MATH 145 Algebra, Lecture Notes

by Stephen New

Chapter 1. Sets and Mathematical Statements

1.1 Remark: A little over 100 years ago, it was found that some mathematical proofs contained paradoxes, and these paradoxes could be used to prove statements that were known to be false. One well-known paradox, outside of the realm of mathematics, is the statement

"This statement is false".

The above statement is true if and only if it is false. It is one form of a paradox known as the **liar's paradox**. After examining some lengthy and convoluted mathematical proofs which contained paradoxes, Bertrand Russell came up with the following mathematical paradox, which is somewhat similar to the liar's paradox:

Let X be the set of all sets, and let $S = \{A \in X \,|\, A \notin A\}$.
Note for example that $\mathbf{Z} \notin \mathbf{Z}$ so $\mathbf{Z} \in S$, and $X \in X$ so $X \notin S$.
Then we have $S \in S$ if and only if $S \notin S$.

This paradox is known as **Russell's paradox**. With Russell's paradox, it was possible to construct a proof by contradiction, which followed all the accepted rules of mathematical proof, of any statement whatsoever. Mathematicians realized that they would need to modify the accepted framework of mathematics in order to ensure that mathematical paradoxes could no longer arise. They were led to consider the following three questions.

1. Exactly what is an allowable mathematical object?
2. Exactly what is an allowable mathematical statement?
3. Exactly what is an allowable mathematical proof?

Eventually, after a great deal of work by many mathematicians, a consensus was reached as to the answers to these three questions. Roughly speaking, the answers are as follows. Essentially every mathematical object is a mathematical **set** (this includes objects that we would not normally consider to be sets, such as integers and functions), and a mathematical set can be constructed using certain specific rules, known as the **Zermelo-Fraenkel Axioms** along with the **Axiom of Choice** which, together, are referred to as the **ZFC axioms**. Mathematical statements are normally expressed using a combination of mathematical symbols and words from a natural language, such as English, but every mathematical statement can be expressed as a so-called **formula** in a certain specific formal symbolic language, called the language of **first-order set theory**, which uses symbols rather than words. Mathematical proofs are likewise normally expressed using a combination of symbols and words, but every mathematical proof can be translated into a very precise symbolic form of proof called a **derivation**. One kind of derivation consists of a finite list of ordered pairs $(\mathcal{S}_n, F_n)$ (which we think of as *proven theorems*), where each S_n is a finite set of formulas (called the *premises*) and each F_n is a single formula (called the *conclusion*), such that each pair $(\mathcal{S}_n, F_n)$ can be obtained from previous pairs $(\mathcal{S}_i, F_i)$ with $i < n$, using certain specific proof rules. In this chapter we shall provide more detailed answers to the the first two of the above three questions, and in the next chapter we shall consider the third question.

1.2 Definition: Every mathematical object is either a (mathematical) **set** or a (mathematical) **class**. Every set or class is a collection of sets. When S is a set or a class and x is a set, we write $x \in S$ to indicate that x is an **element** of S. When A and B are sets, we say that A is **equal** to B, and we write $A = B$, when A and B have the same elements, we say that A is a **subset** of B, and we write $A \subseteq B$ (some books write $A \subset B$), when every element of A is also an element of B, and we say that A is a **proper subset** of B, and we write $A \subset B$, or for emphasis $A \subsetneqq B$, when $A \subseteq B$ but $A \neq B$. When $F(x)$ is a mathematical statement about a set x, we write $\{x \,|\, F(x)\}$ to denote the collection of all sets x for which the statement $F(x)$ is true. When $F(x)$ is a statement about a set x and A is a set, we write $\{x \in A \,|\, F(x)\}$ to denote the collection $\big\{x \,\big|\, x \in A \text{ and } F(x)\big\}$.

A (mathematical) **class** is any collection of sets of the form $\{x \,|\, F(x)\}$ where $F(x)$ is a mathematical statement about x.

A (mathematical) **set** is a collection of sets which can be constructed using certain specific rules, which are known as the **ZFC axioms** (or the **Zermelo-Fraenkel Axioms** along with the **Axiom of Choice**). The ZFC axioms include (or imply) each of the following.

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

Empty Set Axiom: There is set called the **empty set**, denoted by $\emptyset$, with no elements.

Pair Axiom: If A and B are sets then $\{A, B\} = \big\{x \,\big|\, x = A \text{ or } x = B\big\}$ is a set.

Union Axiom: If S is a set of sets then $\bigcup S = \bigcup_{A \in S} A = \big\{x \,\big|\, x \in A \text{ for some } A \in S\big\}$ is a set.

Power Set Axiom: If A is a set then $\mathcal{P}(A) = \big\{X \big| X \subseteq A\big\}$ is a set, which we call the **power set** of A.

Axiom of Infinity: If we define the **natural numbers** to be the sets $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$ and so on, then $\mathbf{N} = \{0, 1, 2, 3, \cdots\}$ is a set.

Separation Axiom: If A is a set and $F(x)$ is a statement about x, $\big\{x \in A \big| F(x)\big\}$ is a set.

Replacement Axiom: If A is a set and $F(x, y)$ is a statement about x and y with the property that for every set x there exists a unique set $y = f(x)$ for which $F(x, y)$ is true, then $\big\{f(x) \,\big|\, x \in A\big\} = \big\{y \,\big|\, \exists\, x \in A \; F(x, y)\big\}$ is a set.

Axiom of Choice: Given a nonempty set S of non-empty disjoint sets, there exists a set C which contains exacly one element from each of the sets in S.

1.3 Definition: For sets A and B, we use the following notation. We denote the **union** of A and B by $A \cup B$, the **intersection** of A and B by $A \cap B$, the set A **remove** B by $A \setminus B$ and the **product** of A and B by $A \times B$, that is

$$\begin{aligned}
A \cup B &= \bigcup\{A, B\} = \big\{x \big| x \in A \text{ or } x \in B\big\},\\
A \cap B &= \big\{x \in A \cup B \big| x \in A \text{ and } x \in B\big\},\\
A \setminus B &= \big\{x \in A \big| x \notin B\big\}, \text{ and}\\
A \times B &= \big\{(x, y) \in \mathcal{P}\big(\mathcal{P}(A \cup B)\big) \big| x \in A \text{ and } b \in B\big\}
\end{aligned}$$

where the **ordered pair** (x, y) is defined to be the set $(x, y) = \big\{\{x\}, \{x, y\}\big\}$. We say that A and B are **disjoint** when $A \cap B = \emptyset$. We also write $A^2 = A \times A$.

1.4 Theorem: *(Properties of Sets) Let $A, B, C \subseteq X$. Then*

(1) (Idempotence) $A \cup A = A$, $A \cap A = A$,
(2) (Identity) $A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$, $A \cup X = X$, $A \cap X = A$,
(3) (Associativity) $(A \cup B) \cup C = A \cup (B \cup C)$ and $(A \cap B) \cap C = A \cap (B \cap C)$,
(4) (Commutativity) $A \cup B = B \cup A$ and $A \cap B = B \cap A$,
(5) (Distributivity) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ and $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,
(6) (De Morgan's Laws) $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$ and $X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B)$.

Proof: We shall provide some proofs later once we have listed some methods of proof.

1.5 Example: When A is a set, we have $\{A, A\} = \{A\}$ by the Equality Axiom and so $\{A\}$ is a set by the Pair Axiom. In particular, since $\emptyset$ is a set, so is $\{\emptyset\}$. Note that $\emptyset \neq \{\emptyset\}$, indeed the set $\emptyset$ has no elements but $\{\emptyset\}$ has one element. Since $\emptyset$ and $\{\emptyset\}$ are sets, so is the set $\big\{\emptyset, \{\emptyset\}\big\}$ by the Pair Axiom. Using the Pair Axiom and the Union Axiom we can then construct the set $\big\{\emptyset\,,\, \{\emptyset\}\big\} \cup \big\{\big\{\emptyset, \{\emptyset\}\big\}\big\} = \big\{\emptyset\,,\, \{\emptyset\}\,,\, \big\{\emptyset, \{\emptyset\}\big\}\big\}$. The first few natural numbers are given by $0 = \emptyset$, $1 = \{0\} = \{\emptyset\}$, $2 = \{0, 1\} = \big\{\emptyset, \{\emptyset\}\big\}$ and $3 = \{0, 1, 2\} = \big\{\emptyset\,,\, \{\emptyset\}\,,\, \{\emptyset, \{\emptyset\}\}\big\}$. Having constructed the natural number n as a set, the number $n+1$ is defined to be the set $n+1 = n \cup \{n\}$ (which is a set by the Pair and Union Axioms). We remark that although we only need to use the Pair Axion and the Union Axiom to construct any given natural number n, we need to use the Axiom of Infinity to conclude that the collection of all natural numbers is a set.

1.6 Example: Since the natural number $3 = \{0, 1, 2\}$ is a set, so is its power set

$$\mathcal{P}(3) = \mathcal{P}\big(\{0, 1, 2\}\big) = \Big\{\emptyset\,,\, \{0\}\,,\, \{1\}\,,\, \{2\}\,,\, \{1, 2\}\,,\, \{0, 2\}\,,\, \{0, 1\}\,,\, \{0, 1, 2\}\Big\}.$$

We remark that when a set A has n elements, its power set $\mathcal{P}(A)$ has 2^n elements. We also remark that when A is a finite set we do not need to use the Power Set Axiom to construct the set $\mathcal{P}(A)$ since we can construct the set $\mathcal{P}(A)$ using the Pair and Union Axioms.

1.7 Remark: It was mentioned earlier that essentially all mathematical objects are sets, including objects that we do not normally consider to be sets such as numbers and functions. We have seen that the natural numbers $0, 1, 2, \cdots$ are defined to be sets. In the following two definitions we indicate how functions and relations are defined to be sets.

1.8 Definition: When A and B are sets, a **function** from A to B is defined to be a set $F \subseteq A \times B$ with the property that for every $x \in A$ there exists a unique element $y \in B$ such that $(x, y) \in F$. We write $F : A \to B$ to indicate that F is a function from A to B, and we write $y = F(x)$ to indicate that $(x, y) \in F$. Thus a function is in fact defined to be equal to what we would normally consider to be its graph.

1.9 Definition: When A is a set, a **binary relation** on A is a subset $R \subseteq A^2$. When $x, y \in A$ we write xRy to indicate that $(x, y) \in R$.

1.10 Example: The operation $+$ on $\mathbf{N}$ is a function $+ : \mathbf{N}^2 \to \mathbf{N}$ and for $x, y \in \mathbf{N}$ we write $x + y$ to denote $+(x, y)$. The relation $<$ on $\mathbf{N}$ is a subset $< \subseteq \mathbf{N}^2$ and for $x, y \in \mathbf{N}$ we write $x < y$ to indicate that $(x, y) \in <$.

1.11 Remark: Using the axioms of set theory, it is possible to construct the set of integers $\mathbf{Z} = \{0, \pm 1, \pm 2, \cdots\}$, the sets of rational numbers $\mathbf{Q} = \big\{\frac{k}{n} \,\big|\, k, n \in \mathbf{Z}, n > 0\big\}$ and the set of real numbers $\mathbf{R}$, along with their usual operations $+, -, \times, \div$ and their usual inequality relations $<, \leq, >, \geq$. This procedure is outlined in the Appendix 1.

1.12 Definition: Every mathematical statement can be expressed in a formal symbolic language called the language of **first-order set theory**, which we shall describe below. For the moment, we describe a very simple formal symbolic language that captures some of the features of mathematics. In the language of **propositional logic**, we use symbols from the **symbol set** $\{\neg\,,\,\wedge\,,\,\vee\,,\,\rightarrow\,,\,\leftrightarrow\,,\,(\,,\,)\}$ along some variable symbols which we denote by $P, Q, R, \cdots$. The symbols (and) are called parentheses or brackets and the other symbols in the symbol set represent English words as follows:

$\neg$	$\wedge$	$\vee$	$\rightarrow$	$\leftrightarrow$
not	and	or	implies	if and only if

The propositional variables $P, Q, R, \cdots$ are intended to represent certain unknown mathematical statements which are assumed to be either true or false, but never both. When X and Y are strings of symbols, we shall write $X \equiv Y$ when X and Y are **identical**.

A **formula**, in propositional logic, is a string of symbols which can be obtained using the following rules:

F1. Every propositional variable symbol is a formula.
F2. If F is a formula then so is the string $\neg F$.
F3. If F and G are formulas then so are the strings $(F \wedge G)$, $(F \vee G)$, $(F \rightarrow G)$ and $(F \leftrightarrow G)$.

A **derivation** for a formula F is a list of formulas $F_1, F_2, \cdots, F_n$ with $F \equiv F_m$ for some $m \leq n$ (usually $F \equiv F_n$) such that each formula F_k is obtained by applying one of the above 3 rules to previous formulas in the list.

We shall often omit the outermost pair of brackets from formulas, for example we might write the formula $(P \vee (Q \rightarrow R))$ as $P \vee (Q \rightarrow R)$.

1.13 Example: The string $F \equiv \neg\big(P \rightarrow \neg(Q \wedge P)\big)$ is a formula. One derivation for F is as follows.

$$P\,,\; Q\,,\; (Q \wedge P)\,,\; \neg(Q \wedge P)\,,\; \big(P \rightarrow \neg(Q \wedge P)\big)\,,\; F.$$

1.14 Definition: An **assignment** of truth-values is a function $\alpha : \{P, Q, R, \cdots\} \rightarrow \{0, 1\}$. When $\alpha(P) = 1$ we say that P is **true** under the assignment α and when $\alpha(P) = 0$ we say that P is **false** under α, and similarly for the variables $Q, R, \cdots$.

Given an assignment α, we define $\alpha(F)$, for any formula F, recursively as follows:

A1. $\alpha(P)$, $\alpha(Q)$ and $\alpha(R)$, and so on, are already known.
A2. If G is a formula then $\alpha(\neg G)$ is defined according to the table

G	$\neg G$
1	0
0	1

A3. If G and H are formulas then $\alpha(G \wedge H)$, $\alpha(G \vee H)$, $\alpha(G \rightarrow H)$ and $\alpha(G \leftrightarrow H)$ are defined according to the table

G	H	$(G \wedge H)$	$(G \vee H)$	$(G \rightarrow H)$	$(G \leftrightarrow H)$
1	1	1	1	1	1
1	0	0	1	0	0
0	1	0	1	1	0
0	0	0	0	1	1

When $\alpha(F) = 1$ we say that F is **true** under the assignment α, and when $\alpha(F) = 0$ we say that F is **false** under α.

1.15 Note: When $\alpha(G) = 0$ we have $\alpha(G \to H) = 1$. This agrees with the way in which we use the word "implies" in mathematics. For example, for every integer x, the statement "if $0 = 1$ then $x = 3$" is considered to be true.

1.16 Example: Let $F \equiv (P \wedge \neg(Q \to P)) \vee (R \leftrightarrow \neg Q)$, and let α be any assignment with $\alpha(P) = 1$, $\alpha(Q) = 0$ and $\alpha(R) = 1$. Determine whether F is true under α.

Solution: We make a derivation $F_1\ F_2\ \cdots\ F_n$ for F, and under each formula F_i we put the truth value $\alpha(F_i)$, which we find using the above definition.

P	Q	R	$Q \to P$	$\neg(Q \to P)$	$P \wedge \neg(Q \to P)$	$\neg Q$	$R \leftrightarrow \neg Q$	F
1	0	1	1	0	0	1	1	1

The table shows that $\alpha(F) = 1$.

1.17 Definition: An assignment on the propositional variables $P_1, \cdots, P_n$ is a function $\alpha : \{P_1, P_2, \cdots, P_n\} \to \{0, 1\}$ (there are 2^n such assignments). A **truth-table**, on the variables $P_1, P_2, \cdots P_n$, for the formula F, is a table in which

T1. The header row is a derivation $F_1\ F_2\ \cdots\ F_n\ \cdots\ F_l$ for F, where the formulas F_k use no propositional variables other than $P_1, \cdots, P_n$, with $F_k = P_k$ for $1 \le k \le n$.

T2. There are 2^n rows (not counting the header row): for each of the 2^n assignments α on $P_1, \cdots, P_n$, there is a row of the form $\alpha(F_1)\ \alpha(F_2)\ \cdots\ \alpha(F_l)$.

T3. The rows are ordered so that first n columns (headed by $P_1, \cdots, P_n$) list the binary numbers in decreasing order from $11\cdots1$ at the top down to $00\cdots0$ at the bottom.

1.18 Example: Make a truth-table on P, Q and R for the formula $F \equiv \neg((P \vee \neg Q) \to R)$.

Solution: We make a table, as in example 1.16, but with $2^3 = 8$ rows.

P	Q	R	$\neg Q$	$P \vee \neg Q$	$(P \vee \neg Q) \to R$	F
1	1	1	0	1	1	0
1	1	0	0	1	0	1
1	0	1	1	1	1	0
1	0	0	1	1	0	1
0	1	1	0	0	1	0
0	1	0	0	0	1	0
0	0	1	1	1	1	0
0	0	0	1	1	0	1

1.19 Definition: Let F and G be formulas and let $\mathcal{S}$ be a set of formulas.

(1) We say F is a **tautology**, and write $\models F$, when for all assignments α we have $\alpha(F) = 1$.
(2) We say that F is a **contradiction** when $\models \neg F$.
(3) We say F is **equivalent** to G, and write $F \cong G$, when for all assignments α, $\alpha(F) = \alpha(G)$.
(4) We say the argument "$\mathcal{S}$ therefore G" is **valid**, or we say that $\mathcal{S}$ **induces** G, or that G is a **consequence** of $\mathcal{S}$, and we write $\mathcal{S} \models G$, when for all assignments α, if $\alpha(F) = 1$ for every $F \in \mathcal{S}$ then $\alpha(G) = 1$. In the case that $\mathcal{S} = \{F_1, F_2, \cdots, F_n\}$ we often omit the set brackets and write $\mathcal{S} \models G$ as $F_1, F_2, \cdots, F_n \models G$. The formulas in $\mathcal{S}$ are called the **premises** of the argument and G is called the **conclusion**.

1.20 Theorem: *Let F, G and $F_1, \cdots, F_n$ be formulas. Then*
(1) $\models F \Longleftrightarrow \emptyset \models F$,
(2) $F \models G \Longleftrightarrow \models (F \to G)$,
(3) $F \cong G \Longleftrightarrow \big(F \models G$ *and* $G \models F\big) \Longleftrightarrow \models (F \leftrightarrow G)$, *and*
(4) $\{F_1, F_2, \cdots, F_n\} \models G \Longleftrightarrow (\cdots((F_1 \wedge F_2) \wedge F_3) \wedge \cdots \wedge F_n) \models G$.

Proof: We shall provide some proofs once we have discussed proof methods.

1.21 Example: Let $F \equiv (P \leftrightarrow ((Q \wedge \neg R) \vee S)) \vee (P \to \neg S)$. Determine whether $\models F$.

Solution: We make a truth-table for F.

P	Q	R	S	$\neg R$	$Q \wedge \neg R$	$(Q \wedge \neg R) \vee S$	$P \leftrightarrow ((Q \wedge \neg R) \vee S)$	$\neg S$	$P \to \neg S$	F
1	1	1	1	0	0	1	1	0	0	1
1	1	1	0	0	0	0	0	1	1	1
1	1	0	1	1	1	1	1	0	0	1
1	1	0	0	1	1	1	1	1	1	1
1	0	1	1	0	0	1	1	0	0	1
1	0	1	0	0	0	0	0	1	1	1
1	0	0	1	1	0	1	1	0	0	1
1	0	0	0	1	0	0	0	1	1	1
0	1	1	1	0	0	1	0	0	1	1
0	1	1	0	0	0	0	1	1	1	1
0	1	0	1	1	1	1	0	0	1	1
0	1	0	0	1	1	1	0	1	1	1
0	0	1	1	0	0	1	0	0	1	1
0	0	1	0	0	0	0	1	1	1	1
0	0	0	1	1	0	1	0	0	1	1
0	0	0	0	1	0	0	1	1	1	1

Since all the entries in the F-column are equal to 1, we have $\models F$.

1.22 Example: Let $F \equiv (P \vee Q) \to R$ and $G \equiv (P \to R) \vee (Q \to R)$. Determine whether $F \cong G$.

Solution: We make a truth-table for F and G,

P	Q	R	$P \vee Q$	F	$P \to R$	$Q \to R$	G
1	1	1	1	1	1	1	1
1	1	0	1	0	0	0	0
1	0	1	1	1	1	1	1
1	0	0	1	0	0	1	1
0	1	1	1	1	1	1	1
0	1	0	1	0	1	0	1
0	0	1	0	1	1	1	1
0	0	0	0	1	1	1	1

The F-column is not the same as the G-column, for example on the 4^{th} row, F is false and G is true, and so $F \ncong G$.

1.23 Example: Let $F \equiv (P \vee \neg Q) \to R$, $G \equiv P \leftrightarrow (Q \wedge R)$, and $H \equiv (Q \to R)$, and let $K = \neg(\neg Q \wedge R)$. Determine whether $\{F, G, H\} \models K$.

Solution: In general, we have $\{F_1, \cdots, F_n\} \models K$ if and only if in a truth-table for the formulas $F_1, \cdots, F_n$ and K, for every row in which $F_1, \cdots, F_n$ are all true, we also have K true. We make a truth-table for F, G, H and K:

P	Q	R	$\neg Q$	$P \vee \neg Q$	F	$Q \wedge R$	G	H	$\neg Q \wedge R$	K
1	1	1	0	1	1	1	1	1	0	1
1	1	0	0	1	0	0	0	0	0	1
1	0	1	1	1	1	0	0	1	1	0
1	0	0	1	1	0	0	0	1	0	1
0	1	1	0	0	1	1	0	1	0	1
0	1	0	0	0	1	0	1	0	0	1
0	0	1	1	1	1	0	1	1	1	0
0	0	0	1	1	0	0	1	1	0	1

On row 7, F G and H are all true but K is false. This implies that $\{F, G, H\} \not\models K$.

1.24 Example: Determine whether $\{P \vee Q, \neg Q, P \to Q\} \models \neg P$.

Solution: We have

P	Q	$P \vee Q$	$\neg Q$	$P \to Q$	$\neg P$
1	1	1	0	1	0
1	0	1	1	0	0
0	1	1	0	1	1
0	0	0	1	1	1

Notice that there are no rows in which the premises are all true. In this situation, we can conclude that $\{P \vee Q, \neg Q, P \to Q\} \models \neg P$ (we don't even need to look at the last column of the table).

1.25 Example: Let F and G be formulas. Consider the following table.

F	G	$F \leftrightarrow G$	$\neg G$	$F \to \neg G$
1	1	1	0	0
1	0	0	1	1
0	1	0	0	1
0	0	1	1	1

Notice that on the first row (when F and G are both true) we have $F \leftrightarrow G$ true and $F \to \neg G$ false. This might seem to imply that $(F \leftrightarrow G) \not\models (F \to \neg G)$, but it does not! For example, if $F \equiv P$ and $G \equiv \neg P \wedge Q$, then we never have F and G both true, so the combination of truth-values shown in the first row of the above table never actually occurs. The above table is not actually a truth-table as defined in 1.17. Rather, it is a table of possible combinations of truth-values which may or may not actually occur.

