \quad f \circ g = id_B}{f \text{ and } g \text{ bijective}} \qquad \frac{f : A \to B \text{ bijective} \quad g : B \to C \text{ bijective}}{g \circ f : A \to C \text{ bijective}}$$

9.6. **Inverse functions.** If $f : A \to B$, then an inverse to f is a function $f^{-1} : B \to A$ such that

$$f^{-1} \circ f = id_A \qquad \text{and} \qquad f \circ f^{-1} = id_B.$$

If an inverse exists, it is unique, but not every function has an inverse. In fact

$$f : A \to B \text{ bijective} \iff f \text{ has an inverse.}$$

This gives rise to the following:

$$\frac{f : A \to B \text{ bijective}}{f^{-1} \circ f = id_A} \qquad \frac{f : A \to B \text{ bijective}}{f \circ f^{-1} = id_B}$$

$$\frac{f : A \to B \text{ bijective}}{\forall\, a \in A.\ f^{-1}(f\, a) = a} \qquad \frac{f : A \to B \text{ bijective}}{\forall\, b \in B.\ f(f^{-1}\, b) = b}$$

We also have the following:

$$\frac{f : A \to B \quad g : B \to A \quad \forall\, a \in A.\ g(f(a)) = a \quad \forall\, b \in B.\ f(g(b)) = b}{f = g^{-1} \text{ and } g = f^{-1}} \qquad \frac{f : A \to B \quad g : B \to A \quad g \circ f = id_A \quad f \circ g = id_B}{f = g^{-1} \text{ and } g = f^{-1}}$$

Finally, we have the following:

$$\frac{f : A \to B \text{ bijective} \quad g : B \to C \text{ bijective}}{g \circ f : A \to C \text{ bijective}}$$

9.7. **Restrictions of functions.** Suppose that $f\colon A \to B$ is a function and $C \subseteq A$. Then we can *restrict* f to C. This results in a function with domain C. The restriction is written $f|_C\colon C \to B$.

The restriction is defined by the rule $f|_C(c) \stackrel{\text{def}}{=} f(c)$ for all $c \in C$. In other words, if $f\colon A \to B$ is defined by a certain rule $a \mapsto \varphi(a)$, then $f|_C\colon C \to B$ is defined by *the same rule*: $c \mapsto \varphi(c)$. The only difference is that this variable c has values only in the subset C.

The restriction of an injective function remains injective, but it is not necessarily true that the restriction of a surjective function remains a surjective.

There is another concept that is used from time to time in mathematics: *restriction of codomain.* Suppose that $f\colon A \to B$ and that $D \subseteq B$. If D is large enough to contain the image $f[A]$ (so that $f[A] \subseteq D$) then we can form a function $f'\colon A \to D$ defined by the same rule as f. In other words $fa = f'a$ for all $a \in A$. The only difference between f and f' is the codomain. (There is no standard notation for the restriction of codomain: we used a here prime). If $f[A]$ is not a subset of D, then the restriction of codomain is not allowed: it results in a function that is not well-defined.

9.8. **Inclusions Functions.** If $A \subseteq B$ then there is an inclusion function $\iota\colon A \to B$ that behaves similarly to an identity function (except the domain is not necessarily equal to the codomain). So $\iota(a) = a$ for all $a \in A$. The function has the effect of *including* A into B. The inclusion function is injective, but not surjective (unless $A = B$).

10. Binary relations

A *binary relation* on a set A is a subset R of the Cartesian product $A \times A$. In other words, R is a set of ordered pairs with first and second coordinate in A.

If $(x, y) \in R$ where R is a relation on A, then we say that x and y are related by R. We often write this as xRy using *infix* notation. If $(x, y) \notin R$ the we say that x and y are not related by R, and write $\neg(xRy)$ or $x \not R\, y$.

We can think of $=$ as giving a binary relation on the set A by defining the relation R to be the set $\{(x, x) \mid x \in A\}$. This set is sometimes called the *diagonal* or the *graph of the identity function*.

10.1. **Types of Binary Relations.** We review the concept of reflexive, symmetric, and transitive relations.

A binary relation is *reflexive* if xRx for all $x \in R$.

A binary relation is *symmetric* means that for all $x, y \in A$ with xRy we have yRx.

A binary relation is *transitive* means that for all $x, y, z \in A$ with xRy and yRz, we have xRz.

10.2. **Equivalence relations.** A binary relation that is (i) reflexive, (ii) symmetric, and (iii) transitive is called an *equivalence relation* on R. For example, $=$ is an equivalence relation.

For the rest of this section we assume that R is an equivalence relation on A. We write xRy as $x \sim y$. In other words, since we have an equivalence relation we have $x \sim x$ for all $x \in A$, $x \sim y \implies y \sim x$ for all $x, y \in A$, and $(x \sim y) \wedge (y \sim z) \implies x \sim z$ for all $x, y, z \in A$.

We define the *equivalence class* $[x]$ to be the following set

$$[x] \stackrel{\text{def}}{=} \{y \in A \mid x \sim y\}$$

The only two equivalence classes can intersect is for them to be equal. In other words, if $[x] \cap [y]$ is not empty, then $[x] = [y]$. By the reflexive law $x \in [x]$, so every element is in an equivalence class, and each equivalence class has at least one element. Thus equivalence classes *partitions* all of A into disjoint, nonempty subsets.

We will need the following (related) laws:

$$[x] = [y] \iff x \sim y$$
$$x \in [x]$$
$$x \in [y] \iff y \in [x]$$
$$x \in [y] \iff [x] = [y]$$

Exercise 1. Prove that if $[x] \cap [y]$ is not empty, then $x \sim y$. Prove it using the definition of equivalence class, plus the symmetric and transitive laws.