1.26 Definition: Let A be a set. Recall that $A^2 = A \times A$. A **unary function** on A is a function $f : A \to A$. A **binary function** on A is a function $g : A^2 \to A$. Many binary functions g are used with **infix notation** which means that we write $g(x, y)$ as xgy. For example, $+$ is a binary function on $\mathbf{N}$ and we write $+(x, y)$ as $x+y$. A **unary relation** on A is a subset $P \subseteq A$, and we write $P(x)$ to indicate that $x \in P$. A **binary relation** on A is a subset $R \subseteq A^2$. When R is written with **prefix notation** we write $R(x, y)$ to indicate that $(x, y) \in R$, and when R is used with **infix notation**, we write xRy to indicate that $(x, y) \in R$. For example, $<$ is a binary relation on the set $\mathbf{N}$, which means that $< \subseteq \mathbf{N}^2$, and we write $x < y$ to indicate that $(x, y) \in <$.

1.27 Definition: We now describe a type of formal symbolic language, called a **first-order language**. Every first-order language uses symbols from the **common symbol set**

$$\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, =, \forall, \exists, (,), ,\}$$

together with some **variable** symbols (such as x, y, z, u, v and w), and possibly some **additional symbols** which might include some **constant** symbols (such as a, b, c, $\emptyset$, 0, 1, e, π), some **function** symbols (such as f, g, h, $\cap$, $\cup$, $+$, $-$, $\times$), and some **relation** symbols (such as P, Q, R, $\in$, $\subset$, $\subseteq$, $<$, $\leq$). The variable symbols are intended to represent elements in a certain set (or class) U, called the **universal set** (or the **universal class**), which is usually understood from the context. The symbol $\forall$ is read as "for all" or "for every" and the symbol $\exists$ is read as "for some" or "there exists".

A **term** in the first-order language is a string of symbols using only variable, constant and function symbols, along with parentheses and commas if necessary, which can be obtained using the following rules.

T1. Every variable symbol is a term and every constant symbol is a term.
T2. If t is a term and f is a unary function symbol then the string $f(t)$ is a term.
T3. If s and t are terms and g is a binary function symbol then the string $g(s, t)$ (or the string (sgt) in the case that g is used with infix notation) is a term. We sometimes omit the brackets from the string (sgt).

Here are some examples of terms (with some brackets omitted):

$$u \ , \ u \cap v \ , \ u \cap (v \cup \emptyset) \ , \ x \ , \ x + 1 \ , \ x \times (y + 1) \ , \ f(x) \ , \ g(x, y) \ , \ g(x + 1, f(y))$$

Each term represents an element in the universal set (or class) U.

A **formula** is a string of symbols which can be obtained using the following rules.

F1. If t is a term and P is a unary relation symbol, then the string $P(t)$ is a formula.
F2. If s and t are terms and R is a binary relation symbol then the string $R(s, t)$ (or the string sRt in the case that R is used with infix notation) is a formula.
F3. If F is a formula then so is the string $\neg F$.
F4. If F and G are formulas then so are the strings $(F \wedge G)$, $(F \vee G)$, $(F \rightarrow G)$ and $(F \leftrightarrow G)$.
F5. If F is a formula and x is a variable symbol, the strings $\forall x \ F$ and $\exists x \ G$ are formulas.

Here are some examples of formulas (with some brackets omitted):

$$u \subseteq v \ , \ u \cap v = \emptyset \ , \ (x = 0 \vee x = 1) \ , \ \forall x \, (x \neq 0 \rightarrow \exists y \ x \times y = 1)$$

A formula is a precise way of expressing a mathematical statement about elements in U.

1.28 Definition: In the language of **first-order number theory**, in addition to symbols from the common symbol set $\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, =, \forall, \exists, (,), ,\}$ along with variable symbols such as x, y, z, we allow ourselves to use additional symbols from the **additional symbol set** $\{0, 1, +, \times, <\}$. (the symbols 0 and 1 are constant symbols, the symbols + and $\times$ are binary function symbols used with infix notation, and the symbol $<$ is a binary relation symbol used with infix notation). Unless explicitly stated otherwise, we do not allow ourselves to use any other symbols (such as 2, $-$, $>$).

1.29 Example: Express each of the following statements about integers as formulas in the language of first-order number theory, taking the universal set to be $U = \mathbf{Z}$.
(a) x is a factor of y.
(b) $z = \min\{x, y\}$.
(c) x is prime.
(d) x is a power of 2.

Solution: The statement "x is a factor of y" can be expressed as $\exists z\ y = x \times z$.
The statement "$z = \min\{x, y\}$" can be expressed as $(x < y \rightarrow z = x) \wedge (\neg\, x < y \rightarrow z = y)$.
The statement "x is prime" can be expressed as

$$1 < x \wedge \forall x\, \forall y\, ((1 < y \wedge 1 < z) \rightarrow \neg\, x = y \times z).$$

The statement "x is a power of 2" is equivalent to the statement "x is positive and every factor of x which is greater than 1 is even" which can be expressed as

$$0 < x \wedge \forall y\, ((1 < y \wedge \exists z\ x = y \times z) \rightarrow \exists z\ y = z + z).$$

1.30 Example: Express each of the following statements about a function $f : \mathbf{R} \rightarrow \mathbf{R}$ as formulas in the language of first-order number theory, taking the universal set to be $U = \mathbf{R}$ and allowing the use of the additional unary function symbol f.
(a) f is nondecreasing.
(b) f is bijective.
(c) $\lim\limits_{x \to a} f(x) = b$.

Solution: The statement "f is nondecreasing" means "for all $x, y \in \mathbf{R}$, if $x \leq y$ then $f(x) \leq f(y)$" which can be expressed as $\forall x\ \forall y\ \big(\neg\, y < x \rightarrow \neg\, f(y) < f(x)\big)$.
The statement "f is bijective" means "for all $y \in \mathbf{R}$ there exists a unique $x \in \mathbf{R}$ such that $y = f(x)$" which can be expressed as $\forall y\ \exists x \big(y = f(x) \wedge \forall z (y = f(z) \rightarrow z = x)\big)$.
The statement "$|x - a| < \delta$" is equivalent to $-\delta < x - a$ and $x - a < \delta$" which can be expressed as the formula $(a < x + \delta \wedge x < a + \delta)$. The statement "$\lim\limits_{x \to a} f(x) = b$" means "for every $\epsilon > 0$ there exists $\delta > 0$ such that for all x, if $0 < |x - a| < \delta$ then $|f(x) - b| < \epsilon$", which can be expressed as the formula

$$\forall \epsilon \big(0 < \epsilon \rightarrow \exists \delta\, \big(0 < \delta \wedge \forall x \big((\neg\, x = a \wedge (a < x + \delta \wedge x < a + \delta)) \rightarrow (b < f(x) + \epsilon \wedge f(x) < b + \epsilon)\big)\big)\big).$$

In the above formula, the symbols ϵ and δ are being used as variable symbols.

1.31 Definition: In the language of **first-order set theory**, in addition to the symbols from the common symbol set along with variable symbols, the only additional symbol that we allow ourselves to use (unless explicitly stated otherwise) is the symbol $\in$, which is a binary relation symbol used with infix notation. When we use the language of first-order set theory, unless indicated otherwise we shall take the universal class to be the class of all sets.

1.32 Remark: Every mathematical statement can, in principle, be expressed in the language of first-order set theory.

1.33 Example: Let u, v and w be sets. The mathematical statement $u \subseteq v$ can be expressed as the formula $\forall x(x \in u \to x \in v)$. The mathematical statement $w = \{u, v\}$ is equivalent (by the Equality Axiom) to the statement "for every set x, we have $x \in w$ if and only if $x \in \{u, v\}$" which can be expressed as the formula $\forall x\big(x \in w \leftrightarrow (x = u \vee x = v)\big)$. The mathematical statement $w = u \cup v$ can be expressed as $\forall x\,\big(x \in w \leftrightarrow (x \in u \vee x \in v)\big)$.

1.34 Example: Each of the ZFC axioms can be expressed as a formula in the language of first-order set theory. Here are a few of the axioms expressed as formulas.

Equality Axiom: $\forall u\,\forall v\,\big(u = v \leftrightarrow \forall x(x \in u \leftrightarrow x \in v)\big)$
Empty Set Axiom: $\exists u\,\forall x\;\neg x \in u$
Pair Axiom: $\forall u\,\forall v\,\exists w\,\forall x\big(x \in w \leftrightarrow (x = u \vee x = v)\big)$
Union Axiom: $\forall u\,\exists w\,\forall x\big(x \in w \leftrightarrow \exists v(v \in u \wedge x \in v)\big)$
Power Set Axiom: $\forall u\,\exists w\,\forall v\big(v \in w \leftrightarrow \forall x(x \in v \to x \in u)\big)$

1.35 Example: Express the mathematical statement $u = 2$ (that is, u is equal to the natural number 2) as a formula in first-order set theory.

Solution: The following statements are equivalent, and the final statement in the list is expressed in the form of a formula:

$$u = 2$$
$$u = \big\{\emptyset, \{\emptyset\}\big\}$$
$$\forall x\,\big(x \in u \leftrightarrow x \in \big\{\emptyset, \{\emptyset\}\big\}\big)$$
$$\forall x\,\big(x \in u \leftrightarrow \big(x = \emptyset \vee x = \{\emptyset\}\big)\big)$$
$$\forall x\,\big(x \in u \leftrightarrow \big(\forall y\,\neg y \in x \;\vee\; \forall y\,(y \in x \leftrightarrow y = \emptyset)\big)\big)$$
$$\forall x\,\big(x \in u \leftrightarrow \big(\forall y\,\neg y \in x \;\vee\; \forall y\,(y \in x \leftrightarrow \forall z\,\neg z \in y)\big)\big)$$

1.36 Remark: As the above example illustrates, although every mathematical statement can, in principle, be expressed as a formula in the language of first-order set theory in practice even fairly simple mathematical statements (such as the statement $u = 2$) can become extremely long and complicated and difficult to read when expressed as formulas. For this reason, as we build mathematics from the foundations of set theory by introducing new concepts and proving new theorems, we continually add new symbols to the symbol set and allow additional notation to be used.

1.37 Example: When the universal set is U and $A \subseteq U$ (in other words A is a unary relation on U), the statement "$x \in A$" can be expressed as the formula $A(x)$. When $S(x)$ is a mathematical statement about the variable x which can be expressed as the formula F, the statement "for all $x \in A$, $S(x)$ is true" can be expressed as the formula $\forall x(A(x) \to F)$, and the statement "there exists $x \in A$ such that $S(x)$" can be expressed as $\exists x(A(x) \wedge F)$.

1.38 Definition: Every occurrence of each variable symbol, which does not immediately follow a quantifier symbol, in a formula H is either **free** or **bound**, as follows. When H is of the form $P(t)$ or $R(s, t)$, every occurrence of each variable symbol is free. When H is of one of the forms $\neg F$, $(F \wedge G)$, $(F \vee G)$, $(F \to G)$ or $(F \leftrightarrow G)$, every occurrence of each variable symbol in H is free or bound in accordance with whether it was free or bound in F or in G. When H is of one of the forms $\forall x\,F$ or $\exists x\,F$, each occurrence of any variable symbol y other than x is free or bound in H in accordance with whether it was free or bound in F, every bound occurrence of x in F remains bound in H (and it is bound in H by the same quantifier symbol which bound it in F), and every free occurrence of x in F becomes bound in H (and it is bound by the initial quantifier symbol).

1.39 Example: The mathematical statement "x is a factor of y", which is a statement about integers x and y, can be expressed as the formula $\exists z\ y = x \times z$ in first-order number theory. In this formula, the variables x and y are free and the variable z is bound by the quantifier. Note that the statement is a statement about x and y but not about z.

1.40 Definition: An **interpretation** for a first-order language is given by specifying a non-empty universal set U, and by specifying exactly which constants, functions and relations are represented by each of the constant, function and relation symbols.

1.41 Note: Until we have chosen an interpretation, a formula F in a first-order language is nothing more than a meaningless string of symbols. Once we have chosen an interpretation, the formula becomes a meaningful statement about its free variables (that is about the elements in the universal set which are represented by the variable symbols which occur freely in F). The truth or falsehood of F may still depend on the values in U assigned to the free variables in F.

1.42 Example: Let F be the formula $\forall y\ x \times y = y \times x$. If we use the interpretation $\mathbf{R}$ (that is is we specify that the universal set is $\mathbf{R}$ and that the function symbol $\times$ represents multiplication) then the the formula F becomes the meaningful statement "the real number x commutes with every real number y" (which is true, no matter what number is assigned to the variable x). If we use the interpretation $\mathbf{R}^3$ (in which the universal set is $\mathbf{R}^3$ and the function symbol $\times$ represents cross product) then the formula becomes the meaningful statement "the vector $x \in \mathbf{R}^3$ has the property that $x \times y = y \times x$ for every vector $y \in \mathbf{R}^3$" (which is true if and only if x is the zero vector).

1.43 Definition: Given an interpretation U, an **assignment** (of values to the variable symbols) in U is a function

$$\alpha : \{\text{variable symbols}\} \to U\,.$$

For a formula F, we write $\alpha(F) = 1$ when F is true under the assignment α in the interpretation U, and we write $\alpha(F) = 0$ when F is false under the assignment α in U.

1.44 Definition: Let F and G be formulas and let $\mathcal{S}$ be a set of formulas.

(1) We say that F is a **tautology**, and we write $\models F$, when for all interpretations U and for all assignments α in U we have $\alpha(F) = 1$.
(2) We say that F is a **contradiction** when $\models \neg F$.
(3) We say that F and G are **equivalent**, and we write $F \cong G$, when for all interpretations U and all assignments α in U we have $\alpha(F) = \alpha(G)$.
(4) We say that the argument "$\mathcal{S}$ therefore G" is **valid**, or we say that $\mathcal{S}$ **induces** G, or that G is a **consequence** of $\mathcal{S}$, and we write $\mathcal{S} \models G$, when for all interpretations U and all assignments α in U, if $\alpha(H) = 1$ for every formula $H \in \mathcal{S}$ then $\alpha(F) = 1$.

1.45 Example: For any term t, we have $\models t = t$. When x and y are variables and f is a unary function symbol, we have $\models \exists y\ y = f(x)$. When a and b are constant symbols, we have $\not\models \neg\, a = b$. For any terms s and t, we have $s = t \cong t = s$. Although it is true in the interpretation $\mathbf{Z}$ that the formula $x \leq y$ has the same meaning as the formula $\neg\, y < x$, we have $x \leq y \not\cong \neg\, y < x$. When x is a variable symbol, a is a constant symbol, t is a term and R is a binary relation symbol, we have $\{\forall x\ xRa\} \models tRa$.

1.46 Note: Unlike the situation in propositional logic, in first order logic there is no routine algorithmic procedure that one can apply to determine whether a given formula is a tautology, or whether two given formulas are equivalent, or whether a given argument is valid. Sometimes we can solve such problems by constructing a mathematical proof.

.

Chapter 2. Mathematical Proof

2.1 Remark: At the end of the last chapter we raised the following questions. Given first-order formulas F and G, how can we determine whether $F \cong G$? Given a set of formulas S and a formula K, how can we determine whether $S \models K$? There is, in general, no routine algorithmic procedure to solve these two problems, but sometimes we can construct a **mathematical proof**. Just as a mathematical statement (normally expressed using a combination of mathematical symbols and words from a natural language such as English) can be expressed as a formula in a very precise formal symbolic language, so too a mathematical proof can be translated into a very precise symbolic form of proof called a **derivation**. In this chapter we shall describe two formal proof systems, one for deriving equivalences and one for deriving valid arguments.

2.2 Note: To state our proof rules precisely, we need to introduce one somewhat subtle concept, namely the concept of **substitution**, which is used in many mathematical proofs. For example, if we know that $a \leq x$ for every $x \in S$ and we know that $b \in S$, then we can conclude that $a \leq b$. In a detailed proof, we would break this into two steps, as follows.

1. Since $\forall x(x \in S \to a \leq x)$ it follows that $b \in S \to a \leq b$.
2. Since $b \in S$ and $b \in S \to a \leq b$ it follows that $a \leq b$.

In the first step, we used a substitution. In the formula $F \equiv (x \in S \to a \leq x)$ we replaced x by b. If write $[F]_{x \mapsto t}$ to denote the formula obtained from F by replacing the variable symbol x by the term t, then the proof rule that was invoked at step 1 was as follows: from $\forall x F$ we can conlude $[F]_{x \mapsto b}$.

When we define $[F]_{x \mapsto t}$, we want it to be the case that (once an interpretation has been chosen) the formula $[F]_{x \mapsto t}$ has the same meaning about t as the original formula F had about x. In general, this cannot be accomplished by simply replacing each occurrence of the symbol x by the term t. For example, in the interpretation $\mathbf{Z}$, the statement "x divides y" can be expressed using the formula $F \equiv \exists z\ y = x \times z$. We would like the formula $[F]_{x \mapsto u}$ to mean "u divides y", and this can be accomplished simply by replacing x by u to obtain $[F]_{x \mapsto u} \equiv \exists z\ y = u \times z$. But we would also like the formula $[F]_{x \mapsto z}$ to mean "z divides y" and if we simply replace x by z the formula becomes $\exists z\ y = z \times z$ which has a totally different meaning (it means "y is a perfect square"). To obtain the desired formula $[F]_{x \mapsto z}$ we first replace the bound variable z in F by the next available variable symbol u, then replace x by z afterwards, as follows

$$[F]_{x \mapsto z} \equiv \big[\exists z\ y = x \times z\big]_{x \mapsto z} \equiv \exists u \big[y = x \times u\big]_{x \mapsto z} \equiv \exists u\ y = z \times u.$$

2.3 Definition: Given a formula F, a variable symbol x, and a term t, we define the formula $[F]_{x \mapsto t}$ as follows. When F is obtained using rule F1 or F2, the formula $[F]_{x \mapsto t}$ is obtained from F by replacing all occurrences of the symbol x by the term t. To deal with rules F3 and F4 we define $\big[\neg F\big]_{x \mapsto t} \equiv \neg [F]_{x \mapsto t}$ and for $* \in \{\wedge, \vee, \to, \leftrightarrow\}$ we define $\big[(F * G)\big]_{x \mapsto t} \equiv \big([F]_{x \mapsto t} * [G]_{x \mapsto t}\big)$. To deal with rule F5, for $K \in \{\forall, \exists\}$ we define $\big[Kx\, F\big]_{x \mapsto t} \equiv Kx\, F$ (note that we do not need to change the formula when all occurrences of x are bound), and for a variable symbol y (which is different than x) we define $\big[Ky\, F\big]_{x \mapsto t}$ as follows. If y does not occur in t, we define $\big[Ky\, F\big]_{x \mapsto t} \equiv Ky[F]_{x \mapsto t}$. If y does occur in t, we define $\big[Ky\, F\big]_{x \mapsto t} \equiv Ku\big[[F]_{y \mapsto u}\big]_{x \mapsto t}$ where u is the first variable symbol which is not x and does not occur in F or in t. The formula $[F]_{x \mapsto t}$ is called the formula obtained from F by **substitution**, by replacing (free occurrences of) x by t.

2.4 Definition: For any formulas F, G and H, and any terms s and t, and any variables x and y, we have the following logical equivalences which are called the **basic equivalences**. The first 24 of these can be verified using truth-tables and the others are accepted axiomatically, without proof.

(Identity)	E1.	$F \cong F$
(Double Negation)	E2.	$F \cong \neg\neg F$
(Commutativity)	E3.	$F \wedge G \cong G \wedge F$
	E4.	$F \vee G \cong G \vee F$
(Associativity)	E5.	$F \wedge (G \wedge H) \cong (F \wedge G) \wedge H$
	E6.	$F \vee (G \vee H) \cong (F \vee G) \vee H$
(DeMorgan's Law)	E7.	$\neg(F \wedge G) \cong (\neg F \vee \neg G)$
	E8.	$\neg(F \vee G) \cong (\neg F \wedge \neg G)$
(Distributivity)	E9.	$F \wedge (G \vee H) \cong (F \wedge G) \vee (F \wedge H)$
	E10.	$F \vee (G \wedge H) \cong (F \vee G) \wedge (F \vee H)$
(Idempotence)	E11.	$F \wedge F \cong F$
	E12.	$F \vee F \cong F$
(Absorption)	E13.	$F \wedge (F \vee G) \cong F$
	E14.	$F \vee (F \wedge G) \cong F$
(Tautology)	E15.	$F \wedge (G \vee \neg G) \cong F$
	E16.	$F \vee (G \vee \neg G) \cong G \vee \neg G$
(Contradiction)	E17.	$F \wedge (G \wedge \neg G) \cong G \wedge \neg G$
	E18.	$F \vee (G \wedge \neg G) \cong F$
(Contrapositive)	E19.	$F \to G \cong \neg G \to \neg F$
(Implication)	E20.	$F \to G \cong \neg F \vee G$
	E21.	$\neg(F \to G) \cong F \wedge \neg G$
(If and Only If)	E22.	$F \leftrightarrow G \cong (F \wedge G) \vee (\neg F \wedge \neg G)$
	E23.	$F \leftrightarrow G \cong (\neg F \vee G) \wedge (F \vee \neg G)$
	E24.	$F \leftrightarrow G \cong (F \to G) \wedge (G \to F)$
(Equality)	E25.	$s = t \ \cong \ t = s$
(Double Quantifier)	E26.	$\forall x \, \forall y \, F \cong \forall y \, \forall x \, F$
	E27.	$\exists x \, \exists y \, F \cong \exists y \, \exists x \, F$
(Negate Quantifier)	E28.	$\neg\forall x \, F \cong \exists x \neg F$
	E29.	$\neg\exists x \, F \cong \forall x \, \neg F$
(Separate Quantifier)	E30.	$\forall x \, (F \wedge G) \cong \forall x \, F \wedge \forall x \, G$
	E31.	$\exists x \, (F \vee G) \cong \exists x \, F \vee \exists x \, G$
(Unused Variable)	E32.	$\forall x \, F \cong F$ if x is not free in F
	E33.	$\exists x \, F \cong F$ if x is not free in F
(Changing Variables)	E34.	$\forall x \, F \cong \forall y \, [F]_{x \mapsto y}$ if y is not free in F
	E35.	$\exists x \, F \cong \exists y \, [F]_{x \mapsto y}$ if y is not free in F

When $F \cong G$ (or $G \cong F$) is one of the above basic equivalences, and H is a formula which contains F as a sub-formula, and K is the formula obtained from H by replacing F by G, we say that K is obtained from G by **applying the basic equivalence** $F \cong G$ (or the equivalence $G \cong F$).

2.5 Definition: Given equivalent formulas F and G, a **derivation** of the equivalence $F \cong G$ is a list of formulas $F_1, F_2, \cdots, F_l$ with $F_1 = F$ and $F_l = G$ such that each formula F_{k+1} is obtained from the previous formula F_k by applying one of the basic equivalences.

2.6 Example: Let F and G be formulas. Make a derivation for $F \wedge (F \to G) \cong F \wedge G$.

Solution: Here is one possible derivation.

$F \wedge (F \to G)$	$\cong F \wedge (\neg F \vee G)$	Implication E20
	$\cong (F \wedge \neg F) \vee (F \wedge G)$	Distributivity E9
	$\cong (F \wedge G) \vee (F \wedge \neg F)$	Commutativity E4
	$\cong F \wedge G$	Contradiction E18

2.7 Example: Let F, G and H be formulas. Find a derivation for distributivity of $\vee$ over $\wedge$ from the right, that is for the logical equivalence $(F \wedge G) \vee H \cong (F \vee H) \wedge (G \vee H)$.

Solution: Here is a derivation.

$(F \wedge G) \vee H$	$\cong H \vee (F \wedge G)$	Commutativity E4
	$\cong (H \vee F) \wedge (H \vee G)$	Distributivity E10
	$\cong (F \vee H) \wedge (H \vee G)$	Commutativity E4
	$\cong (F \vee H) \wedge (G \vee H)$	Commutativity E4

2.8 Example: Derive the logical equivalence $F \to (G \to H) \cong (F \wedge G) \to H$.

Solution:

$F \to (G \to H)$	$\cong \neg F \vee (G \to H)$	Implication E20
	$\cong \neg F \vee (\neg G \vee H)$	Implication E20
	$\cong (\neg F \vee \neg G) \vee H$	Associativity E6
	$\cong \neg(F \wedge G) \vee H$	DeMorgan's Law E7
	$\cong (F \wedge G) \to H$	Implication E20

2.9 Example: Derive the logical equivalence $(F \wedge G) \to H \cong (F \to H) \vee (G \to H)$.

Solution:

$(F \wedge G) \to H$	$\cong \neg(F \wedge G) \vee H$	Implication E20
	$\cong (\neg F \vee \neg G) \vee H$	DeMorgan's Law E7
	$\cong (\neg F \vee \neg G) \vee (H \vee H)$	Idempotence E12
	$\cong ((\neg F \vee \neg G) \vee H) \vee H$	Associativity E6
	$\cong (\neg F \vee (\neg G \vee H)) \vee H$	Associativity E6
	$\cong (\neg F \vee (H \vee \neg G)) \vee H$	Commutativity E4
	$\cong ((\neg F \vee H) \vee \neg G) \vee H$	Associativity E6
	$\cong (\neg F \vee H) \vee (\neg G \vee H)$	Associativity E6
	$\cong (F \to H) \vee (\neg G \vee H)$	Implication E20
	$\cong (F \to H) \vee (G \to H)$	Implication E20

2.10 Example: Make a derivation for the equivalence $\exists x\,(F \to G) \cong \forall x\,F \to \exists x\,G$.

Solution: Here is a derivation.

$\exists x\,(F \to G)$	$\cong \exists x\,(\neg F \vee G)$	Implication E20
	$\cong \exists x \neg F \vee \exists x\,G$	Separating Quantifier E30
	$\cong \neg \forall x\,F \vee \exists x\,G$	Negating Quantifier E27
	$\cong \forall x\,F \to \exists x\,G$	Implication E20

2.11 Remark: We now give several examples of proofs which use standard mathematical language and notation. As an exercise, you should try to justify each step in each proof. Some steps make use of a definition, and other steps use one or more of the basic equivalences. In Example 2.15, two of the steps make use of the fact that $\alpha(F), \alpha(G) \in \{0, 1\}$.

We make several remarks about the symbols used in these proofs. Many of the symbols used in our formal symbolic language are not normally used in more standard mathematical language. The symbols $\rightarrow$ and $\leftrightarrow$ are more commonly written as $\Longrightarrow$ and $\Longleftrightarrow$. The symbols $\wedge$ and $\vee$ are more commonly expressed using the words "and" and "or". The negation symbol $\neg$ is usually either expressed in words (using expressions involving the word "not") or is indicated by crossing out a binary relation symbol, for example by writing $\neg\, s{=}t$ as $s \neq t$ and by writing $\neg\, s{\in}t$ as $s \notin t$. Also note that the same symbol $\Longleftrightarrow$ which is used in place of the symbol $\rightarrow$ is also often used to replace the symbol $\cong$.

2.12 Example: Let A and B be sets. Show that $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$.

Solution: We have

$$\begin{aligned} A = B &\Longleftrightarrow \forall x \left(x \in A \Longleftrightarrow x \in B\right) \\ &\Longleftrightarrow \forall x \left((x \in A \Longrightarrow x \in B) \text{ and } (x \in B \Longrightarrow x \in A)\right) \\ &\Longleftrightarrow \forall x \,(x \in A \Longrightarrow x \in B) \text{ and } \forall x (x \in B \Longrightarrow x \in A) \\ &\Longleftrightarrow A \subseteq B \text{ and } B \subseteq A. \end{aligned}$$

2.13 Example: Prove that for all sets A, B and C we have $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (this is part of Theorem 1.4).

Solution: Let A, B and C be sets. Then for all x we have

$$\begin{aligned} x \in A \cap (B \cup C) &\Longleftrightarrow x \in A \text{ and } x \in (B \cup C) \\ &\Longleftrightarrow x \in A \text{ and } \left(x \in B \text{ or } x \in C\right) \\ &\Longleftrightarrow \left(x \in A \text{ and } x \in B\right) \text{ or} \left(x \in A \text{ and } x \in C\right) \\ &\Longleftrightarrow x \in A \cap B \text{ or } x \in A \cap C \\ &\Longleftrightarrow x \in (A \cap B) \cup (A \cap C). \end{aligned}$$

2.14 Example: Prove that for sets $A, B \subseteq X$ we have $X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B)$ (this is also part of Theorem 1.4).

Solution: Let A, B and X be sets with $A, B \subseteq X$. Then for all x we have

$$\begin{aligned} x \in X \setminus (A \cup B) &\Longleftrightarrow x \in X \text{ and } x \notin (A \cup B) \\ &\Longleftrightarrow x \in X \text{ and } \left(x \notin A \text{ and } x \notin B\right) \\ &\Longleftrightarrow \left(x \in X \text{ and } x \in X\right) \text{ and } \left(x \notin A \text{ and } x \notin B\right) \\ &\Longleftrightarrow \left(x \in X \text{ and } x \notin A\right) \text{ and } \left(x \in X \text{ and } x \notin B\right) \\ &\Longleftrightarrow x \in X \setminus A \text{ and } x \in X \setminus B \\ &\Longleftrightarrow x \in (X \setminus A) \cap (X \setminus B). \end{aligned}$$

2.15 Example: Prove that for all formulas F and G, $F \cong G \Longleftrightarrow \left(F \models G \text{ and } G \models F\right)$ (this is part of Theorem 1.20),

Solution: Let F and G be formulas. Then

$$\begin{aligned} \left(F \models G \text{ and } G \models F\right) \Longleftrightarrow &\text{ for all assignments } \alpha \ \left(\alpha(F){=}1 \Longrightarrow \alpha(G){=}1\right) \\ &\text{and for all assignments } \alpha \ \left(\alpha(G){=}1 \Longrightarrow \alpha(F){=}1\right) \end{aligned}$$

$\iff$ for all assignments α $\big(\big(\alpha(F)=1 \Longrightarrow \alpha(G)=1\big)$ and $\big(\alpha(G)=1 \Longrightarrow \alpha(F)=1\big)\big)$

$\iff$ for all assignments α $\big(\alpha(F)=1 \iff \alpha(G)=1\big)$

$\iff$ for all assignments α $\big(\big(\alpha(F)=1$ and $\alpha(G)=1\big)$ or $\big(\alpha(F) \neq 1$ and $\alpha(G) \neq 1\big)\big)$

$\iff$ for all assignments α $\big(\big(\alpha(F)=1$ and $\alpha(G)=1\big)$ or $\big(\alpha(F) = 0$ and $\alpha(G) = 0\big)\big)$

$\iff$ for all assignments α $\alpha(F) = \alpha(G)$

$\iff F \cong G$.

2.16 Definition: For any formulas F, G and H, any sets of formulas $\mathcal{S}$ and $\mathcal{T}$, any terms s and t, and any variable symbols x and y, the following rules are called the **basic validity rules**. We accept these rules axiomatically, without proof.

(Premise)	V1.	If $F \in \mathcal{S}$ then $\mathcal{S} \models F$
(Adding Premises)	V2.	If $\mathcal{S} \models F$ and $\mathcal{S} \subseteq \mathcal{T}$ then $\mathcal{T} \models F$
(The Chain Rule)	V3.	If $\mathcal{S} \models F$ and $\mathcal{S} \cup \{F\} \models G$ then $\mathcal{S} \models G$
(Proof by Cases)	V4.	If $\mathcal{S} \cup \{F\} \models G$ and $\mathcal{S} \cup \{\neg F\} \models G$ then $\mathcal{S} \models G$
(Contradiction)	V5.	If $\mathcal{S} \cup \{\neg F\} \models G$ and $\mathcal{S} \cup \{\neg F\} \models \neg G$ then $\mathcal{S} \models F$
	V6.	If $\mathcal{S} \cup \{F\} \models G$ and $\mathcal{S} \cup \{F\} \models \neg G$ then $\mathcal{S} \models \neg F$
(Conjunction)	V7.	If $\mathcal{S} \models F$ and $\mathcal{S} \models G$ then $\mathcal{S} \models F \wedge G$
	V8.	If $\mathcal{S} \cup \{F, G\} \models H$ then $\mathcal{S} \cup \{F \wedge G\} \models H$
	V9.	If $\mathcal{S} \models F \wedge G$ then $\mathcal{S} \models F$
	V10.	If $\mathcal{S} \models F \wedge G$ then $\mathcal{S} \models G$
(Disjunction)	V11.	If $\mathcal{S} \cup \{\neg F\} \models G$ then $\mathcal{S} \models F \vee G$
	V12.	If $\mathcal{S} \cup \{\neg G\} \models F$ then $\mathcal{S} \models F \vee G$
	V13.	If $\mathcal{S} \cup \{F\} \models H$ and $\mathcal{S} \cup \{G\} \models H$ then $\mathcal{S} \cup \{F \vee G\} \models H$
	V14.	If $\mathcal{S} \models F$ then $\mathcal{S} \models F \vee G$
	V15.	If $\mathcal{S} \models G$ then $\mathcal{S} \models F \vee G$
	V16.	If $\mathcal{S} \models F \vee G$ and $\mathcal{S} \models \neg F$ then $\mathcal{S} \models G$
	V17.	If $\mathcal{S} \models F \vee G$ and $\mathcal{S} \models \neg G$ then $\mathcal{S} \models F$
(Implication)	V18.	If $\mathcal{S} \cup \{F\} \models G$ then $\mathcal{S} \models F \to G$
	V19.	If $\mathcal{S} \cup \{\neg G\} \models \neg F$ then $\mathcal{S} \models F \to G$
	V20.	If $\mathcal{S} \cup \{\neg F\} \models H$ and $\mathcal{S} \cup \{G\} \models H$ then $\mathcal{S} \cup \{F \to G\} \models H$
	V21.	If $\mathcal{S} \models \neg F$ then $\mathcal{S} \models F \to G$
	V22.	If $\mathcal{S} \models G$ then $\mathcal{S} \models F \to G$
	V23.	If $\mathcal{S} \models F \to G$ and $\mathcal{S} \models F$ then $\mathcal{S} \models G$
	V24.	If $\mathcal{S} \models F \to G$ and $\mathcal{S} \models \neg G$ then $\mathcal{S} \models \neg F$
(If and Only If)	V25.	If $\mathcal{S} \models F \to G$ and $\mathcal{S} \models G \to F$ then $\mathcal{S} \models F \leftrightarrow G$
	V26.	If $\mathcal{S} \cup \{F, G\} \models H$ and $\mathcal{S} \cup \{\neg F, \neg G\} \models H$ then $\mathcal{S} \cup \{F \leftrightarrow G\} \models H$
	V27.	If $\mathcal{S} \models F$ and $\mathcal{S} \models G$ then $\mathcal{S} \models F \leftrightarrow G$
	V28.	If $\mathcal{S} \models \neg F$ and $\mathcal{S} \models \neg G$ then $\mathcal{S} \models F \leftrightarrow G$
	V29.	If $\mathcal{S} \models F \leftrightarrow G$ and $\mathcal{S} \models F$ then $\mathcal{S} \models G$
	V30.	If $\mathcal{S} \models F \leftrightarrow G$ and $\mathcal{S} \models G$ then $\mathcal{S} \models F$
	V31.	If $\mathcal{S} \models F \leftrightarrow G$ and $\mathcal{S} \models \neg F$ then $\mathcal{S} \models \neg G$
	V32.	If $\mathcal{S} \models F \leftrightarrow G$ and $\mathcal{S} \models \neg G$ then $\mathcal{S} \models \neg F$

(Equality) V33 $\mathcal{S} \models t = t$
V34 If $\mathcal{S} \models s = t$ then $\mathcal{S} \models t = s$
V35 If $\mathcal{S} \models r = s$ and $\mathcal{S} \models s = t$ then $\mathcal{S} \models r = t$
V36 If $S \models s = t$ and $\mathcal{S} \models [F]_{x \mapsto s}$ then $\mathcal{S} \models [F]_{x \mapsto t}$

(Forall) V37 If $S \models [F]_{x \mapsto y}$ where y is not free in $\mathcal{S} \cup \{\forall x\, F\}$, then $\mathcal{S} \models \forall x\, F$
V38 If $\mathcal{S} \cup \left\{[F]_{x \mapsto t}\right\} \models G$ then $\mathcal{S} \cup \{\forall x\, F\} \models G$
V39 If $\mathcal{S} \models \forall x\, F$ then $\mathcal{S} \models [F]_{x \mapsto t}$

(Exists) V40. If $\mathcal{S} \models [F]_{x \mapsto t}$ then $\mathcal{S} \models \exists x\, F$
V41. If $\mathcal{S} \cup \left\{[F]_{x \mapsto y}\right\} \models G$ where y is not free in $\mathcal{S} \cup \{G, \exists x\, F\}$, then $\mathcal{S} \cup \{\exists x\, F\} \models G$

(Equivalence) V42. If $F \cong G$ and $\mathcal{S} \models F$ then $\mathcal{S} \models G$
V43. If $F \cong G$ and $\mathcal{S} \cup \{F\} \models H$ then $\mathcal{S} \cup \{G\} \models H$

Rule V13 is also called **Proof by Cases**, Rule V19 is called the **Contrapositive Rule**, Rule V23 is called **Modus Ponens**, and rule V36 is called the **Substitution Rule**.

2.17 Definition: Given a set of formulas $\mathcal{S}$ and a formula F such that $\mathcal{S} \models F$, a **derivation** of the valid argument $\mathcal{S} \models F$ is a list of valid arguments $\mathcal{S}_1 \models F_1$, $\mathcal{S}_2 \models F_2$, $\cdots$, $\mathcal{S}_l \models F_l$ with $\mathcal{S}_l = \mathcal{S}$ and $F_l = F$, such that each valid argument $\mathcal{S}_k \models F_k$ is obtained from previous valid arguments $\mathcal{S}_j \models F_j$ with $j < k$ using one of the Basic Validity Rules. The equivalence rules V42 and V43 are only used in the case that the equivalence $F \cong G$ is obtained by applying one of the 35 basic equivalences. Except for the equivalence rules, the Basic Validity Rules are not applied to subformulas.

2.18 Note: The basic validity rules correspond to standard methods of proof which are used routinely in mathematics. Here are the basic validity rules stated less formally (and less precisely) in standard mathematical language.

V1 (Premise) If we suppose F then we can conclude F.
V2 (Adding Premises) If we can prove G without F then we can prove G with F.
V3 (Chain Rule) If we can prove F and, with F we can prove G, then we can prove G.
V4 (Proof by Cases) To prove F by cases, choose a formula G, then consider two cases. For the first case, suppose that G is true then prove F and, for the second case, suppose that G is false then prove F.

V5 (Contradiction 1) To prove that F is true we can suppose, for a contradiction, that F is false, choose a formula G, then prove that G is true and that G is false.
V6 (Contradiction 2) To prove that F is false we can suppose, for a contradiction, that F is true, choose a formula G, then prove that G is true and that G is false.

V7 (Conjunction 1) To prove $F \wedge G$, we prove F and we prove G.
V8 (Conjunction 2) To prove that $F \wedge G$ implies H we suppose F and G then prove H.
V9 (Conjunction 3) From $F \wedge G$ we can conclude F.
V10 (Conjunction 4) From $F \wedge G$ we can conclude G.

V11 (Disjunction 1) To prove $F \vee G$ we can suppose that F is false then prove G.
V12 (Disjunction 2) To prove $F \vee G$ we can suppose that G is false then prove F.
V13 (Disjunction 3) To prove that $F \vee G$ implies H we consider two cases. For the first case, we suppose F then prove H and, for the second case, we suppose G then prove H.
V14 (Disjunction 4) From F we can conclude $F \vee G$.
V15 (Disjunction 5) From G we can conclude $F \vee G$.
V16 (Disjunction 6) From $F \vee G$ and $\neg F$ we can conclude G.
V17 (Disjunction 7) From $F \vee G$ and $\neg G$ we can conclude F.

V18 (Implication 1) To prove $F \to G$ we can suppose F then prove G.
V19 (Implication 2) To prove $F \to G$ we can suppose $\neg G$ then prove $\neg F$.
V20 (Implication 3) To prove that $F \to G$ implies H, we consider two cases. For the first case, suppose $\neg F$ then prove H and, for the second case, suppose G then prove H.
V21 (Implication 4) From $\neg F$ we can conclude $F \to G$.
V22 (Implication 5) From G we can conclude $F \to G$.
V23 (Implication 6) From $F \to G$ and F we can conclude G.
V24 (Implication 7) From $F \to G$ and $\neg G$ we can conclude $\neg F$.

V25 (If and Only If 1) To prove $F \leftrightarrow G$, we prove $F \to G$ and we prove $G \to F$
V26 (If and Only If 2) To prove that $F \leftrightarrow G$ implies H, first we suppose that F and G are both true then prove H, and then we suppose that F and G are both false then prove H.
V27 (If and Only If 3) From F and G we can conclude $F \leftrightarrow G$.
V28 (If and Only If 4) From $\neg F$ and $\neg G$ we can conclude $F \leftrightarrow G$.
V29 (If and Only If 5) From $F \leftrightarrow G$ and F we can conclude G.
V30 (If and Only If 6) From $F \leftrightarrow G$ and G we can conclude F.
V31 (If and Only If 7) From $F \leftrightarrow G$ and $\neg F$ we can conclude $\neg G$.
V32 (If and Only If 8) From $F \leftrightarrow G$ and $\neg G$ we can conclude $\neg F$.
V33 (Equality 1) We can always conclude that $t = t$.
V34 (Equality 2) From $s = t$ we can conclude that $t = s$.
V35 (Equality 3) From $r = s$ and $s = t$ we can conclude that $r = t$.
V36 (Substitution) From $s = t$ and $[F]_{x \mapsto s}$ we can conclude $[F]_{x \mapsto t}$.

V37 (Forall 1) To prove $\forall x\, F$, we choose a variable symbol y about which we have not made any assumptions (in the case that we have not made any assumptions about x we can take $y \equiv x$) and we write "let y be arbitrary", then we prove the statement $[F]_{x \mapsto y}$.
V38 (Forall 2) To prove that $\forall x\, F$ implies G, we choose a term t, suppose that $[F]_{x \mapsto t}$ is true, then prove G.
V39 (Forall 3) From $\forall x F$ we can conclude $[F]_{x \mapsto t}$.

V40 (Exists 1) To prove $\exists x\, F$, we choose a term t then we prove the statement $[F]_{x \mapsto t}$.
V41 (Exists 2) To prove that $\exists x\, F$ implies G, we choose a variable symbol y about which we have made no assumptions and which does not occur in G (in the case that we have not made any assumptions about x and x does not occur in G we can take $y \equiv x$), we suppose $[F]_{x \mapsto y}$ and write "choose y so that $[F]_{x \mapsto y}$ is true", then we prove G.

V42 (Equivalence 1) If F is equivalent to G, then to prove F we can prove G.
V43 (Equivalence 2) If F is equivalent to G, then we can replace the premise F by G.

2.19 Note: Recall that the statement $\forall\, x \in A\ F$ can be expressed as $\forall x\big(x \in A \to F\big)$. To prove this statement, in the case that we have made no assumptions about x, we write "let x be arbitrary" [to use V37] then we suppose $x \in A$ [to use V18] then we prove F (in the case that we have made assumptions about x, we write "let y be arbitrary", suppose $y \in A$, then prove $[F]_{x \mapsto y}$). Rather than writing "let x be arbitrary and suppose $x \in A$" we usually write "let $x \in A$ be arbitrary" or simply "let $x \in A$". Similarly, to prove a statement of the form "for every function $f : A \to B$ we have F" we would begin the proof by writing "let $f : A \to B$ be arbitrary" or simply "let $f : A \to B$".

2.20 Note: In standard mathematical proofs, the proof rules are often used implicitly, but in the next few examples we shall state explicitly (in square brackets) which rule is being used at each step in our proof.

2.21 Example: Let F, G and H be formulas. Prove that $\{F \to (G \wedge H), (F \wedge G) \vee H\} \models H$.

Solution: We need to prove that for every assignment α, if $\alpha\big(F \to (G \wedge H)\big) = 1$ and $\alpha\big((F \wedge G) \vee H\big) = 1$ then $\alpha(H) = 1$. Here is a step-by-step proof in which we indicate which proof rule is being used at each step.

1. Let α be an arbitrary assignment [to use V37 and V18].
2. Suppose that $F \to (G \wedge H)$ is true (under α), and that $(F \wedge G) \vee H$ is true [to use V8].
3. Suppose, for a contradiction, that H is false [to use V5].
4. Since $(F \wedge G) \vee H$ is true and H is false, it follows that $F \wedge G$ is true [by V17].
5. Since $F \wedge G$ is true, F is true [by V9].
6. Since F is true and $F \to (G \wedge H)$ is true, it follows that $G \wedge H$ is true [by V23].
7. Since $G \wedge H$ is true, H is true [by V10].
8. Since H true and H false, we have the desired contradiction. Thus H is true [by V5].
9. Thus if $\alpha\big(F \to (G \wedge H)\big) = 1$ and $\alpha\big((F \wedge G) \vee H\big) = 1$ then $\alpha(H) = 1$ [by V8].
10. Since α was arbitrary, we have $\{F \to (G \wedge H), (F \wedge G) \vee H\} \models H$ [by V37 and V18].

Now here is the same proof presented in the form of a derivation of valid arguments.

1.	$\{F \to (G \wedge H), (F \wedge G) \vee H\,,\, \neg H\} \models \neg H$	V1
2.	$\{F \to (G \wedge H), (F \wedge G) \vee H\,,\, \neg H\} \models (F \wedge G) \vee H$	V1
3.	$\{F \to (G \wedge H), (F \wedge G) \vee H\,,\, \neg H\} \models (F \wedge G)$	V17 on lines 2 and 1
4.	$\{F \to (G \wedge H), (F \wedge G) \vee H\,,\, \neg H\} \models F$	V9 on 3
5.	$\{F \to (G \wedge H), (F \wedge G) \vee H\,,\, \neg H\} \models F \to (G \wedge H)$	V1
6.	$\{F \to (G \wedge H), (F \wedge G) \vee H\,,\, \neg H\} \models G \wedge H$	V23 on 5 and 4
7.	$\{F \to (G \wedge H), (F \wedge G) \vee H\,,\, \neg H\} \models H$	V10 on 6
8.	$\{F \to (G \wedge H), (F \wedge G) \vee H\} \models H$	V5 on 7 and 1

2.22 Example: Prove that $\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\} \models H \wedge \neg F$.

Solution: Here is a derivation of valid arguments.

1.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\,,\, \neg H\} \models \neg H$	V1
2.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\,,\, \neg H\} \models \neg(F \vee \neg G)$	V1
3.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\,,\, \neg H\} \models (F \vee \neg G) \to H$	V23 on 1 and 2
4.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\,,\, \neg H\} \models \neg F \wedge \neg\neg G$	V42 with E8 on 3
5.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\,,\, \neg H\} \models \neg F$	V9 on 4
6.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\,,\, \neg H\} \models \neg\neg G$	V10 on 4
7.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\,,\, \neg H\} \models G$	V42 with E2 on 6
8.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\,,\, \neg H\} \models G \wedge \neg H$	V7 on 7 and 1
9.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\,,\, \neg H\} \models F \leftrightarrow (G \wedge \neg H)$	V1
10.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\,,\, \neg H\} \models F$	V30 on 9
11.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\} \models H$	V5 on 10 and 5
12.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\} \models \neg\neg H$	V42 with E2 on 11
13.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\} \models \neg G \vee \neg\neg H$	V15 on 12
14.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\} \models \neg(G \wedge \neg H)$	V42 with E7 on 13
15.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\} \models F \leftrightarrow (G \wedge \neg H)$	V1 (or V2 on 9)
16.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\} \models H$	V5 on 10 and 5
17.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\} \models \neg F$	V32 on 16
18.	$\{((F \vee \neg G) \to H)\,,\, (F \leftrightarrow (G \wedge \neg H))\} \models H$	V7 on 11 and 17

2.23 Example: Prove that $\models \forall x\, \exists y\, \big(\neg y = f(x) \to yRf(x)\big)$.

Solution: We need to prove that for every non-empty set U, for every function $f : U \to U$, and for every binary relation relation $R \subseteq U^2$, for all $x \in U$ there exists $y \in U$ such that if $y \neq f(x)$ then $(y, f(x)) \in R$. Here is a step-by-step proof.

1. Let U be a set, let $f : U \to U$ and let $R \subseteq U^2$ [to use V37 and V18].
2. Let $x \in U$ be arbitrary [to use V37].
3. Choose $y = f(x)$ [to use V40 where the chosen term is $t = f(x)$].
4. Since $y = f(x)$, the formula $\neg y = f(x)$ is false. [by V42 with E2]
5. Since $\neg y = f(x)$ is false, the statement $\neg y = f(x) \to (y, f(x)) \in R$ is true [by V21].
6. This proves that there exists $y \in U$ such that if $y \neq f(x)$ then $(y, f(x)) \in R$ [by V40].
7. Since $x \in U$ was arbitrary, we have proven that for all $x \in U$ there exists $y \in U$ such that if $y \neq f(x)$ then $(y, f(x)) \in R$ [by V37].
8. Since U and f and R were arbitrary, our proof is complete [by V37 and V18].

Here is the same proof presented formally as a derivation of valid arguments.

1.	$\models f(x) = f(x)$	V33
2.	$\models \neg\neg\, f(x) = f(x)$	V42 with E2 on 1
3.	$\models \big(\neg\, f(x) = f(x) \to f(x)Rf(x)\big)$	V24 on 2
4.	$\models \exists y\big(\neg\, y = f(x) \to yRf(x)\big)$	V40 with $t = f(x)$ on 3
5.	$\models \forall x\, \exists y\, \big(\neg y = f(x) \to yRf(x)\big)$	V37 on 4

2.24 Exercise: Prove that $\models \forall x\big(\exists y\ \neg xRy \vee \exists y\ yRx\big)$.

2.25 Example: Prove that $\big\{\forall x\ g(x, a) = x\big\} \models \forall x\, \big(\forall y\ g(x, y) = y \to x = a\big)$.

Solution: First we provide a proof using standard mathematical language. Let U be a nonempty set, let $a \in U$ and let $g : U^2 \to U$. Suppose that for all $x \in U$ we have $g(x, a) = x$. We need to prove that for all $x \in U$, if $g(x, y) = y$ for every $y \in U$ then $x = a$. Let $x \in U$ be arbitrary, and note that $g(x, a) = x$. Suppose that for every $y \in U$ we have $g(x, y) = y$. Then in particular, taking $y = a$, we have $g(x, a) = a$. Thus $x = g(x, a) = a$, as required.

We now convert the above proof into a derivation of valid arguments:

1.	$\big\{\forall x\ g(x, a) = x\,,\ \forall y\ g(x, y) = y\big\} \models\ \forall y\ g(x, y) = y$	V1
2.	$\big\{\forall x\ g(x, a) = x\,,\ \forall y\ g(x, y) = y\big\} \models\ g(x, a) = a$	V39 on 1
3.	$\big\{\forall x\ g(x, a) = x\,,\ \forall y\ g(x, y) = y\big\} \models\ \forall x\ g(x, a) = x$	V1
4.	$\big\{\forall x\ g(x, a) = x\,,\ \forall y\ g(x, y) = y\big\} \models\ g(x, a) = x$	V39 on 3
5.	$\big\{\forall x\ g(x, a) = x\,,\ \forall y\ g(x, y) = y\big\} \models\ x = g(x, a)$	V34 on 4
6.	$\big\{\forall x\ g(x, a) = x\,,\ \forall y\ g(x, y) = y\big\} \models\ x = a$	V35 on 5, 2
7.	$\big\{\forall x\ g(x, a) = x\big\} \models\ (\forall y\, g(x, y) = y \to x = a)$	V18 on 6
8.	$\big\{\forall x\ g(x, a) = x\big\} \models\ \forall x\, (\forall y\ g(x, y) = y \to x = a)$	V37 on 7

Note that, at the final step, we were able to apply Rule V37 on line 7 because the variable symbol x is not free in the formula $\forall x\ g(x, a) = x$.

2.26 Note: Any statement of the form $\forall\, x \in \emptyset\ F$ is true. Indeed the statement $\forall\, x \in \emptyset\ F$ is equivalent (by definition) to the statement $\forall x(x \in \emptyset \to F)$. For every x, the statement $x \in \emptyset$ is false, and so the statement $x \in \emptyset \to F$ is true. A statement of this form is said to be **vacuously true**.

2.27 Exercise: Let F be a formula and let the symbol $\emptyset$ be a constant symbol. Make a derivation of valid arguments to show that $\{\forall x \neg x \in \emptyset\} \models \forall x (x \in \emptyset \to F)$.

2.28 Exercise: Let F be a formula and let $\emptyset$ be the empty set. Prove that $\models F \iff \emptyset \models F$.

2.29 Example: Prove that the class of all sets is not a set.

Solution: Here is a proof in standard mathematical language.

1. Let u be the class of all sets.
2. Suppose, for a contradiction, that u is a set .
3. Let $w = \{x \in u | x \notin x\}$ and note that w is a set by a Separation Axiom.
4. We claim that $w \in w$. Suppose, for a contradiction, that $w \notin w$.
5. Since $w \in u$ and $w \notin w$ we have $w \in w$ by the definition of w.
6. Since $w \in w$ and $w \notin w$ we have the desired contradiction, so $w \in w$, as claimed.
7. We claim that $w \notin w$. Suppose, for a contradiction, that $w \in w$.
8. Since $w \in u$ and $w \in w$ we have $w \notin w$, by the definition of w.
9. Since $w \in w$ and $w \notin w$ we have the desired contradiction, so $w \notin w$, as claimed.
10. Since $w \in w$ and $w \notin w$, we have the desired contradiction, so u is not a set, as claimed.

Note that the statement "the class u of all sets is a set" can be expressed as $\exists u \forall x \; x \in u$. Also, note that on line 3 we used the Separation Axiom $\forall u \exists w \forall x \left(x \in w \leftrightarrow (x \in u \wedge \neg x \in x\right)$. The above proof can be converted into a derivation of the valid argument

$$\left\{\forall u \exists w \forall x \left(x \in w \leftrightarrow (x \in u \wedge \neg x \in x)\right)\right\} \models \neg \exists u \forall x \; x \in u \, .$$

Here is a derivation which is a bit similar to the above proof.

1.	$\left\{w \in w \leftrightarrow (w \in u \wedge \neg w \in w) \, , \, w \in u \, , \, w \in w\right\} \models w \in w$	V1
2.	$\left\{w \in w \leftrightarrow (w \in u \wedge \neg w \in w) \, , \, w \in u \, , \, w \in w\right\} \models w \in w \leftrightarrow (w \in u \wedge \neg w \in w)$	V1
3.	$\left\{w \in w \leftrightarrow (w \in u \wedge \neg w \in w) \, , \, w \in u \, , \, w \in w\right\} \models (w \in u \wedge \neg w \in w)$	V29
4.	$\left\{w \in w \leftrightarrow (w \in u \wedge \neg w \in w) \, , \, w \in u \, , \, w \in w\right\} \models \neg w \in w$	V12
5.	$\left\{w \in w \leftrightarrow (w \in u \wedge \neg w \in w) \, , \, w \in u\right\} \models \neg w \in w$	V6
6.	$\left\{w \in w \leftrightarrow (w \in u \wedge \neg w \in w) \, , \, w \in u\right\} \models w \in u$	V1
7.	$\left\{w \in w \leftrightarrow (w \in u \wedge \neg w \in w) \, , \, w \in u\right\} \models w \in u \wedge \neg w \in w$	V10
8.	$\left\{w \in w \leftrightarrow (w \in u \wedge \neg w \in w) \, , \, w \in u\right\} \models w \in w \leftrightarrow (w \in u \wedge \neg w \in w)$	V1
9.	$\left\{w \in w \leftrightarrow (w \in u \wedge \neg w \in w) \, , \, w \in u\right\} \models w \in w$	V30
10.	$\left\{w \in w \leftrightarrow (w \in u \wedge \neg w \in w)\right\} \models \neg w \in u$	V6
11.	$\left\{w \in w \leftrightarrow (w \in u \wedge \neg w \in w)\right\} \models \exists x \neg x \in u$	V40
12.	$\left\{\forall x \left(x \in w \leftrightarrow (x \in u \wedge \neg x \in x)\right)\right\} \models \exists x \neg x \in u$	V38
13.	$\left\{\exists w \forall x \left(x \in w \leftrightarrow (x \in u \wedge \neg x \in x)\right)\right\} \models \exists x \neg x \in u$	V41
14.	$\left\{\forall u \exists w \forall x \left(x \in w \leftrightarrow (x \in u \wedge \neg x \in x)\right)\right\} \models \exists x \neg x \in u$	V39
15.	$\left\{\forall u \exists w \forall x \left(x \in w \leftrightarrow (x \in u \wedge \neg x \in x)\right)\right\} \models \forall u \exists x \neg x \in u$	V37
16.	$\left\{\forall u \exists w \forall x \left(x \in w \leftrightarrow (x \in u \wedge \neg x \in x)\right)\right\} \models \forall u \neg \forall x \, x \in u$	V45, E28
17.	$\left\{\forall u \exists w \forall x \left(x \in w \leftrightarrow (x \in u \wedge \neg x \in x)\right)\right\} \models \neg \exists u \forall x \, x \in u$	V45, E29

2.30 Example: For $a, b \in \mathbf{Z}$ we write $a \big| b$ when a is a factor of b, Prove that for $a, b, c \in \mathbf{Z}$, if $a|b$ and $b|c$ then $a|c$.

Solution: Here is a proof, in standard mathematical language, in which we do not bother to explicitly list all of our assumptions and we do not bother to indicate which proof rules are being used.

1. Suppose that $a|b$ and that $b|c$.
2. Since $a|b$ we can choose $k \in \mathbf{Z}$ so that $b = ak$.
3. Since $b|c$ we can choose $l \in \mathbf{Z}$ so that $c = bl$.
4. Then we have $c = bl = (ak)l = a(kl)$.
5. Thus $a|c$.

Note that $a|b$ can be expressed as $\exists x\ b = a \times x$ and $b|c$ can be expressed as $\exists x\ c = b \times x$ and $a|c$ can be expressed as $\exists x\ c = a \times x$. Also notice that on line 4 of the above proof, we implicitly made use of the fact that multiplication is associative which can be expressed as $\forall x\, \forall y\, \forall z\ (x \times y) \times z = x \times (y \times z)$. We now translate the above standard mathematical proof into a more detailed step-by-step proof which shows that

$$\big\{\exists x\ b = (a \times x)\,,\ \exists x\ c = b \times x\,,\ \forall x\, \forall y\, \forall z\ ((x \times y) \times z) = (x \times (y \times z))\big\} \models \exists x\ c = a \times x\,.$$

We need to prove that for every non-empty set U, for every binary function $\times$, and for every choice of $a \in U$ and $b \in U$, if there exists $x \in U$ such that $b = a \times x$ and if there exists $x \in U$ such that $c = b \times x$ and if for all $x, y, z \in U$ we have $(x \times y) \times z = x \times (y \times z)$, then there exists $x \in U$ such that $c = a \times x$. Here is a proof.

1. Let U be a set, let $\times$ be a binary function on U, and let $a, b \in U$ [to use V37 and V18].
2. Suppose $\exists\, x\ b = a \times x$, suppose $\exists\, x\ c = b \times x$, and suppose $\forall x\, \forall y\, \forall z\ (x \times y) \times z = x \times (y \times z)$ [to use V8 and V18].
3. Since $\exists x\ b = a \times x$ we can choose x so that $b = a \times x$ [to use V41].
4. Since $\exists x\ c = b \times x$ we can choose y so that $c = b \times y$ [to use V41].
5. Since $c = b \times y$ and $b = a \times x$ it follows that $c = (a \times x) \times y$ [by V36].
6. Since $\forall x\, \forall y\, \forall z\ (x \times y) \times z = x \times (y \times z)$, we have $\forall y\, \forall z\ (a \times y) \times z = a \times (y \times z)$ [by V39].
7. Since $\forall y\, \forall z\ (a \times y) \times z = a \times (y \times z)$, it follows that $\forall z\ (a \times x) \times z = a \times (x \times z)$ [by V39].
8. Since $\forall z\ (a \times x) \times z = a \times (x \times z)$ it follows that $(a \times x) \times y = a \times (x \times y)$ [by V39].
9. Since $c = (a \times x) \times y$ and $(a \times x) \times y = a \times (x \times y)$, it follows that $c = a \times (x \times y)$ [by V35].
10. Since $c = a \times (x \times y)$ it follows that $\exists x\ c = a \times x$ [by V40].
11. Since U and $\times$ and a and b were arbitrary, the proof is complete [by V37, V8 and V18].

Here is a similar proof presented formally as a derivation of valid arguments.

1.	$\big\{b = a \times x\,,\ c = b \times y\,,\ (a \times x) \times y = a \times (x \times y)\big\} \models\ b = a \times x$	V1
2.	$\big\{b = a \times x\,,\ c = b \times y\,,\ (a \times x) \times y = a \times (x \times y)\big\} \models\ c = b \times y$	V1
3.	$\big\{b = a \times x\,,\ c = b \times y\,,\ (a \times x) \times y = a \times (x \times y)\big\} \models\ c = (a \times x) \times y$	V36
4.	$\big\{b = a \times x\,,\ c = b \times y\,,\ (a \times x) \times y = a \times (x \times y)\big\} \models\ (a \times x) \times y = a \times (x \times y)$	V1
5.	$\big\{b = a \times x\,,\ c = b \times y\,,\ (a \times x) \times y = a \times (x \times y)\big\} \models\ c = a \times (x \times y)$	V35
6.	$\big\{b = a \times x\,,\ c = b \times y\,,\ (a \times x) \times y = a \times (x \times y)\big\} \models\ \exists x\ c = a \times x$	V40
7.	$\big\{b = a \times x\,,\ c = b \times y\,,\ \forall z\ (a \times x) \times z = a \times (x \times z)\big\} \models\ \exists x\ c = a \times x$	V38
8.	$\big\{b = a \times x\,,\ c = b \times y\,,\ \forall y\, \forall z\ (a \times y) \times z = a \times (y \times z)\big\} \models\ \exists x\ c = a \times x$	V38
9.	$\big\{b = a \times x\,,\ c = b \times y\,,\ \forall x\, \forall y\, \forall z\ (x \times y) \times z = x \times (y \times z)\big\} \models\ \exists x\ c = a \times x$	V38
10.	$\big\{b = a \times x\,,\ \exists x\ c = b \times x\,,\ \forall x\, \forall y\, \forall z\ (x \times y) \times z = x \times (y \times z)\big\} \models\ \exists x\ c = a \times x$	V41
11.	$\big\{\exists x\ b = a \times x\,,\ \exists x\ c = b \times x\,,\ \forall x\, \forall y\, \forall z\ (x \times y) \times z = x \times (y \times z)\big\} \models\ \exists x\ c = a \times x$	V41

2.31 Example: Prove that for $a, b \in \mathbf{Z}$, if $a|b$ and $b|a$ then $b = \pm a$.

Proof: Here is a proof in standard mathematical language.

1. Suppose that $a|b$ and $b|a$.
2. Since $a|b$ we can choose $k \in \mathbf{Z}$ so that $b = ak$.
3. Since $b|a$ we can choose $l \in \mathbf{Z}$ so that $a = bl$.
4. Then we have $a = bl = (ak)l = a(kl)$, that is $a \cdot 1 = a \cdot kl$.
5. Since $a \cdot 1 = a \cdot kl$, it follows that either $a = 0$ or $1 = kl$.
6. In the case that $a = 0$ we have $b = a \cdot k = 0 \cdot k = 0$ and so $b = a$.
7. In the case that $1 = kl$, it follows that either $k = l = 1$ or $k = l = -1$.
8. When $k = l = 1$ we have $b = a \cdot k = a \cdot 1 = a$ and
9. when $k = l = -1$ we have $b = a \cdot k = a \cdot (-1) = -a$.
10. In all cases, either $b = a$ or $b = -a$, that is $b = \pm a$.

It is a bit challenging to convert the above proof into a derivation of valid arguments, not because the proof itself is particularly difficult, but because the proof makes use of many algebraic properties of the integers and, because we are so familiar with those algebraic properties, it is easy to overlook the fact that some of these properties need to be included as premises in order to obtain a valid argument. Also it is difficult to decide exactly which properties need to be included as premises and which properties can then be proven from those premises. Here are some of the properties that were used implicitly in the proof.

On line 4 we used associativity of multiplication to obtain $(ak)l = a(kl)$ and we used the fact that $a = a \cdot 1$. On line 5 we used the fact that if $au = av$ then either $a = 0$ or $u = v$. On line 6 we used the fact that $0 \cdot k = 0$. On line 7 we used the fact that if $kl = 1$ then either $k = l = 1$ or $k = l = -1$ (incidentally, this algebraic property does not hold in $\mathbf{Q}$ or in $\mathbf{R}$). On line 8 we used the fact that $a \cdot 1 = a$ (which was also used on line 5) and on line 9 we used the fact that $a \cdot (-1) = -a$.

Another slight complication is that, in a derivation of valid arguments, all of the statements must be expressed as formulas in a first-order language, say first-order number theory. On several lines in our proof we use some mathematical notation, with which all students will no doubt be familiar, but which is not used explicitly in first order number theory. Namely, we use the negative sign $-$ to write -1 and $-a$. The statement $b = -a$ can be expressed by the formula $b + a = 0$, and the statement $k = l = -1$ can be expressed as $\big(k = l \ \wedge \ l + 1 = 0\big)$, but it is more challenging to decide how to express the statement $a \cdot (-1) = -a$ as a formula; since our proof uses the fact that if $k = -1$ then $a \cdot k = -a$, we might choose to express the statement using the formula $\big(k + 1 = 0 \ \rightarrow \ a \times k + a = 0\big)$.

2.32 Remark: In the next chapter, we shall carefully gather together and list all of the basic algebraic properties of $\mathbf{Z}$, $\mathbf{Q}$ and $\mathbf{R}$ which are needed in all of the subsequent proofs in this course. They are also needed in all proofs in all other mathematics courses. These algebraic properties can either be accepted axiomatically, without proof, or they can all be painstakingly proven. Our approach will be to accept them axiomatically.

In order to prove all of the basic algebraic properties, it would be necessary to begin by carefully and precisely defining the sets $\mathbf{Z}$, $\mathbf{Q}$ and $\mathbf{R}$ (by constructing them explicitly using the ZFC axioms of set theory), and then also carefully and precisely defining all of the algebraic operations, such as addition and multiplication, which are used in these sets. Only after the operations have been defined is it possible to prove that they satisfy their well-known algebraic properties. The procedure by which one can define the sets $\mathbf{Z}$, $\mathbf{Q}$ and $\mathbf{R}$ along with their operations $+$ and $\times$, and also their ordering $\leq$, is outlined briefly in Appendix 1.

.

Chapter 3. Rings, Fields, and Orders

3.1 Remark: In this chapter (and the next) we shall gather together and list all of the basic algebraic properties which hold in $\mathbf{Z}$, $\mathbf{Q}$ and $\mathbf{R}$ which are needed in mathematical proofs. In addition to the sets $\mathbf{Z}$, $\mathbf{Q}$ and $\mathbf{R}$, there are many other algebraic systems (involving sets of mathematical objects along with operations, such as addition and multiplication, which act on these objects) which are studied and used in mathematics. For example, we can add and multiply two functions together, or we can add and multiply two matrices together. Some of the algebraic properties which hold in $\mathbf{Z}$ or in $\mathbf{R}$ also hold in some of these other algebraic systems. Rings and fields, as defined below, are algebraic systems with operations which satisfy some familiar algebraic properties.

3.2 Definition: A **ring** (with identity) is a set R with distinct elements $0, 1 \in R$, called the **zero** and **identity** elements, and binary operations $+, \times : R^2 \to R$, called **addition** and **multiplication**, where for $a, b \in R$ we write $+(a, b)$ as $a + b$ and we write $\times(a, b)$ as $a \times b$ or $a \cdot b$ or ab, such that

R1. $+$ is associative: for all $a, b, c \in R$ we have $(a + b) + c = a + (b + c)$,
R2. $+$ is commutative: for all $a, b \in R$ we have $a + b = b + a$,
R3. 0 is an additive identity: for all $a \in R$ we have $a + 0 = a$,
R4. every $a \in R$ has an additive inverse: for all $a \in R$ there exists $b \in R$ such that $a + b = 0$,
R5. $\times$ is associative: for all $a, b, c \in R$ we have $(ab)c = a(bc)$,
R6. 1 is a multiplicative identity: for all $a \in R$ we have $a \cdot 1 = a = 1 \cdot a$, and
R7. $\times$ is distributive over $+$: for all $a, b, c \in R$ we have $a(b+c) = ab+ac$ and $(a+b)c = ac+bc$.

A ring R is called **commutative** when

R8. $\times$ is commutative: for all $a, b \in R$ we have $ab = ba$.

A **field** is a commutative ring R in which

R9. every $0 \neq a \in R$ has an inverse: for all $0 \neq a \in R$ there exists $b \in R$ such that $ab = 1$.

3.3 Theorem: $\mathbf{Z}$ *is a commutative ring and* $\mathbf{Q}$ *and* $\mathbf{R}$ *are fields with* $\mathbf{Z} \subseteq \mathbf{Q} \subseteq \mathbf{R}$.

Proof: We accept this axiomatically, without proof.

3.4 Exercise: Show that the set $\mathbf{Z}[\sqrt{2}] = \left\{a + b\sqrt{2} \,\middle|\, a, b \in \mathbf{Z}\right\}$ is a commutative ring and that the set $\mathbf{Q}[\sqrt{2}] = \left\{a + b\sqrt{2} \,\middle|\, a, b \in \mathbf{Q}\right\}$ is a field.

3.5 Example: Let n be a positive integer. There is a ring, denoted by $\mathbf{Z}_n$, which is called the ring of **integers modulo** n. Later, we shall define $\mathbf{Z}_n$ precisely but, for now, we provide an informal description. We let $\mathbf{Z}_n = \{0, 1, 2, \cdots, n-1\}$ and, for $a, b \in \mathbf{Z}_n$, we define the sum $a + b \in \mathbf{Z}_n$ to be the remainder when the integer $a + b \in \mathbf{Z}$ is divided by n, and we define the product $ab \in \mathbf{Z}_n$ to be the remainder when the integer ab is divided by n. For example, in $\mathbf{Z}_6$ addition and multiplication are given by the following tables.

$+$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

$\times$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

3.6 Example: There is a field, denoted by $\mathbf{C}$, which is called the **field of complex numbers**. Later we shall study the field $\mathbf{C}$ in more detail but, for now, we provide a brief introduction. We define $\mathbf{C} = \mathbf{R}^2 = \big\{(x,y)\big| x \in \mathbf{R}, y \in \mathbf{R}\big\}$. In $\mathbf{C}$, we write $0 = (0,0)$, $1 = (1,0)$ and $i = (0,1)$ and for $x, y \in \mathbf{R}$ we write $x = (x,0)$, $iy = (0,y)$ and $x+iy = (x,y)$. For $a, b, c, d \in \mathbf{R}$ we define

$$\begin{aligned}(a+ib) + (c+id) &= (a+c) + i(b+d),\\ (a+ib)\cdot(c+id) &= (ac-bd) + i(ad+bc).\end{aligned}$$

Note that in $\mathbf{C}$ we have $i^2 = -1$ and for $0 \le a \in \mathbf{R}$ we have $\left(i\sqrt{a}\right)^2 = -a$.

3.7 Exercise: Show that the set $\mathbf{Z}[i] = \big\{a+ib\big|a,b \in \mathbf{Z}\big\}$ is a commutative ring and that the set $\mathbf{Q}(i) = \big\{a+ib\big|a,b \in \mathbf{Q}\big\}$ is a field.

3.8 Example: For sets R and S, we write R^S for the set of all functions $f : S \to R$. When R is a ring, the set R^S is a ring with addition and multiplication defined, for $x \in S$, by $(f+g)(x) = f(x) + g(x)$ and $(fg)(x) = f(x)g(x)$. The zero and identity elements are the constant function 0 (given by $0(x) = 0$ for all $x \in S$) and the constant function 1 (given by $1(x) = 1$ for all $x \in S$.

3.9 Example: When R is a set and $n \in \mathbf{Z}^+ = \{1, 2, 3, \cdots\}$, we define $R^n = R^{\{1,2,\cdots,n\}}$, which is the set of all functions $a : \{1, 2, \cdots, n\} \to R$. An element of R^n is called an **n-tuple** with entries in R. We write $a = (a_1, a_2, \cdots, a_n)$ to indicate that a is the n-tuple $a : \{1, 2, \cdots, n\} \to R$ given by $a(k) = a_k \in R$ for each index $k \in \{1, 2, \cdots, n\}$. The elements $a_k \in R$ are called the **entries** of the n-tuple a. We have

$$R^n = \big\{(a_1, a_2, \cdots, a_n)\big|\text{ each } a_k \in R\big\}.$$

When R is a ring, R^n is a ring with addition and multiplication given by $(a+b)_k = a_k + b_k$ and $(ab)_k = a_k b_k$ and with 0 and 1 given by $0 = (0, 0, \cdots, 0)$ and $1 = (1, 1, \cdots, 1)$.

For a set R, we let $R^\infty = R^{\{1,2,3,\cdots\}}$, which is the set of all functions $a : \mathbf{Z}^+ \to R$. An element of R^∞ is called a **sequence** with entries in R. We write $a = (a_k)_{k\ge1} = (a_1, a_2, a_3, \cdots)$ to indicate that a is the sequence given by $a(k) = a_k$ for all $k \in \mathbf{Z}^+$. The elements $a_k \in R$ are called the **entries** of the sequence a. Thus we have

$$R^\infty = \big\{(a_1, a_2, a_3, \cdots)\big|\text{ each } a_k \in R\big\}.$$

When R is a ring, R^∞ is a ring with addition and multiplication given by $(a+b)_k = a_k + b_k$ and $(ab)_k = a_k b_k$ and with 0 and 1 given by $0 = (0, 0, 0, \cdots)$ and $1 = (1, 1, 1, \cdots)$.

More generally, given a set R, a **sequence** with entries in R is a function of the form $a : \{m, m+1, m+2, \cdots\} \to R$ for some $m \in \mathbf{Z}$, and we write $a = (a_n)_{n\ge m}$ to indicate that a is the sequence with $a(k) = a_k$ for all $k \ge m$.

3.10 Remark: In foundational mathematics, as outlined briefly in Appendix 1, the sets $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{C}$ are constructed using the ZFC axioms. In this procedure, the natural numbers are defined to be the sets $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$ and in general, $n = \{0, 1, 2, \cdots, n-1\}$. In foundational mathematics, the sets R^n and R^∞ are usually defined (slightly differently than we defined them above) by $R^n = R^{\{0,1,\cdots,n-1\}}$ and $R^\infty = R^{\mathbf{N}} = R^{\{0,1,2,\cdots\}}$.

3.11 Example: Let R be a ring. A **formal power series** in the variable x with coefficients in R is an expression of the form

$$f = f(x) = \sum_{k\geq 0} c_k x^k = \sum_{k=0}^{\infty} c_k x^k = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + \cdots$$

with each $c_k \in R$ The elements $c_k \in R$ are called the **coefficients** of f. The set of all formal power series in x with coefficients in R is denoted by $R[[x]]$. The set $R[[x]]$ is a ring with addition and multiplication defined as follows: for $f(x) = \sum_{k\geq 0} a_k x^k$ and $g(x) = \sum_{k\geq 0} b_k x^k$, we define $(f+g)(x) = \sum_{k\geq 0} (a_k + b_k)x^k$ and $(fg)(x) = \sum_{n\geq 0} c_n x^n$, where $c_n = \sum_{k=0}^{n} a_k b_{n-k} = a_0 b_n + a_1 b_{n-1} + \cdots a_n b_0$. A **polynomial** in the variable x with coefficients in R is a formal power series with only finitely many nonzero coefficients, that is a power series of the form

$$f = f(x) = \sum_{k\geq 0} c_k x^k = \sum_{k=0}^{n} c_k x^k = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n$$

for some $n \in \mathbf{N}$, where each $c_k \in R$ and $c_k = 0$ for all $k > n$. In this case, if $c_n \neq 0$ then we say that n is the **degree** of the polynomial f and we write $\deg(f) = n$. The set of all polynomials in x with coefficients in R is denoted by $R[x]$, and it is a ring using the same operations used in $R[[x]]$.

3.12 Remark: In foundational mathematics, the above definition would not be considered to be rigorous, because all mathematical objects must be defined to be sets (which can be constructed using the ZFC axioms). The above definition states that a formal power series is an "expression " of a certain form, but it does not define what such an expression actually is, as a set. To be rigorous, the power series which we denoted by $f(x) = \sum_{k=1}^{\infty} c_k x^k$ would be defined to be *equal* to the function $c \in R^{\mathbf{N}}$ which is given by $c(k) = c_k$. Using this definition, the variable symbol x is irrelevant and we have $R[[x]] = R[[y]]$. Also note that, using this definition, the sets $R[[x]]$ and $R^{\mathbf{N}}$ are *equal*, but the multiplication operation used in $R[[x]]$ is not equal to the multiplication operation used in $\mathbf{R}^{\mathbf{N}}$.

3.13 Example: Let R be a ring and let $m, n \in \mathbf{Z}^+$. An $m \times n$ **matrix** with entries in R is an expression of the form

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}$$

with each $a_{k,l} \in R$. The element $a_{k,l}$ is called the (k,l) entry of A, and we write $A_{k,l} = a_{k,l}$. The set of all $m \times n$ matrices with entries in R is denoted by $M_{m\times n}(R)$. We define the sum and product of two matrices as follows. For $A, B \in M_{m\times n}(R)$, we define $A+B \in M_{m\times n}(R)$ to be the matrix with entries $(A+B)_{k,l} = A_{k,l} + B_{k,l}$. For $A \in M_{m\times p}(R)$ and $B \in M_{p\times n}(R)$ we define $AB \in M_{m\times n}(R)$ to be the matrix with entries $(AB)_{k,l} = \sum_{j=1}^{p} A_{k,j} B_{j,l}$. When R is a ring and $n \in \mathbf{Z}^+$, the set of square matrices $M_n(R) = M_{n\times n}(R)$ is a non-commutative ring using these operations. The zero element is the **zero matrix** O with entries $O_{k,l} = 0$ for all k, l, and the identity element is the **identity matrix** I with entries $I_{k,k} = 1$ and $I_{k,l} = 0$ when $k \neq l$. The ring $M_n(R)$ plays an important role in linear algebra.

3.14 Definition: Let R be a ring. For $a, b \in R$, if $ab = ba = 1$ then we say that a is an **inverse** of b and that b is an **inverse** of a. For $a \in R$, if there exists $b \in R$ such that $ab = ba = 1$ then we say that a is **invertible** or that a is a **unit**.

3.15 Example: In $\mathbf{Z}$, the numbers 1 and -1 are invertible, and each is equal to its own inverse. In a field, every nonzero element is invertible (by R9). In $\mathbf{Z}[\sqrt{2}]$, the elements $a = 3 + 2\sqrt{2}$ and $b = 3 - 2\sqrt{2}$ are inverses of each other. The multiplication table in Example 3.5 shows that the only invertible elements in $\mathbf{Z}_6$ are 1 and 5, and each is equal to its own inverse. In $\mathbf{Z}_7$, every nonzero element is invertible, indeed we have $1 \cdot 1 = 1$, $2 \cdot 4 = 4 \cdot 2 = 1$, $3 \cdot 5 = 5 \cdot 3 = 1$ and $6 \cdot 6 = 1$. Verify that in $\mathbf{Z}[[x]]$, the power series $f(x) = 1 - x$ and $g(x) = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + \cdots$ are inverses. Verify that in $M_2(\mathbf{Z})$, the matrices $A = \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}$ and $B = \begin{pmatrix} -5 & 3 \\ 2 & -1 \end{pmatrix}$ are inverses.

3.16 Theorem: *(Uniqueness of Identity and Inverse) Let R be a ring. Then*

(1) the additive identity 0 is unique in the sense that if $e \in R$ has the property that $a + e = a$ for all $a \in R$ then $e = 0$,
(2) the additive inverse of $a \in G$ is unique in the sense that for all $a, b, c \in G$ if $a + b = 0$ and $a + c = 0$ then $b = c$,
(3) the multiplicative identity 1 is unique in the sense that if $u \in R$ has the property that $au = ua = a$ for all $a \in G$ then $u = 1$, and
(4) if $a \in R$ has an inverse, then it is unique in the sense that for all $a, b, c \in G$ if $ab = ba = 1$ and $ac = ca = 1$ then $b = c$.

Proof: Let us prove Part (1). Let $e \in R$ and suppose that $a + e = a$ for all $a \in R$. Then

$$\begin{aligned} e &= e + 0 \quad \text{, by R3,} \\ &= 0 + e \quad \text{, by R2,} \\ &= 0 \quad \text{, since } a + e = a \text{ for all } a \in R \text{ so in particular } 0 + e = 0. \end{aligned}$$

Let us prove Part (2). Let $a, b, c \in R$ and suppose that $a + b = 0$ and $a + c = 0$. Then

$$\begin{aligned} b &= b + 0 \text{ , by R3,} \\ &= 0 + b \text{ , by R2,} \\ &= (a + c) + b \text{ , since } a + c = 0, \\ &= (c + a) + b \text{ , by R2,} \\ &= c + (a + b) \text{ , by R1,} \\ &= c + 0 \text{ , since } a + b = 0, \\ &= c \text{ , by R3.} \end{aligned}$$

The proof of Parts (3) and (4) is left as an exercise.

3.17 Exercise: Convert the proof of Part (1) of the above theorem into a derivation of valid arguments to show that

$$\big\{\forall x\, \forall y\ x+y = y+x\,,\ \forall x\ x+0 = x\big\} \models \forall u\big(\forall x\ x+u = x \rightarrow u = 0\big).$$

3.18 Notation: Let R be a ring. For $a \in R$ we denote the unique additive inverse of $a \in R$ by $-a$, and for $a, b \in R$ we write $b - a$ for $b + (-a)$. When $a \in R$ is invertible, we denote its unique multiplicative inverse by a^{-1}. When F is a field and $a \neq 0$ we also write a^{-1} as $\frac{1}{a}$, and when $a, b \in F$ with $a \neq 0$, we write $b \div a = b/a = \frac{b}{a} = b\,a^{-1}$.

3.19 Definition: Let R be a ring. For $a, b \in R$, if $a \neq 0$ and $b \neq 0$ and $ab = 0$ then we say that a and b are **zero divisors**.

3.20 Theorem: *(Properties of Rings) Let R be a ring. Then for all $a, b, c \in R$,*

(1) if $a + b = a + c$ then $b = c$,
(2) if $a + b = a$ then $b = 0$,
(3) if $a + b = 0$ then $b = -a$,
(4) $a \cdot 0 = 0 = 0 \cdot a$ for all $a \in R$,
(5) $-(-a) = a$ for all $a \in R$,
(6) $(-a)b = -(ab) = a(-b)$ for all $a, b \in R$,
(7) $(-a)(-b) = ab$ for all $a, b \in R$,
(8) $(-1)a = -a$ for all $a \in R$,
(9) $a(b - c) = ab - ac$ and $(a - b)c = ac - bc$ for all $a, b, c \in R$,
(10) if $ab = 1$ and $bc = 1$ then b is invertible and $a = c = b^{-1}$,
(11) if a and b are invertible then so is ab and we have $(ab)^{-1} = b^{-1}a^{-1}$,
(12) if $ab = ac$, or if $ba = ca$, then either $a = 0$, or a is a zero divisor, or $b = c$,
(13) if a is a unit then a is not a zero divisor, and
(14) if R is contained in a field (and uses the same operations) then R has no zero divisors.

Proof: We give a few sample proofs. To prove Part (1), suppose that $a + b = a + c$. Let $d = -a$ so that $a + d = 0$ (we can do this by R4). Then

$$\begin{aligned}
b &= b + 0 \text{ , by R3,} \\
&= b + (a + d) \text{ , since } a + d = 0, \\
&= (b + a) + d \text{ , by R1,} \\
&= (a + b) + d \text{ , by R2,} \\
&= (a + c) + d \text{ , since } a + b = a + c, \\
&= (c + a) + d \text{ , by R2,} \\
&= c + (a + d) \text{ , by R1,} \\
&= c + 0 \text{ , since } a + d = 0, \\
&= c \text{ , by R3.}
\end{aligned}$$

To prove Part (2), suppose that $a + b = a$. Since $a + b = a$ and $a = a + 0$ (by R3), we have $a + b = a + 0$, and so $b = 0$ by Part (1).

To prove Part (3), suppose that $a + b = 0$. Since $a + b = 0$ and $0 = a + (-a)$, we have $a + b = a + (-a)$, and so $b = -a$ by Part (1).

To prove Part (4), note that since $0 = 0 + 0$ (by R3) we have $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$ (by R7). Since $0 \cdot a + 0 \cdot a = 0 \cdot a$ it follows that $0 \cdot a = 0$, by Part (1). Similarly, $a \cdot 0 = 0$.

To prove Part (5), note that $(-a) + a = a + (-a) = 0$ so $a = -(-a)$ by Part (3).

To prove Part (8), note that

$$\begin{aligned}
a + (-1)a &= 1 \cdot a + (-1)a \text{ , by R6,} \\
&= (1 + (-1)) \cdot a \text{ , by R7,} \\
&= 0 \cdot a \text{ , since } 1 + (-1) = 0, \\
&= 0 \text{ , by Part (4).}
\end{aligned}$$

Since $a + (-1)a = 0$ we have $(-1)a = -a$ by Part (3).

3.21 Example: In $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{C}$ there are no zero divisors. In $\mathbf{Z}_6$ we have $2 \cdot 3 = 0$ so 2 and 3 are zero divisors. In $\mathbf{Z}^2$, given $0 \neq a, b \in \mathbf{Z}$ we have $(a, 0) \cdot (0, b) = (0, 0)$ and so $(a, 0)$ and $(0, b)$ are zero divisors. In $M_2(\mathbf{Z})$, for $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & -4 \\ -1 & 2 \end{pmatrix}$ we have $AB = O$ and so A and B are zero divisors.

3.22 Definition: A commutative ring with no zero divisors is called an **integral domain**.

3.23 Definition: An **order** on a set X is a binary relation $\leq$ on X such that

O1. (Totality) for all $x, y \in X$, either $x \leq y$ or $y \leq x$,
O2. (Antisymmetry) for all $x, y \in X$, if $x \leq y$ and $y \leq x$ then $x = y$, and
O3. (Transitivity) for all $x, y, z \in X$, if $x \leq y$ and $y \leq z$ then $x \leq z$.

An **ordered set** is a set X with on order $\leq$.

3.24 Theorem: *Each of* $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$ *and* $\mathbf{R}$ *is an ordered set using its standard order* $\leq$ *with* $\mathbf{N} \subseteq \mathbf{Z} \subseteq \mathbf{Q} \subseteq \mathbf{R}$ *(and the orders coincide so that for example when* $a, b \in \mathbf{N}$ *we have* $a \leq b$ *in* $\mathbf{N}$ *if and only if* $a \leq b$ *in* $\mathbf{R}$*).*

Proof: We accept the truth of this theorem axiomatically, without proof.

3.25 Notation: When $\leq$ is an order on X, we write $x < y$ when $x \leq y$ and $x \neq y$, we write $x \geq y$ when $y \leq x$ and we write $x > y$ when $y < x$.

3.26 Theorem: *Let* $\leq$ *be an order on a set* X*. Then*

(1) for all $x, y \in X$*, we have* $x \leq y \iff \big(x < y \text{ or } x = y\big)$,
(2) for all $x, y \in X$ *exactly one of the following 3 statements holds:*

$$x = y,\ x < y \text{ or } y < x.$$

(3) for all $x, y, z \in X$*, if* $x < y$ *and* $y < z$ *then* $x < z$.

Solution: We shall only prove Parts (1) and (2). Let $x, y \in X$. By the definition of $<$, to prove Part (1) we need to show that $x \leq y \iff \big((x \leq y \text{ and } x \neq y) \text{ or } x = y\big)$. Suppose first that $x \leq y$. Note that either $x = y$ or $x \neq y$. If $x = y$ then the statement $\big((x \leq y \text{ and } x \neq y) \text{ or } x = y\big)$ is true. If $x \neq y$ then we have $(x \leq y \text{ and } x \neq y)$ and so again the statement $\big((x \leq y \text{ and } x \neq y) \text{ or } x = y\big)$ is true. This completes the proof that $x \leq y \Longrightarrow \big((x \leq y \text{ and } x \neq y) \text{ or } x = y\big)$. Suppose, conversely, that either $(x \leq y \text{ and } x \neq y)$ or $x = y$. If $x \leq y$ and $x \neq y$ then of course $x \leq y$. Suppose that $x = y$. By applying O1 in the case that $y = x$, we find that $(x \leq x \text{ or } x \leq x)$ or, more simply, $x \leq x$. Since $x = y$ and $x \leq x$ if follows (by substitution) that $x \leq y$. This completes the proof that $\big((x \leq y \text{ and } x \neq y) \text{ or } x = y\big) \Longrightarrow x \leq y$.

Let us prove Part (2). First we show that at least one of the 3 statement holds. We need to show that either $(x = y \text{ or } x < y)$ or $y < x$. By Part (1), this is equivalent to showing that either $x \leq y$ or $y < x$. Suppose that $y \not< x$. By the definition of $y < x$ we are supposing that it is not the case that $(y \leq x \text{ and } y \neq x)$ or, equivalently, we are supposing that either $y \not\leq x$ or $y = x$. In the case that $y \not\leq x$, it follows by O1 that $x \leq y$. In the case that $y = x$ we already showed above that $x \leq y$. This completes the proof that at least 1 of the 3 statements holds. It is not possible to have $x = y$ and $x < y$ because, by definition, when $x < y$ we have $x \neq y$. Similarly it is not possible to have $x = y$ and $y < x$. Suppose, for a contradiction, that $x < y$ and $y < x$. Since $x < y$ we have $x \leq y$ (by the definition of $x < y$). Since $y < x$ we have $y \leq x$ (by definition). Since $x \leq y$ and $y \leq x$ we have $x = y$ by O2. But since $x < y$ we have $x \neq y$ (by definition) and this gives the desired contradiction.

3.27 Definition: An **ordered field** is a field F with an order $\leq$ such that

O4. (Compatibility with $+$) for all $x, y, z \in F$, if $x \leq y$ then $x + z \leq y + z$, and

O5. (Compatibility with $\times$) for all $x, y \in F$, if $0 \leq x$ and $0 \leq y$ then $0 \leq xy$.

When F is an ordered field and $x \in F$ we say that x is **positive** when $x > 0$, we say x is **negative** when $x < 0$, we say x is **nonpositive** when $x \leq 0$, and we say x is **nonnegative** when $x \geq 0$.

3.28 Theorem: R *is an ordered field.*

Proof: We accept this axiomatically, without proof.

3.29 Corollary: *Any field F, which is a subset of* **R** *and uses the same operations and the same order, is an ordered field.*

3.30 Example: Q and $\mathbf{Q}[\sqrt{2}]$ are ordered fields.

3.31 Theorem: *(Properties of Ordered Fields) Let F be an ordered field. Then for all $x, y, z \in F$*

(1) if $x \geq 0$ then $-x \leq 0$, and if $x \leq 0$ then $-x \geq 0$,
(2) if $x \geq 0$ and $y \leq z$ then $xy \leq xz$,
(3) if $x \leq 0$ and $y \leq z$ then $xz \leq xy$,
(4) $0 \leq x^2$,
(5) we have $0 < 1$ and $-1 < 0$, and
(6) if $0 < x$ then $0 < \frac{1}{x}$ and if $0 < x \leq y$ then $0 < \frac{1}{y} \leq \frac{1}{x}$.

Proof: We shall provide two proofs for each of the first two parts. The first proof will be brief, using standard mathematical language, and will use some of the properties from Definition 3.2 and Theorem 3.20 implicitly. The second proof will be more detailed, indicating explicitly which property is being used at each step of the proof. Then we provide brief proofs for each of the remaining parts.

For the first proof of Part (1), let $x \in F$. If $x \geq 0$, that is if $0 \leq x$, then by O4 we have $0 + (-x) \leq x + (-x)$ hence $-x \leq 0$, and if $0 \leq x$ then by O4 we have $0 + (-x) \leq x + (-x)$ hence $-x \leq 0$.

We now repeat the above prof of Part (1) adding additional detail. Let $x \in F$ be arbitrary. Let $u = -x$ so that $x + u = 0$ (using R4 and the notation from 3.18). First suppose that $x \geq 0$, which means that $0 \leq x$ by Notation 3.27. Then

$$\begin{aligned} 0 + u &\leq x + u \text{ , by O4,} \\ 0 + u &\leq 0 \text{ , since } x + u = 0, \\ u + 0 &\leq 0 \text{ , since } 0 + u = u + 0 \text{ by R2,} \\ u &\leq 0 \text{ , since } u + 0 = u \text{ by R3,} \\ -x &\leq 0 \text{ , since } u = -x. \end{aligned}$$

Next suppose that $x \leq 0$. Then

$$\begin{aligned} x + u &\leq 0 + u \text{ , by O4,} \\ 0 &\leq 0 + u \text{ , since } x + u = 0, \\ 0 &\leq u + 0 \text{ , since } 0 + u = u + 0 \text{ by R2,} \\ 0 &\leq u \text{ , since } u + 0 = u \text{ by R3.} \\ u &\geq 0 \text{ , by Notation 3.27,} \\ -x &\geq 0 \text{ , since } u = -x. \end{aligned}$$

To prove Part (2), let $x, y, z \in F$ and suppose that $0 \leq x$ and $y \leq z$. Since $y \leq z$, by O4 we have $y + (-y) \leq z + (-y)$, hence $0 \leq z - y$. Since $0 \leq x$ and $0 \leq z - y$, by O5 we have $0 \leq x(z - y)$, and hence by Theorem 3.20 Part (9) we have $0 \leq xz - xy$. By O4 it follows that $0 + xy \leq (xz - xy) + xy$. Thus

$$xy = xy + 0 = 0 + xy \leq (xz - xy) + xy = xz + (-xy + xy) = xz + 0 = xz.$$

We now provide a second proof of Part (2), adding additional detail to the above proof. Also, we shall avoid using Part (9) of Theorem 3.20, since we did not prove it. Instead, we shall use Part (4) which we did prove.

Let $x, y, z \in F$ be arbitrary. Suppose that $x \geq 0$, that is $0 \leq x$, and suppose that $y \leq z$. Let $u = -y$ so that $y + u = 0$ (using R4 and Notation 3.18). Then

$$\begin{aligned}
y + u &\leq z + u \text{ , since } y \leq z, \text{ by O4,} \\
0 &\leq z + u \text{ , since } y + u = 0, \\
0 &\leq x(z + u) \text{ , since } 0 \leq x \text{ and } 0 \leq z + u, \text{ by O5,} \\
0 &\leq xz + xu \text{ , by R7,} \\
0 + xy &\leq (xz + xu) + xy \text{ , by O4,} \\
0 + xy &\leq xz + (xu + xy) \text{ , by R1,} \\
0 + xy &\leq xz + x(u + y) \text{ , by R7,} \\
0 + xy &\leq xz + x(y + u) \text{ , by R2,} \\
0 + xy &\leq xz + x \cdot 0 \text{ , since } y + u = 0, \\
0 + xy &\leq xz + 0 \text{ , by Theorem 3.20 Part (4),} \\
0 + xy &\leq xz \text{ , by R3,} \\
xy + 0 &\leq xz \text{ , by R2,} \\
xy &\leq xz \text{ , by R3.}
\end{aligned}$$

To prove Part (3), let $x, y, z \in F$ and suppose that $x \leq 0$ and $y \leq z$. Since $x \leq 0$ we have $0 \leq -x$ by Part (1). Since $y \leq z$, by O4 we have $y - y \leq z - y$, that is $0 \leq z - y$. Since $0 \leq -x$ and $0 \leq z - y$, by O5 we have $0 \leq (-x)(z - y)$. Using some properties of rings, it follows that $0 \leq xy - xz$ hence, by O4, $0 + xz \leq (xy - xz) + xz$, and hence $xz \leq xy$.

We prove Part (4) by considering two cases. Let $x \in F$ be arbitrary. By O1 we know that either $x \leq 0$ or $0 \leq x$. If $0 \leq x$ then by O5 we have $0 \leq x \cdot x$, that is $0 \leq x^2$. If $x \leq 0$ then by Part (1) we have $0 \leq -x$ and so, by O5, we have $0 \leq (-x)(-x)$ hence, by Part (7) of Theorem 3.20, we have $0 \leq x \cdot x$, that is $0 \leq x^2$. In either case we find that $0 \leq x^2$, as required.

By Part (4) we have $0 \leq 1^2 = 1$. Since $0 \leq 1$ and $0 \neq 1$, we have $0 < 1$. We leave the proof that $-1 < 0$ as an exercise.

To prove Part (6), let $x, y \in F$ with $0 < x \leq y$. Suppose, for a contradiction, that $\frac{1}{x} \leq 0$. Since $0 \leq x$ and $\frac{1}{x} \leq 0$ it follows from Part (2) that $x \cdot \frac{1}{x} \leq x \cdot 0$, and hence $1 \leq 0$. But we know from Part (4) that $0 < 1$ and so we have the desired contradiction. Since it is not the case that $\frac{1}{x} \leq 0$ we must have $0 < \frac{1}{x}$ by O1 and O2. Since $0 \leq x$ and $x \leq y$ we have $0 \leq y$ by O3. If we had $y = 0$ then we would have $y = 0 < x$ and $x \leq y$ which is not possible by O1 and O2. Since $0 \leq y$ and $y \neq 0$ we have $0 < y$. As above, since $0 < y$ we have $0 < \frac{1}{y}$. It remains to show that $\frac{1}{y} \leq \frac{1}{x}$. If $x = y$ then $\frac{1}{x} = \frac{1}{y}$. Suppose that $x < y$. If we had $\frac{1}{x} \leq \frac{1}{y}$ then, since $0 \leq x$ and $0 \leq y$, it would follow from O5 that $\frac{1}{x}\, xy \leq \frac{1}{y}\, xy$, so that $y \leq x$, which contradicts the fact that $x < y$. Thus $\frac{1}{y} < \frac{1}{x}$.

3.32 Note: The various properties of ordered fields, which were stated in terms of the order relation $\leq$, have analogous counterparts involving the strict order relation $<$. As an exercise, verify that the following properties hold when F is an ordered field and $x, y, z \in F$.

(1) If $x < y$ then $x + z < y + z$,
(2) if $x > 0$ and $y > 0$ then $xy > 0$,
(3) if $x > 0$ then $-x < 0$, and if $x < 0$ then $-x > 0$,
(4) if $x > 0$ and $y < z$ then $xy < xz$.
(5) if $x < 0$ and $y < z$ then $xy > xz$,
(6) if $x \neq 0$ then $x^2 > 0$,
(7) if $0 < x < y$ then $0 < \frac{1}{y} < \frac{1}{x}$.
(8) if $x < y < 0$ then $\frac{1}{y} < \frac{1}{x} < 0$.

3.33 Note: Note that it is not possible to define an order $\leq$ on $\mathbf{C}$ which makes $\mathbf{C}$ into an ordered field because if $\mathbf{C}$ was an ordered field then since $0 < 1$ we would have $-1 < 0$ (by Part 3) but since $-1 = i^2$ we would also have $-1 > 0$ (by Part 6).

3.34 Definition: Let F be an ordered field. For $a \in F$ we define the **absolute value** of a to be

$$|a| = \begin{cases} a \text{ if } a \geq 0, \\ -a \text{ if } a \leq 0. \end{cases}$$

3.35 Theorem: *(Properties of Absolute Value) Let F be an ordered field. For all $x, y \in F$*

(1) (Positive Definiteness) $|x| \geq 0$ with $|x| = 0 \Longleftrightarrow x = 0$,
(2) (Symmetry) $|x - y| = |y - x|$,
(3) (Multiplicativeness) $|xy| = |x|\,|y|$
(4) (Triangle Inequality) $\big||x| - |y|\big| \leq |x + y| \leq |x| + |y|$, and
(5) (Approximation) for $a, b \in F$ with $b \geq 0$ we have $|x - a| \leq b \Longleftrightarrow a - b \leq x \leq a + b$.

Proof: The proof is left as an exercise.

3.36 Theorem: *(Basic Order Properties of $\mathbf{Z}$)*

(1) For all $n \in \mathbf{Z}$ we have $n \in \mathbf{N}$ if and only if $n \geq 0$.
(2) For all $k, n \in \mathbf{Z}$ we have $k \leq n$ if and only if $k < n + 1$. Equivalently, for all $n \in \mathbf{Z}$ there does not exist $k \in \mathbf{Z}$ with $n < k < n + 1$.

Proof: This theorem is accepted as true axiomatically, without proof.

3.37 Example: Prove that for all $k, l \in \mathbf{Z}$, if $kl = 1$ then either $k = l = 1$ or $k = l = -1$.

Solution: Let $k, l \in \mathbf{Z}$ and suppose that $kl = 1$. By Theorem 3.26 Part (2), applied several times, exactly 1 of the following 7 possibilities holds: $k < -1$, $k = -1$, $-1 < k < 0$, $k = 0$, $0 < k < -1$, $k = -1$ or $-1 < k$. Note that $k \neq 0$ since if we had $k = 0$ then we would hove $1 = kl = 0 \cdot l = 0$. Also, we cannot have $-1 < k < 0$ or $0 < k < 1$ by Theorem 3.36 Part (2), so we are left with the following 4 possibilities: $k < -1$, $k = -1$, $k = 1$ or $1 < k$. Suppose, for a contradiction, that $1 < k$. By Theorem 3.31 Part (5) we have $0 < 1 < k$, so by Note 3.32 Part (7) we have $0 < \frac{1}{k} < \frac{1}{1}$. This implies that $0 < l < 1$ $\big($because $kl = 1$ so that $l = \frac{1}{k}$ and $1 \cdot 1 = 1$ so that $\frac{1}{1} = 1\big)$, but this is not possible by Theorem 3.36 Part (2). Similarly, if we had $k < -1$ then we would have $\frac{1}{-1} < \frac{1}{k} < 0$ and hence $-1 < l < 0$, which is impossible by Theorem 3.36 Part (2). Thus we have eliminated 5 of the 7 possibilities leaving only the 2 possibilities $k = \pm 1$. Finally note that when $k = 1$ we have $1 = kl = 1 \cdot l = l$ and when $k = -1$ we have $1 = kl = (-1)l = -l$.

3.38 Remark: At this stage we have have either proven, or accepted axiomatically, all of the algebraic properties of $\mathbf{Z}$ which were used in our proof in Example 2.31 at the end of the previous chapter.

3.39 Definition: Let X be an ordered set and let $A \subseteq X$. We say that A is **bounded above** (in X) when there exists an element $b \in X$ such that $x \leq b$ for all $x \in A$, and in this case we say that b is an **upper bound** for A (in X). We say that A is **bounded below** (in X) when there exists an element $a \in X$ such that $a \leq x$ for all $x \in A$, and in this case we say that a is a **lower bound** for A (in X). We say that A is **bounded** (in X) when A is bounded above and bounded below.

3.40 Definition: Let X be an ordered set and let $A \subseteq X$. We say that A has a **supremum** (or a **least upper bound**) (in X) when there exists an element $b \in X$ such that b is an upper bound for A with $b \leq c$ for every upper bound $c \in X$ for A, and in this case we say that b is the **supremum** (or the **least upper bound**) of A (in X) (note that if the supremum exists then it is unique by antisymmetry) and we write $b = \sup A$. When the supremum $b = \sup A$ exists and we have $b \in A$, then we also say that b is the **maximum element** of A and we write $b = \max A$.

We say that A has an **infimum** (or a **greatest lower bound**) (in X) when there exists an element $a \in X$ such that a is a lower bound for A with $c \leq a$ for every lower bound c for A, and in this case we say that a is the **infimum** (or the **greatest lower bound**) of A (in X) and we write $a = \inf A$. When $a = \inf A \in A$ we also say that a is the **minimum element** of A and we write $a = \min A$.

3.41 Example: Let $A = (0, \infty) = \{x \in \mathbf{R} \,|\, 0 < x\}$ and $B = [1, \sqrt{2}) = \{x \in \mathbf{R} \,|\, 1 \leq x < \sqrt{2}\}$. The set A is bounded below but not bounded above. The numbers -1 and 0 are both lower bounds for A and we have $\inf A = 0$. The set A has no minimum element and no maximum element. The set B is bounded above and below. The numbers 0 and 1 are both lower bounds for B and the numbers $\sqrt{2}$ and 3 are both upper bounds for B. We have $\inf B = 1$ and $\sup B = \sqrt{2}$. The set B has a minimum element, namely $\min B = \inf B = 1$, but B has no maximum element.

3.42 Theorem: *(Least Upper Bound and Greatest Lower Bound Properties of* $\mathbf{R}$*)*

(1) Every nonempty subset of $\mathbf{R}$ *which is bounded above in* $\mathbf{R}$ *has a supremum in* $\mathbf{R}$*.*
(2) Every nonempty subset of $\mathbf{R}$ *which is bounded below in* $\mathbf{R}$ *has an infimum in* $\mathbf{R}$*.*

Proof: We accept this axiomatically, without proof.

3.43 Theorem: *(Approximation Property of Supremum and Infimum) Let* $\emptyset \neq A \subseteq \mathbf{R}$*.*

(1) If $b = \sup A$ *then for all* $0 < \epsilon \in \mathbf{R}$ *there exists* $x \in A$ *with* $b - \epsilon < x \leq b$*, and*
(2) if $a = \inf A$ *then for all* $0 < \epsilon \in \mathbf{R}$ *there exists* $x \in A$ *with* $a \leq x < a + \epsilon$*.*

Proof: We prove Part (1). Let $b = \sup A$. Let $\epsilon > 0$. Suppose, for a contradiction, that there is no element $x \in A$ with $b - \epsilon < x$, or equivalently that for all $x \in A$ we have $b - \epsilon \geq x$. Let $c = b - \epsilon$. Note that c is an upper bound for A since $x \leq b - \epsilon = c$ for all $x \in A$. Since $b = \sup A$ and c is an upper bound for A we have $b \leq c$. But since $\epsilon > 0$ we have $b > b - \epsilon = c$ giving the desired contradiction. This proves that there exists $x \in A$ with $b - \epsilon < x$. Choose such an element $x \in A$. Since $b = \sup A$ we know that b is an upper bound for A and hence $b \geq x$. Thus we have $b - \epsilon < x \leq b$, as required.

3.44 Theorem: *(Well-Ordering Properties of* $\mathbf{Z}$ *in* $\mathbf{R}$*)*

(1) Every nonempty subset of $\mathbf{Z}$ *which is bounded above in* $\mathbf{R}$ *has a maximum element.*
(2) Every nonempty subset of $\mathbf{Z}$ *which is bounded below in* $\mathbf{R}$ *has a minimum element, in particular every nonempty subset of* $\mathbf{N}$ *has a minimum element.*

Proof: We prove Part (1). Let A be a nonempty subset of $\mathbf{Z}$ which is bounded in $\mathbf{R}$. By Theorem 3.39, A has a supremum in $\mathbf{R}$. Let $n = \sup A$. We must show that $n \in A$. Suppose, for a contradiction, that $n \notin A$. By the Approximation Property (using $\epsilon = 1$), we can choose $a \in A$ with $n - 1 < a \leq n$. Note that $a \neq n$ since $a \in A$ and $n \notin A$ and so we have $a < n$. By the Approximation Property again (using $\epsilon = n - a$) we can choose $b \in A$ with $a < b \leq n$. Since $a < b$ we have $b - a > 0$. Since $n - 1 < a$ and $b \leq n$ we have $1 = n - (n-1) > b - a$. But then we have $b - a \in \mathbf{Z}$ with $0 < b - a < 1$ which contradicts the Basic Order Properties of $\mathbf{Z}$. Thus $n \in A$ so A has a maximum element.

3.45 Theorem: *(Floor and Ceiling Properties of* $\mathbf{Z}$ *in* R*)*

(1) (Floor Property) For every $x \in \mathbf{R}$ *there exists a unique* $n \in \mathbf{Z}$ *with* $x - 1 < n \leq x$.
(2) (Ceiling Property) For every $x \in \mathbf{R}$ *there exists a unique* $m \in \mathbf{Z}$ *with* $x \leq m < x + 1$.

Proof: We prove Part (1). First we prove uniqueness. Let $x \in \mathbf{R}$ and suppose that $n, m \in \mathbf{Z}$ with $x - 1 < n \leq x$ and $x - 1 < m \leq x$. Since $x - 1 < n$ we have $x < n + 1$. Since $m \leq x$ and $x < n + 1$ we have $m < n + 1$ hence $m \leq n$. Similarly, we have $n \leq m$. Since $n \leq m$ and $m \leq n$, we have $n = m$. This proves uniqueness.

Next we prove existence. Let $x \in \mathbf{R}$. First let us consider the case that $x \geq 0$. Let $A = \{k \in \mathbf{Z} \,|\, k \leq x\}$. Note that $A \neq \emptyset$ because $0 \in A$ and A is bounded above in $\mathbf{R}$ by x. By The Well-Ordering Property of $\mathbf{Z}$ in $\mathbf{R}$, A has a maximum element. Let $n = \max A$. Since $n \in A$ we have $n \in \mathbf{Z}$ and $n \leq x$. Also note that $x - 1 < n$ since $x - 1 \geq n \Longrightarrow x \geq n + 1 \Longrightarrow n + 1 \in A \Longrightarrow n \neq \max A$. Thus for $n = \max A$ we have $n \in \mathbf{Z}$ with $x - 1 < n \leq x$, as required.

Next consider the case that $x < 0$. If $x \in \mathbf{Z}$ we can take $n = x$. Suppose that $x \notin \mathbf{Z}$. We have $-x > 0$ so, by the previous paragraph, we can choose $m \in \mathbf{Z}$ with $-x - 1 < m \leq -x$. Since $m \in \mathbf{Z}$ but $x \notin \mathbf{Z}$ we have $m \neq -x$ so that $-x - 1 < m < -x$ and hence $x < -m < x + 1$. Thus we can take $n = -m - 1$ to get $x - 1 < n < x$. This completes the proof of Part (1).

3.46 Definition: Given $x \in \mathbf{R}$ we define the **floor** of x to be the unique $n \in \mathbf{Z}$ with $x - 1 < n \leq x$ and we denote the floor of x by $\lfloor x \rfloor$. The function $f : \mathbf{R} \to \mathbf{Z}$ given by $f(x) = \lfloor x \rfloor$ is called the **floor function**.

3.47 Theorem: *(Archimedean Properties of* $\mathbf{Z}$ *in* $\mathbf{R}$*)*

(1) For every $x \in \mathbf{R}$ *there exists* $n \in \mathbf{Z}$ *with* $n > x$.
(2) For every $x \in \mathbf{R}$ *there exists* $m \in \mathbf{Z}$ *with* $m < x$.

Proof: Let $x \in \mathbf{R}$. Let $n = \lfloor x \rfloor + 1$ and $m = \lfloor x \rfloor - 1$. Since $x - 1 < \lfloor x \rfloor$ we have $x < \lfloor x \rfloor + 1 = n$ and since $\lfloor x \rfloor \leq x$ we have $m = \lfloor x \rfloor - 1 \leq x - 1 < x$.

3.48 Theorem: *(Density of* $\mathbf{Q}$ *in* $\mathbf{R}$*) For all* $a, b \in \mathbf{R}$ *with* $a < b$ *there exists* $q \in \mathbf{Q}$ *with* $a < q < b$.

Proof: Let $a, b \in \mathbf{R}$ with $a < b$. By the Archimedean Property, we can choose $n \in \mathbf{Z}$ with $n > \frac{1}{b-a} > 0$. Then $n(b - a) > 1$ and so $nb > na + 1$. Let $k = \lfloor na + 1 \rfloor$. Then we have $na < k \leq na + 1 < nb$ hence $a < \frac{k}{n} < b$. Thus we can take $q = \frac{k}{n}$ to get $a < q < b$.

.

Chapter 4. Recursion and Induction

4.1 Theorem: *(Mathematical Induction) Let $F(n)$ be a statement about $n \in \mathbf{Z}$ and let $m \in \mathbf{Z}$. Suppose that $F(m)$ is true. Suppose that for all $n \in \mathbf{Z}$ with $n \geq m$, if $F(n)$ is true then $F(n+1)$ is true. Then $F(n)$ is true for all $n \in \mathbf{Z}$ with $n \geq m$.*

Proof: Let $S = \{k \in \mathbf{Z} | k \geq m \text{ and } F(k) \text{ is false}\}$. To prove that $F(n)$ is true for all $n \geq m$, we shall prove that $S = \emptyset$. Suppose, for a contradiction, that $S \neq \emptyset$. Since $S \neq \emptyset$ and S is bounded below by m, it follows from the Well-Ordering Property of $\mathbf{Z}$ that S has a minimum element. Let $a = \min(S)$. Since $a \in S$ it follows that $a \geq m$ and $F(a)$ is false. Since $F(m)$ is true and $F(a)$ is false, it follows that $a \neq m$. Since $a \geq m$ and $a \neq m$ it follows that $a > m$ and so $a - 1 \geq m$. We claim that $F(a-1)$ is true. Suppose, for a contradiction, that $F(a-1)$ is false. Since $a - 1 \geq m$ and $F(a-1)$ is false, it follows that $a-1 \in S$. Since $a = \min(S)$ and $a-1 \in S$, we have $a \leq a-1$. But we know that $a > a-1$ so we have obtained the desired contradiction (to the assumption that $F(a-1)$ is false). Thus $F(a-1)$ is true, as claimed. Since $a - 1 \geq m$ and $F(a-1)$ is true, it follows by the hypothesis in the statement of the theorem that $F(a)$ is true. But, as mentioned earlier, since $a \in S$ we know that $F(a)$ is false, so we have obtained the desired contradiction (to the assumption that $S \neq \emptyset$). Thus $S = \emptyset$, as required.

4.2 Note: It follows, from the above theorem, that in order to prove that $F(n)$ is true for all $n \geq m$, we can do the following.

1. Prove that $F(m)$ is true (this is called proving the **base case**).
2. Let $n \geq m$ and suppose that $F(n)$ is true (this is called the **induction hypothesis**).
3. Prove that $F(n+1)$ is true.

Alternatively, we can prove that $F(n)$ is true for all $n \geq m$ as follows: prove that $F(m)$ is true, let $n > m$ and suppose that $F(n-1)$ is true, then prove that $F(n)$ is true.

4.3 Definition: For a sequence $(a_n)_{n \geq m}$, a formula for a_n in terms of n is called a **closed-form** formula for a_n, and a formula for a_n in terms of n along with previous terms a_k with $k < n$ is called a **recursion** formula for a_n.

4.4 Example: For the sequence $(a_n)_{n \geq 0}$ given in closed-form by $a_n = n^2$ for $n \geq 0$, we have

$$(a_n)_{n \geq 0} = (0, 1, 4, 9, 16, 25, 36, \cdots).$$

For the sequence $(a_n)_{n \geq 0}$ defined recursively by $a_0 = 2$ and $a_{n+1} = 2a_n - 1$ for $n \geq 0$, we have

$$(a_n)_{n \geq 0} = (2, 3, 5, 9, 17, 33, 62, \cdots).$$

The **Fibonacci sequence** is defined recursively by $a_0 = 0$, $a_1 = 1$ and $a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$, so we have

$$(a_n)_{n \geq 0} = (0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \cdots).$$

4.5 Example: When we write

$$S_n = \sum_{k=m}^{n} a_k = a_m + a_{m+1} + \cdots + a_n$$

what we really mean is that the sequence S_n is defined recursively by $S_m = a_m$ and $S_n = S_{n-1} + a_n$ for all $n > m$.

4.6 Example: When we write

$$P_n = \prod_{k=m}^{n} a_k = a_m \cdot a_{m+1} \cdot \ldots \cdot a_n$$

what we really mean is that the sequence P_n is defined recursively by $P_m = a_m$ and $P_n = P_{n-1} \cdot a_n$ for $n > m$.

4.7 Example: When we say that $n!$ (read as n **factorial**) is defined for $n \in \mathbf{N}$ by $0! = 1$ and $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$, what we really mean is that $n!$ is defined recursively by $0! = 1$ and $n! = n \cdot (n-1)!$ for $n \geq 1$.

4.8 Example: Let $a_1 = 1$ and let $a_{n+1} = \frac{n}{n+1}\, a_n + 1$ for all $n \geq 1$. Find a closed-form formula for a_n.

Solution: Using the given recursion formula, the first few terms in the sequence $(a_n)_{n\geq 1}$ are as follows.

n	1	2	3	4	5	6
a_n	1	$\frac{3}{2}$	2	$\frac{5}{2}$	3	$\frac{7}{2}$

It appears, from the table, that $a_n = \frac{n+1}{2}$ for all $n \geq 1$. Here is a proof by induction. When $n = 1$ we have $\frac{n+1}{2} = \frac{1+1}{2} = 1 = a_1 = a_n$. Let $n \geq 1$ be arbitrary and suppose, inductively, that $a_n = \frac{n+1}{2}$ (for this one particular value of n). Then we have

$$\begin{aligned} a_{n+1} &= \tfrac{n}{n+1}\, a_n + 1 \text{ by the recursion formula} \\ &= \tfrac{n}{n+1} \cdot \tfrac{n+1}{2} + 1 \text{ by the induction hypothesis} \\ &= \tfrac{n}{2} + 1 = \tfrac{(n+1)+1}{2} \text{ , as required.} \end{aligned}$$

By induction, it follows that $a_n = \frac{n+1}{2}$ for all $n \geq 1$.

4.9 Example: Find $\prod_{k=2}^{n} \left(1 - \frac{1}{k^2}\right)$.

Solution: Let $P_n = \prod_{k=2}^{n} \left(1 - \frac{1}{k^2}\right)$ for $n \geq 2$. This means that the sequence $(P_n)_{n\geq 2}$ is defined recursively by $P_2 = 1 - \frac{1}{4} = \frac{3}{4}$ and $P_n = P_{n-1}\left(1 - \frac{1}{n^2}\right)$ for all $n \geq 3$. The first few values of P_n are as follows. $P_2 = 1 - \frac{1}{4} = \frac{3}{4}$, $P_3 = P_2\left(1 - \frac{1}{9}\right) = \frac{3}{4} \cdot \frac{8}{9} = \frac{2}{3}$, $P_4 = P_3\left(1 - \frac{1}{16}\right) = \frac{2}{3} \cdot \frac{15}{16} = \frac{5}{8}$, $P_5 = P_4\left(1 - \frac{1}{25}\right) = \frac{5}{8} \cdot \frac{24}{25} = \frac{3}{5}$, and $P_6 = P_5\left(1 - \frac{1}{36}\right) = \frac{3}{5} \cdot \frac{35}{36} = \frac{7}{12}$. It appears, from these first few values, that $P_n = \frac{n+1}{2n}$ for all $n \geq 2$. When $n = 2$ we have $\frac{n+1}{2n} = \frac{3}{4} = P_2$. Let $n \geq 3$ and suppose, inductively, that $P_{n-1} = \frac{n}{2(n-1)}$. Then

$$\begin{aligned} P_n &= P_{n-1}\left(1 - \tfrac{1}{n^2}\right) \text{ , by the recursion formula for } P_n \\ &= \tfrac{n}{2(n-1)} \cdot \tfrac{n^2-1}{n^2} \text{ , by the induction hypothesis} \\ &= \tfrac{n}{2(n-1)} \cdot \tfrac{(n-1)(n+1)}{n^2} = \tfrac{n+1}{2n} \text{ , as required.} \end{aligned}$$

By induction, it follows that $P_n = \frac{n+1}{2n}$ for all $n \geq 2$.

4.10 Exercise: Find $\sum_{k=1}^{n} k$ and find $\sum_{k=1}^{n} k^3$.

4.11 Exercise: Let $(a_n)_{n\geq 0}$ be the Fibonacci sequence.

(a) Show that $a_0 + a_1 + a_2 + \cdots + a_n = a_{n+2} - 1$ for all $n \geq 0$.
(b) Show that ${a_0}^2 + {a_1}^2 + {a_2}^2 + \cdots + {a_n}^2 = a_n a_{n+1}$ for all $n \geq 0$.
(c) Show that $a_{n-1}a_{n+1} = {a_n}^2 + (-1)^n$ for all $n \geq 1$.
(d) Show that ${a_{n-1}}^2 + {a_n}^2 = a_{2n-1}$ for all $n \geq 1$.

4.12 Theorem: *(Strong Mathematical Induction) Let $F(n)$ be a statement about $n \in \mathbf{Z}$ and let $m \in \mathbf{Z}$. Suppose that for all $n \in \mathbf{Z}$ with $n \geq m$, if $F(k)$ is true for all $k \in \mathbf{Z}$ with $m \leq k < n$ then $F(n)$ is true. Then $F(n)$ is true for all $n \geq m$.*

Proof: Let $G(\ell)$ be the statement "$F(k)$ is true for all $k \in \mathbf{Z}$ with $m \leq k < \ell$". Note that $G(m)$ is true vacuously because there are no values of $k \in \mathbf{Z}$ with $m \leq k < m$. Let $\ell \geq m$ and suppose, inductively, that $G(\ell)$ is true or, in other words, suppose that $F(k)$ is true for all $k \in \mathbf{Z}$ with $m \leq k < \ell$. Since $F(k)$ is true for all $k \in \mathbf{Z}$ with $m \leq k < \ell$, it follows from the hypothesis in the statement of the theorem that $F(\ell)$ is true. Since $F(k)$ is true for all $k \in \mathbf{Z}$ with $m \leq k < \ell$ and $F(\ell)$ is true, it follows that $F(k)$ is true for all $k \in \mathbf{Z}$ with $m \leq k < \ell + 1$ or, in other words, it follows that $G(\ell + 1)$ is true, as required. By induction, it follows that $G(\ell)$ is true for all $\ell \geq m$.

Let $n \geq m$ be arbitrary. Since $G(\ell)$ is true for all $\ell \geq m$, in particular $G(n + 1)$ is true, so $F(k)$ is true for all $k \in \mathbf{Z}$ with $m \leq k < n + 1$. Since $F(k)$ is true for all $k \in \mathbf{Z}$ with $m \leq k < n + 1$, in particular $F(n)$ is true. Since $n \geq m$ was arbitrary, it follows that $F(n)$ is true for all $n \geq m$.

4.13 Note: In order to prove that $F(n)$ is true for all $n \geq m$, we can do the following.

1. Let $n \geq m$ and suppose that $F(k)$ is true for all $k \in \mathbf{Z}$ with $m \leq k < n$.
2. Prove that $F(n)$ is true.

Although strong induction, used as above, does not require the proof that $F(m)$ is true (the base case), there are situations in which one ore more base cases must be verified to make this method of proof valid. For example, if a sequence $(x_n)_{n\geq 1}$ is defined by specifying the values of x_1 and x_2 and by giving a recursion formula for x_n in terms of x_{n-1} and x_{n-2} for all $n \geq 3$, then in order to prove that x_n satisfies the closed-form formula $x_n = f(n)$ for all $n \geq 1$ it suffices to prove that $x_1 = f(1)$ and $x_2 = f(2)$ (two base cases) and to prove that for all $n \geq 3$, if $x_{n-1} = f(n - 1)$ and $x_{n-2} = f(x_{n-2})$ then $x_n = f(n)$.

4.14 Example: Let $a_0 = a_1 = 2$ and let $a_n = 2a_{n-1} + 3a_{n-2}$ for all $n \geq 2$. Find a closed-form formula for a_n.

Solution: The first few values of a_n are as follows.

n	0	1	2	3	4	5
a_n	2	2	10	26	82	242

It appears that $a_n = 3^n + (-1)^n$ for all $n \geq 0$. Here is a proof by induction. When $n = 0$ we have $3^n + (-1)^n = 3^0 + (-1)^0 = 1 + 1 = 2 = a_0 = a_n$ and when $n = 1$ we have $3^n + (-1)^n = 3^1 + (-1)^1 = 3 - 1 = 2 = a_1 = a_n$. Let $n \geq 2$ and suppose, inductively, that $a_k = 3^k + (-1)^k$ for all $k \in \mathbf{Z}$ with $0 \leq k < n$ (in particular for $k = n - 1$ and $k = n - 2$). Then

$$\begin{aligned} a_n &= 2a_{n-1} + 3a_{n-2} \text{ , by the recursion formula for } a_n \\ &= 2\big(3^{n-1} + (-1)^{n-1}\big) + 3\big(3^{n-2} + (-1)^{n-2}\big) \text{ , by the induction hypothesis} \\ &= 2 \cdot 3^{n-1} - 2(-1)^n + 3^{n-1} + 3(-1)^n = 3^n + (-1)^n \text{ , as required.} \end{aligned}$$

By induction, it follows that $a_n = 3^n + (-1)^n$ for all $n \geq 0$.

4.15 Note: One shortcoming with the method that we used in the above example is that we needed to guess a closed-form formula for the sequence and, for many sequences, such a closed-form formula can be extremely difficult to guess. For this reason it is useful to develop a method which allows us to calculate such a closed-form formula.

4.16 Theorem: *(Quadratic Linear Recursion) Let $a, b, p, q \in \mathbf{R}$ with $q \neq 0$ and define $(x_n)_{n \geq m}$ recursively by $x_m = a$, $x_{m+1} = b$ and $x_n = px_{n-1} + qx_{n-2}$ for all $n \geq m+2$. Let $f(x) = x^2 - px - q$ and suppose that $f(x)$ factors as $f(x) = (x-u)(x-v)$ for some $u, v \in \mathbf{R}$ with $u \neq v$. Then there exist unique numbers $A, B \in \mathbf{R}$ such that $x_n = Au^n + Bv^n$ for all $n \geq m$.*

Proof: Since $x^2 - px - q = f(x) = (x-u)(x-v) = x^2 - (u+v)x + pq$ we have $u + v = p$ and $uv = -q$. Since $q \neq 0$ and $uv = -q$ it follows that $u \neq 0$ and $v \neq 0$.

In order to have $x_n = Au^n + Bv^n$ for all $n \geq m$ we must have $Au^m + Bv^m = x_m = a$ (1) and $Au^{m+1} + Bv^{m+1} = b$ (2). Multiplying Equation (1) by v and subtracting Equation (2) gives $A(vu^m - u^{m+1}) = av - b$, so we must choose $A = \frac{av-b}{u^m(v-u)}$. Multiplying Equation (1) by u and subtracting (2) gives $B(uv^m - v^{m+1}) = au - b$, so we must choose $B = \frac{au-b}{v^m(u-v)}$.

Let $A = \frac{av-b}{u^m(v-u)}$ and $B = \frac{au-b}{v^m(u-v)}$. We claim that $x_n = Au^n + Bv^n$ for all $n \geq m$. Here is a proof by induction. When $n = m$ we have

$$Au^n + Bv^n = Au^m + Bv^m = \tfrac{av-b}{v-u} + \tfrac{au-b}{u-v} = \tfrac{av-b-au+b}{v-u} = a = x_m = x_n$$

and when $n = m+1$ we have

$$Au^n + Bv^n = Au^{m+1} + Bv^{m+1} = \tfrac{(av-b)u}{v-u} + \tfrac{(au-b)v}{u-v} = \tfrac{auv-bu-auv+bv}{v-u} = b = x_{m+1} = x_n.$$

Let $n \geq m+2$ and suppose, inductively, that $x_k = Au^k + Bv^k$ for all $k \in \mathbf{Z}$ with $m \leq k < n$ (in particular for $k = n-1$ and $k = n-2$. Then

$$\begin{aligned}
x_n &= px_{n-1} + qx_{n-2} \text{ , by the recursion formula for } x_n \\
&= (u+v)x_{n-1} - (uv)x_{n-2} \text{ , since } u+v = p \text{ and } uv = -q \\
&= (u+v)\big(Au^{n-1} + Bv^{n-1}\big) - (uv)\big(Au^{n-2} + Bv^{n-2}\big) \text{ , by the induction hypothesis} \\
&= A\big((u+v)u^{n-1} - (uv)u^{n-2}\big) + B\big((u+v)v^{n-1} - (uv)v^{n-2}\big) \\
&= Au^n + Bv^n \text{ , as required.}
\end{aligned}$$

It follows, by induction, that $x_n = Au^n + Bv^n$ for all $n \geq m$.

4.17 Theorem: *(Linear Recursion) Let $a_0, a_1, a_2, \cdots, a_{d-1}$ and $c_0, c_1, c_2, \cdots, c_{d-1}$ be real (or complex) numbers with $c_0 \neq 0$. Let $(x_n)_{n \geq m}$ be the sequence defined recursively by $x_m = a_0$, $x_{m+1} = a_1$, $x_{m+2} = a_2$, $\cdots$, $x_{m+d-1} = a_{d-1}$ and*

$$x_n + c_{d-1}x_{n-1} + c_{d-2}x_{n-2} + \cdots + c_1x_{n-d+1} + c_0x_{n-d} = 0.$$

Let $f(x) = x^d + c_{d-1}x^{d-1} + c_{d-2}x^{d-2} + \cdots + c_1x + c_0$ and suppose that $f(x)$ factors as $f(x) = \prod_{i=1}^{\ell} (x-u_i)^{k_i}$ where the u_i are distinct real (or complex) numbers. Then there exist unique polynomials $p_1(x), p_2(x), \cdots, p_\ell(x)$, with each $p_i(x)$ of degree less than k_i, such that

$$x_n = \sum_{i=1}^{\ell} p_i(n){u_i}^n \text{ for all } n \geq m.$$

Proof: This is a stronger version of the previous theorem, but we omit the proof.

4.18 Example: Let $x_0 = 4$, $x_1 = -1$ and $x_n = 3x_{n-1} + 10x_{n-2}$ for $n \geq 2$. Find a closed-form formula for x_n.

Solution: The first few terms x_n are as follows.

n	0	1	2	3	4
4	-1	37	101	673	3029

It seems difficult to guess a closed-form formula for x_n from the information in the above table. Instead, we make use of the above theorem with $p = 3$ and $q = 10$. Let

$$f(x) = x^2 - px - q = x^2 - 3x - 10 = (x-5)(x+2).$$

From the above theorem, we know that for some $A, B \in \mathbf{R}$ we have $x_n = A(5)^n + B(-2)^n$ for all $n \geq 0$. In particular, we must have $A(5)^0 + B(-2)^0 = x_0$, that is $A + B = 4$ (1) and we must have $A(5)^1 + B(-2)^1 = x_1$, that is $5A - 2B = -1$ (2). Multiply Equation (1) by 2 and add Equation (2) to get $7A = 7$ so that $A = 1$, and multiply Equation (1) by 5 and subtract Equation (2) to get $7B = 21$ so that $B = 3$. Thus $x_n = 5^n + 3(-2)^n$ for all $n \geq 0$.

4.19 Exercise: Find a closed-form formula for the terms of the Fibonacci sequence.

4.20 Note: Suppose that we choose k of n objects, When the objects are chosen with replacement (so that repetition is allowed) and the order of the chosen objects matters (so the chosen objects form an ordered k-tuple), the number of ways to choose k of n objects is equal to n^k (since we have n choices for each of the k objects). For example, the number of ways to roll 3 six-sided dice is equal to $6^3 = 216$.

When the objects are chosen without replacement (so that the k chosen objects are distinct) and the order matters, the number of ways to choose k of n objects is equal to $n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!}$ (since we have n choices for the first object and $n-1$ choices for the second object and so on). In particular, the number of ways to arrange n objects in order (to form an ordered n-tuple) is equal to $n!$.

When the objects are chosen without replacement and the order does not matter (so the chosen objects form a k-element set), the number of ways to choose k of n objects is equal to $\frac{n!}{k!(n-k)!} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$ (since each k-element set can be ordered in $k!$ ways to form $k!$ ordered k-tuples, and there are $\frac{n!}{(n-k)!}$ such ordered k-tuples). For example, the number of 4-element subsets of the set $\{1,2,3,4,5,6,7\}$ is equal to $\frac{7!}{4!3!} = \frac{7\cdot6\cdot5\cdot4}{4\cdot3\cdot2\cdot1} = 7\cdot5 = 35$.

4.21 Definition: For $n, k \in \mathbf{N}$ with $0 \leq k \leq n$, we define the **binomial coefficient** $\binom{n}{k}$, read as "n choose k", by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}.$$

4.22 Theorem: *(Pascal's Triangle) For $k, n \in \mathbf{N}$ with $0 \leq k \leq n$ we have*

$$\binom{n}{0} = \binom{n}{n} = 1\,,\ \binom{n}{k} = \binom{n}{n-k} \text{ and } \binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}.$$

Proof: The formulas $\binom{n}{0} = \binom{n}{n} = 1$ and $\binom{n}{k} = \binom{n}{n-k}$ are immediate from the definition of $\binom{n}{k}$ (since $0! = 1$) and we have

$$\begin{aligned}\binom{n}{k} + \binom{n}{k+1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-k-1)!} = \frac{(k+1)n!}{(k+1)!(n-k)!} + \frac{(n-k)n!}{(k+1)!(n-k)!} \\ &= \frac{(k+1+n-k)n!}{(k+1)!(n-k)!} = \frac{(n+1)!}{(k+1)!((n+1)-(k+1))!} = \binom{n+1}{k+1}.\end{aligned}$$

4.23 Exercise: Make a table displaying the values $\binom{n}{k}$ for $0 \leq k \leq n \leq 10$. The table forms a triangle of positive integers in which each entry is obtained by adding two of the entries above.

4.24 Notation: Let R be a ring and let $a \in R$. For $k \in \mathbf{Z}^+$ we write $ka = a + a + \cdots + a$ with k terms in the sum, and we write $(-k)a = k(-a)$, and we write $a^k = a \cdot a \cdot \ldots \cdot a$ with k terms in the product. For $0 \in \mathbf{Z}$ we write $0a = 0$ and $a^0 = 1$. When $a \in R$ is a unit, for $k \in \mathbf{Z}^+$ we write $a^{-k} = (a^{-1})^k$.

4.25 Exercise: Let R be a ring and let $a, b \in R$. Show that for all $k, l \in \mathbf{Z}$ we have $(-k)a = -(ka)$, $(k+l)a = ka + la$ and $(ka)(lb) = (kl)(ab)$. Show that for all $k, l \in \mathbf{Z}^+$ we have $a^{k+l} = a^k a^l$. Show that if $ab = ba$ then for all $k, l \in \mathbf{Z}^+$ we have $(ab)^k = a^k b^k$. Show that if a is a unit, then for all $k, l \in \mathbf{Z}$ we have $a^{-k} = (a^k)^{-1}$ and $a^{k+l} = a^k a^l$.

4.26 Theorem: *(Binomial Theorem) Let R be a ring, let $a, b \in R$ with $ab = ba$, and let $n \in \mathbf{N}$. Then*

$$\begin{aligned}(a+b)^n &= \sum_{k=0}^{n} \tbinom{n}{k} a^k b^{n-k} = \sum_{k=0}^{n} \tbinom{n}{k} a^{n-k} b^k \\ &= \tbinom{n}{0} a^n + \tbinom{n}{1} a^{n-1} b + \tbinom{n}{2} a^{n-2} b^2 + \cdots + \tbinom{n}{n-1} a\, b^{n-1} + \tbinom{n}{n} b^n.\end{aligned}$$

Proof: We shall prove, by induction, that $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$ for all $n \geq 0$.

When $n = 0$ we have $\sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = \binom{0}{0} a^0 b^0 = 1 = (a+b)^0 = (a+b)^n$.

When $n = 1$ we have $\sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = \binom{1}{0} a^1 b^0 + \binom{1}{1} a^0 b^1 = a + b = (a+b)^1 = (a+b)^n$.

Let $n \geq 1$ and suppose, inductively that $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$. Then

$$\begin{aligned}(a+b)^{n+1} &= (a+b)(a+b)^n = (a+b) \sum_{k=0}^{n} \tbinom{n}{k} a^{n-k} b^k \\ &= (a+b)\Big(\tbinom{n}{0} a^n + \tbinom{n}{1} a^{n-1} b + \tbinom{n}{2} a^{n-2} b^2 + \cdots + \tbinom{n}{n-1} a\, b^{n-1} + \tbinom{n}{n} b^n \Big) \\ &= \tbinom{n}{0} a^{n+1} + \tbinom{n}{1} a^n b + \tbinom{n}{2} a^{n-1} b^2 + \cdots + \tbinom{n}{n-1} a^2\, b^{n-1} + \tbinom{n}{n} a\, b^n \\ &\qquad + \tbinom{n}{0} a^n b + \tbinom{n}{1} a^{n-1} b^2 + \cdots + \tbinom{n}{n-2} a^2 b^{n-1} + \tbinom{n}{n-1} a\, b^n + \tbinom{n}{n} b^{n+1} \\ &= a^{n+1} + \big(\tbinom{n}{0} + \tbinom{n}{1}\big) a^n b + \big(\tbinom{n}{1} + \tbinom{n}{2}\big) a^{n-1} b + \cdots \\ &\qquad + \big(\tbinom{n}{n-2} + \tbinom{n}{n-1}\big) a^2 b^{n-1} + \big(\tbinom{n}{n-1} + \tbinom{n}{n}\big) a\, b^n + b^{n+1} \\ &= \tbinom{n+1}{0} a^{n+1} + \tbinom{n+1}{1} a^n b + \tbinom{n+1}{2} a^{n-1} b^2 + \cdots + \tbinom{n+1}{n-1} a^2\, b^n + \tbinom{n+1}{n+1} a\, b^n \\ &= \sum_{k=0}^{n+1} \tbinom{n+1}{k+1} a^{n+1-k} b^k\end{aligned}$$

as required, since $\binom{n}{0} = 1 = \binom{n+1}{0}$ and $\binom{n}{n} = 1 = \binom{n+1}{n+1}$ and $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$ for all k with $0 \leq k \leq n$. By induction, we have $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$ for all $n \geq 0$.

Finally note that, by interchanging a and b, we also have $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$ for all $n \geq 0$.

4.27 Example: Expand $(x+2)^6$.

Solution: By the Binomial Theorem, we have

$$\begin{aligned}(x+2)^6 &= \tbinom{6}{0}x^6 + \tbinom{6}{1}x^5 2^1 + \tbinom{6}{2}x^4 2^2 + \tbinom{6}{3}x^3 2^3 + \tbinom{6}{4}x^2 2^4 + \tbinom{6}{5}x^1 2^5 + \tbinom{6}{6}2^6 \\ &= x^6 + 6\cdot 2\,x^5 + 15\cdot 4\,x^4 + 20\cdot 8\,x^3 + 15\cdot 16\,x^2 + 6\cdot 32\,x + 64 \\ &= x^6 + 12x^5 + 60x^4 + 160x^3 + 240x^2 + 192x + 64.\end{aligned}$$

4.28 Example: Find the coefficient of x^8 in the expansion of $\left(5x^3 - \frac{2}{x^2}\right)^{11}$.

Solution: By the Binomial Theorem, we have

$$\left(5x^3 - \tfrac{2}{x^2}\right)^{11} = \sum_{k=0}^{11} \tbinom{11}{k}(5x^3)^{11-k}(2x^{-2})^k = \sum_{k=0}^{11}(-1)^k\tbinom{11}{k}\,5^{11-k}2^k\,x^{3(11-k)-2k}.$$

In order to get $3(11-k)-2k = 8$, we need $33-5k=8$, so we choose the term with $k=5$. Thus the coefficient of x^8 is equal to

$$(-1)^5\tbinom{11}{5}\,5^6\,2^5 = -\tfrac{11\cdot 10\cdot 9\cdot 8\cdot 7}{5\cdot 4\cdot 3\cdot 2\cdot 1}\,5^6\,2^5 = -11\cdot 3\cdot 7\cdot 10^6 = -231{,}000{,}000.$$

4.29 Example: Find $\sum_{k=0}^{n}\binom{2n}{2k}\frac{1}{2^k}$.

Solution: Notice that

$$\left(1+\tfrac{1}{\sqrt{2}}\right)^{2n} = \tbinom{2n}{0}+\tbinom{2n}{1}\tfrac{1}{\sqrt{2}}+\tbinom{2n}{2}\tfrac{1}{2}+\tbinom{2n}{3}\tfrac{1}{2\sqrt{2}}+\tbinom{2n}{4}\tfrac{1}{2^2}+\cdots+\tbinom{2n}{2n-1}\tfrac{1}{2^{n-1}\sqrt{2}}+\tbinom{2n}{2n}\tfrac{1}{2^n}$$

$$\left(1-\tfrac{1}{\sqrt{2}}\right)^{2n} = \tbinom{2n}{0}-\tbinom{2n}{1}\tfrac{1}{\sqrt{2}}+\tbinom{2n}{2}\tfrac{1}{2}-\tbinom{2n}{3}\tfrac{1}{2\sqrt{2}}+\tbinom{2n}{4}\tfrac{1}{2^2}-\cdots-\tbinom{2n}{2n-1}\tfrac{1}{2^{n-1}\sqrt{2}}+\tbinom{2n}{2n}\tfrac{1}{2^n}$$

Add these and divide by 2 to get

$$\tfrac{1}{2}\left(\left(1+\tfrac{1}{\sqrt{2}}\right)^{2n}+\left(1-\tfrac{1}{\sqrt{2}}\right)^{2n}\right) = \tbinom{2n}{0}+\tbinom{2n}{2}\tfrac{1}{2}+\tbinom{2n}{4}\tfrac{1}{2^2}+\cdots+\tbinom{2n}{2n}\tfrac{1}{2^n} = \sum_{k=0}^{n}\tbinom{2n}{2k}\tfrac{1}{2^k}.$$

4.30 Exercise: Let n be a positive integer. By calculating $\sum_{k=0}^{n}\left((k+1)^{m+1}-k^{m+1}\right)$ in two different ways, find a recursion formula for the sum $S_m = \sum_{k=0}^{n} k^m$.

4.31 Exercise: There are n points on a circle around a disc. Each of the $\binom{n}{2}$ pairs of points is connected by a line segment. No three of these line segments have a common point of intersection. Determine the number of regions into which the disc is divided by the line segments.

4.32 Exercise: Let $n \in \mathbf{Z}^+$. Show that every positive real number has a unique positive n^{th} root. When n is odd, show that every real number has a unique real n^{th} root.

4.33 Notation: When $n \in \mathbf{Z}^+$ and $x \in \mathbf{R}$ (with $x \geq 0$ in the case that n is even) we denote the unique n^{th} root of x by $x^{1/n}$ or by $\sqrt[n]{x}$. In the case $n=2$ and $x \geq 0$, we also write $x^{1/2}$ as $\sqrt{x}$.

4.34 Exercise: (The Quadratic Formula) Show that for all $a, b, c, x \in \mathbf{R}$ with $a \neq 0$ we have

$$ax^2+bx+c = 0 \iff b^2-4ac \geq 0 \text{ and } x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}.$$

Be careful to notice (and prove) any familiar algebraic properties of $\mathbf{R}$, which you need to use in your proof, but which have not yet been mentioned in these course notes.

.

Chapter 5. Factorization of Integers

5.1 Definition: For $a, b \in \mathbf{Z}$ we say that a **divides** b (or that a is a **factor** of b, or that b is a **multiple of** a**), and we write** $a\big|b$**, when** $b = ak$ **for some** $k \in \mathbf{Z}$**.**

5.2 Theorem: *(Basic Properties of Divisors) Let* $a, b, c \in \mathbf{Z}$*. Then*

(1) $a|0$ *for all* $a \in \mathbf{Z}$ *and* $0|a \Longleftrightarrow a = 0$*,*
(2) $a|1 \Longleftrightarrow a = \pm 1$ *and* $1|a$ *for all* $a \in \mathbf{Z}$*.*
(3) If $a|b$ *and* $b|c$ *then* $a|c$*.*
(4) If $a|b$ *and* $b|a$ *then* $b = \pm a$*.*
(5) If $a|b$ *then* $|a| \leq |b|$*.*
(6) If $a|b$ *and* $a|c$ *then* $a\big|(bx + cy)$ *for all* $x, y \in \mathbf{Z}$*.*

Proof: Some of these properties hold in all rings while other properties are specific to $\mathbf{Z}$. Property (1) holds in all rings because when R is a ring and $a \in R$ we have $a \cdot 0 = 0$. Part, but not all, of Property (2) also holds in all rings. In any ring R we have $1|a$ for all $a \in R$ because $1 \cdot a = a$. Also in any ring R, because $1 \cdot 1 = 1$ and $(-1)(-1) = 1$ it follows that if $a = \pm 1$ then $a|1$. However, it is not the case that in every ring R and for all $a \in R$, if $a|1$ then $a = \pm 1$. For example, in the ring $\mathbf{Z}_8$ we have $1^2 = 3^2 = 5^2 = 7^2 = 1$ so that $1|1$, $3|1$, $5|1$ and $7|1$ (or equivalently $\pm 1|1$ and $\pm 3|1$). Let us prove that for all $a \in \mathbf{Z}$, if $a|1$ then $a = \pm 1$. Let $a \in \mathbf{Z}$ be arbitrary. Suppose that $a|1$. Choose $k \in \mathbf{Z}$ such that $1 = a \cdot k$. We claim that because $a \cdot k = 1$ it follows that either $a = k = 1$ or $a = k = -1$. Either $a < -1$ or $a = -1$ or $a = 0$ or $a = 1$ or $a > 1$. If $a > 1$ then

5.3 Theorem: *(The Division Algorithm) Let* $a, b \in \mathbf{Z}$ *with* $b \neq 0$*. Then there exist unique integers* q *and* r *such that*

$$a = qb + r \text{ and } 0 \leq r < |b|.$$

The integers q *and* r *are called the* **quotient** *and* **remainder** *when* a *is divided by* b*.*

Proof: We begin by proving that such integers q and r exist. Later we will show that the values of q and r are unique. Case 1: suppose that $b > 0$ and $a \geq 0$. Consider the sequence $0, b, 2b, 3b, \cdots$. Eventually the terms in the sequence become larger than a. Choose $q \geq 0$ so that $qb \leq a$ and $(q+1)b > a$. Let $r = a - qb$ so that $b = qa + r$. Since $qb \leq a$ we have $r = a - qb \geq 0$. Since $(q+1)b > a$, we have $qb + b > a$, and so $r = a - qb < b = |b|$.

Case 2: suppose that $b > 0$ and $a < 0$. Consider the sequence $0, -b, -2b, -3b, \cdots$. Eventually the terms in the sequence become smaller than a. Choose $p \geq 0$ so that $-(p-1)b > a$ and $-pb \leq a$. Let $q = -p$ so that $(q+1)b > a$ and $qb \leq a$. As above, we let $r = a - qb$ to get $a = qb + r$ and $0 \leq r < b = |b|$.

Case 3: suppose that $b < 0$ and $a \in \mathbf{Z}$. By the above two paragraphs, we can choose integers p and r so that $a = p|b| + r = -pb + r$ with $0 \leq r < |b|$, then we let $q = -p$ so that $a = qb + r$. In all three cases, we have shown that there exist integers q and r such that $a = qb + r$ and $0 \leq r < |b|$.

It remains to verify that the values of q and r are unique. Suppose that $a = qb + r$ with $0 \leq r < |b|$ and $a = pb + s$ with $0 \leq s < |b|$. Suppose, for a contradiction, that $r \neq s$ and say $r < s$ so that we have $0 \leq r < s < |b|$. Since $a = qb + r = pb + s$ we have $s - r = qb - pb = (q - p)b$ so that $b|(s - r)$. Since $b\big|(s - r)$ we have $|b| \leq |s - r| = s - r$ (by one of the basic properties of divisors). But since $s < |b|$ and $r \geq 0$ we have $s - r < |b|$ giving the desired contradiction. Thus we have $r = s$. Since $r = s$ and $s - r = (q - p)b$ we have $0 = (q - p)b$ hence $p = q$ (since $b \neq 0$).

5.4 Note: For $a, b \in \mathbf{Z}$, when we write $a = qb + r$ with $q, r \in \mathbf{Z}$ and $0 \leq r < |b|$, we have $b|a$ if and only if $r = 0$. Indeed if $r = 0$ then $a = qb$ so that $b|a$ and, conversely, if $b|a$ with say $a = pb = pb + 0$, then we must have $q = p$ and $r = 0$ by the uniqueness of the quotient and remainder.

5.5 Definition: Let $a, b \in \mathbf{Z}$. A **common divisor** of a and b is an integer d such that $d|a$ and $d|b$. When a and b are not both 0, we denote the **greatest common divisor** of a and b by $\gcd(a, b)$. For convenience, we also define $\gcd(0, 0) = 0$.

5.6 Theorem: *(Basic Properties of the Greatest Common Divisor) Let* $a, b, q, r \in \mathbf{Z}$.
(1) $\gcd(a, b) = \gcd(b, a)$.
(2) $\gcd(a, b) = \gcd(|a|, |b|)$.
(3) If $a|b$ *then* $\gcd(a, b) = |a|$. *In particular,* $\gcd(a, 0) = |a|$.
(4) If $b = qa + r$ *then* $\gcd(a, b) = \gcd(a, r)$.

Proof: The proof is left as an exercise.

5.7 Theorem: *(The Euclidean Algorithm With Back-Substitution) Let* a *and* b *be integers and let* $d = \gcd(a, b)$. *Then there exist integers* s *and* t *such that* $as + bt = d$. *The proof provides explicit procedures for finding* d *and for finding* s *and* t.

Proof: We can find d using the following procedure, called the **Euclidean Algorithm**. If $b|a$ then we have $d = |b|$. Otherwise, let $r_{-1} = a$ and $r_0 = b$ and use the division algorithm repeatedly to obtain integers q_i and r_i such that

$$\begin{array}{cc}
r_{-1} = a = q_1 b + r_1 & 0 < r_1 < |a| \\
r_0 = b = q_2 r_1 + r_2 & 0 < r_2 < r_1 \\
r_1 = q_3 r_2 + r_3 & 0 < r_3 < r_2 \\
\vdots & \vdots \\
r_{k-2} = q_k r_{k-1} + r_k & 0 < r_k < r_{k-1} \\
\vdots & \vdots \\
r_{n-2} = q_n r_{n-1} + r_n & 0 < r_n < r_{n-1} \\
r_{n-1} = q_{n+1} r_n + r_{n+1} & r_{n+1} = 0\,.
\end{array}$$

Since $r_{n-1} = q_{n+1} r_n$ we have $r_n \big| r_{n-1}$ so $\gcd(r_{n-1}, r_n) = r_n$. Since $r_{k-2} = q_k r_{k-1} + r_k$ we have $\gcd(r_{k-2}, r_{k-1}) = \gcd(r_{k-1}, r_k)$ and so

$$d = \gcd(a, b) = \gcd(b, r_1) = \gcd(r_1, r_2) = \cdots = \gcd(r_{n-2}, r_{n-1}) = \gcd(r_{n-1}, r_n) = r_n\,.$$

Having found d using the Euclidean algorithm, as above, we can find s and t using the following procedure, which is known as **Back-Substitution**. If $b|a$ so that $d = |b|$, then we can take $s = 0$ and $t = \pm 1$ to get $as + bt = d$. Otherwise, we let

$$s_0 = 1 \ , \ s_1 = -q_n \ , \text{ and } s_{\ell+1} = s_{\ell-1} - q_{n-\ell} s_\ell \text{ for } 1 \le \ell \le n-1$$

and then we can take $s = s_{n-1}$ and $t = s_n$ to get $as + bt = d$, because, writing $k = n - \ell$,

$$\begin{aligned}
d = r_n &= r_{n-2} - q_n r_{n-1} = s_1 r_{n-1} + s_0 r_{n-2} \\
&\ \ \vdots \\
&= \cdots = s_\ell r_{n-\ell} + s_{\ell-1} r_{n-\ell-1} = s_{n-k} r_k + s_{n-k-1} r_{k-1} \\
&= s_{n-k}\big(r_{k-2} - q_k r_{k-1}\big) + s_{n-k-1} r_{k-1} = \big(s_{n-k-1} - q_k s_{n-k}\big) r_{k-1} + s_{n-k} r_{k-2} \\
&= (s_{\ell-1} - q_{n-\ell} s_\ell) r_{n-\ell-1} + s_\ell r_{n-\ell-2} = s_{\ell+1} r_{n-\ell-1} + s_\ell r_{n-\ell-2} \\
&\ \ \vdots \\
&= \cdots = s_n r_0 + s_{n-1} r_{-1} = s_n b + s_{n-1} a \, .
\end{aligned}$$

5.8 Example: Let $a = 5151$ and $b = 1632$. Find $d = \gcd(a, b)$ and then find integers s and t so that $as + bt = d$.

Solution: The Euclidean Algorithm gives

$$\begin{aligned}
5151 &= 3 \cdot 1632 + 255 \\
1632 &= 6 \cdot 255 + 102 \\
255 &= 2 \cdot 102 + 51 \\
102 &= 2 \cdot 51 + 0
\end{aligned}$$

so $d = 51$. Using the quotients $q_1 = 3$, $q_2 = 6$ and $q_3 = 2$, Back-Substitution gives

$$\begin{aligned}
s_0 &= 1 \\
s_1 &= -q_3 = -2 \\
s_2 &= s_0 - q_2 s_1 = 1 - 6(-2) = 13 \\
s_3 &= s_1 - q_1 s_2 = -2 - 3(13) = -41 \, ,
\end{aligned}$$

so we take $s = s_2 = 13$ and $t = s_3 = -41$. (It is a good idea to check that indeed we have $(1632)(-41) + (5151)(13) = 51$).

5.9 Example: Let $a = 754$ and $b = -3973$. Find $d = \gcd(a, b)$ then find integers s and t such that $as + bt = d$.

Solution: The Euclidean Algorithm gives

$3973 = 5{\cdot}754{+}203$, $754 = 3{\cdot}203{+}145$, $203 = 1{\cdot}145{+}58$, $145 = 2{\cdot}58{+}29$, $58 = 2{\cdot}29{+}0$

so that $d = 29$. Then Back-Substitution gives rise to the sequence

$$1 \ , \ \ -2 \ , \ \ 3 \ , \ \ -11 \ , \ \ 58$$

so we have $(754)(58) + (3973)(-11) = 29$, that is $(754)(58) + (-3973)(11) = 29$. Thus we can take $s = 58$ and $t = 11$.

5.10 Theorem: *(More Properties of the Greatest Common Divisor) Let $a, b, c, d \in \mathbf{Z}$.*
(1) If $c|a$ and $c|b$ then $c \Big| \gcd(a, b)$.
(3) We have $\gcd(a, b) = 1$ if and only if there exist $x, y \in \mathbf{Z}$ such that $ax + by = 1$.
(4) If $d = \gcd(a, b) \neq 0$ then $\gcd\left(\frac{a}{d}, \frac{b}{d}\right) = 1$.
(5) If $a|bc$ and $\gcd(a, b) = 1$ then $a|c$.

Proof: We prove Part (5). Suppose that $a|bc$ and $\gcd(a, b) = 1$. Since $a|bc$ we can choose $k \in \mathbf{Z}$ so that $bc = ak$. Since $\gcd(a, b) = 1$, by the Euclidean Algorithm with Back-Substitution, we can choose $s, t \in \mathbf{Z}$ with $as + bt = 1$. Then we have

$$c = c \cdot 1 = c(as + bt) = acs + bct = acs + akt = a(cs + kt),$$

and so $a|c$, as required.

5.11 Definition: A **diophantine equation** is a polynomial equation in which the variables represent integers. Some diophantine equations are fairly easy to solve while others can be extremely difficult.

5.12 Theorem: *(Linear Diophantine Equations) Let $a, b, c \in \mathbf{Z}$ with $(a, b) \neq (0, 0)$. Let $d = \gcd(a, b)$ and note that $d \neq 0$. Consider the Diophantine equation $ax + by = c$.*
(1) The equation has a solution $(x, y) \in \mathbf{Z}^2$ if and only if $d|c$, and
(2) if $(u, v) \in \mathbf{Z}^2$ is one solution to the equation then the general solution is given by

$$(x, y) = (u, v) + k\big(-\tfrac{b}{d}, \tfrac{a}{d}\big) \text{ for some } k \in \mathbf{Z}.$$

Proof: Suppose that the equation $ax+by = c$ has a solution $(x, y) \in \mathbf{Z}^2$. Choose $(s, t) \in \mathbf{Z}^2$ so that $as + bt = c$. Since $d|a$ and $d|b$, it follows that $d|(ax + by)$ for all $x, y \in \mathbf{Z}$, so in particular $d|(as+bt)$, that is $d|c$. Conversely, suppose that $d|c$, say $c = d\ell$ with $\ell \in \mathbf{Z}$. Use the Euclidean Algorithm with Back-Substitution to find $s, t \in \mathbf{Z}$ such that $as + bt = d$. Multiply by ℓ to get $a(s\ell) + b(t\ell) = d\ell = c$. Thus we can take $x = e\ell$ and $y = t\ell$ to obtain a solution $(x, y) \in \mathbf{Z}^2$ to the equation $ax + by = c$. This proves Part (1)

Now suppose that $(u, v) \in \mathbf{Z}^2$ is a solution to the given equation, so we have $au+bv = c$. To prove Part (2), we need to prove that for all $k \in \mathbf{Z}$, if we let $(x, y) = (u, v) + k\big(-\frac{b}{d}, \frac{a}{d}\big)$ then (x, y) is a solution to $ax + by = c$ and, conversely, that if (x, y) is a solution then there exists $k \in \mathbf{Z}$ such that $(x, y) = (u, v) + k\big(-\frac{b}{d}, \frac{a}{d}\big)$.

Let $k \in \mathbf{Z}$ and let $(x, y) = (u, v) + k\big(-\frac{b}{d}, \frac{a}{d}\big)$. Then $x = u - \frac{kb}{d}$ and $y = v + \frac{ka}{d}$ and so

$$ax + by = a\big(u - \tfrac{kb}{d}\big) + b\big(v + \tfrac{ka}{d}\big) = (au + bv) - \tfrac{kab}{d} + \tfrac{kab}{d} = au + bv = c.$$

Conversely, let (x, y) be a solution to the given equation, so we have $ax + by = c$. Suppose that $a \neq 0$ (we leave the case $a = 0$ as an exercise). Since $ax + by = c$ and $au + bu = c$ we have $ax + by = au + bv$ and so $a(x - u) = -b(y - v)$. Divide both sides by d to get $\frac{a}{d}(x - u) = -\frac{b}{d}(y - v)$. Since $\frac{a}{d} \Big| \frac{b}{d}(y - v)$ and $\gcd\left(\frac{a}{d}, \frac{b}{d}\right) = 1$, it follows that $\frac{a}{d} \Big| (y - v)$. Choose $k \in \mathbf{Z}$ so that $y - v = \frac{ka}{d}$. Since $a \neq 0$ and $a(x - u) = -b(y - v) = -\frac{kab}{d}$, we have $x - u = -\frac{kb}{d}$ and so $(x, y) = (u, v) + k\big(-\frac{b}{d}, \frac{a}{d}\big)$, as required.

5.13 Example: Let $a = 426$, $b = 132$ and $c = 42$. Find all $x, y \in \mathbf{Z}$ such that $ax + by = c$.

Solution: The Euclidean Algorithm gives

$$426 = 3 \cdot 132 + 30 \ , \ 132 = 4 \cdot 30 + 12 \ , \ 30 = 2 \cdot 12 + 6 \ , \ 12 = 2 \cdot 6 + 0$$

so that $d = \gcd(a, b) = 6$. Note that $d|c$, indeed $c = d\ell$ with $\ell = 7$, so a solution does exist. Back-Substitution gives the sequence

$$1 \ , \ -2 \ , \ 9 \ , \ -29$$

so we have $a(9) + b(-29) = d$. Multiply by $\ell = 7$ to get $a(63) + b(-203) = c$, so one solution is given by $(x, y) = (63, -203)$. Since $\frac{a}{d} = \frac{426}{6} = 71$ and $\frac{b}{d} = \frac{132}{6} = 22$, The general solution is $(x, y) = (63, -203) + k(-22, 71)$.

5.14 Exercise: Let $a = 4123$, $b = 17689$ and $c = 798$. Find all $x, y \in \mathbf{Z}$ with $0 \leq y \leq 100$ such that $ax + by = c$.

5.15 Example: A **Pythagorean triple** is a solution (x, y, z) with $x, y, z \in \mathbf{Z}^+$ to the equation $x^2 + y^2 = z^2$. Note that when (x, y, z) is a Pythagorean triple with $z \neq 0$, we have $\left(\frac{x}{z}\right)^2 + \left(\frac{y}{z}\right)^2 = 1$ so that the point $\left(\frac{x}{z}, \frac{y}{z}\right)$ is a point on the unit circle with rational coordinates. Let S be the unit circle $\{(x, y) | x^2 + y^2 = 1\}$ and let $T = S \setminus \{(0, 1)\}$. The **stereographic projection** from T to $\mathbf{R}$ is the function $F : T \to \mathbf{R}$ defined as follows: given $(x, y) \in T$, let L be the line through $(0, 1)$ and (x, y), and define $F(x, y) = u$ where $(u, 0)$ is the point of intersection of L with the x-axis. Let us find a formula for F and a formula for its inverse $G : \mathbf{R} \to T$.

Given $(x, y) \in T$, the line L from $(0, 1)$ to (x, y) is given parametrically by $(u, v) = (0, 1) + t((x, y) - (0, 1)) = (tx, 1 + t(y - 1))$. This line meets the x-axis when $0 = v = 1 + t(y - 1)$, that is when $t = \frac{1}{1-y}$, and the resulting point of intersection is at $(u, v) = (tx, 1 + t(y - 1)) = \left(\frac{x}{1-y}, 0\right)$. Thus the map F is given by $u = F(x, y) = \frac{x}{1-y}$.

Given a point $(u, 0)$ on the x-axis, the line M through $(0, 1)$ and $(u, 0)$ is given parametrically by $(x, y) = (0, 1) + t((u, 0) - (0, 1)) = (tu, 1 - t)$. The point $(x, y) = (tu, 1 - t)$ lies on S when $1 = x^2 + y^2 = (tu)^2 + (1 - t)^2 = t^2u^2 + 1 - 2t + t^2$, that is when $(u^2 + 1)t^2 = 2t$, or equivalently when $t = 0$ or $t = \frac{2}{u^2+1}$. When $t = 0$ the resulting point is $(x, y) = (tu, 1 - t) = (0, 1)$ and when $t = \frac{2}{u^2+1}$ the resulting point is $(x, y) = (tu, 1 - t) = \left(\frac{2u}{u^2+1}, \frac{u^2-1}{u^2+1}\right)$. Thus the inverse map G is given by $(x, y) = G(u) = \left(\frac{2u}{u^2+1}, \frac{u^2-1}{u^2+1}\right)$.

Notice that if $(x, y) \in T$ with $x, y \in \mathbf{Q}$ then $u = F(x, y) \in \mathbf{Q}$ and that, conversely, if $u \in \mathbf{Q}$ then $(x, y) = G(u) \in \mathbf{Q}^2$. It follows that we have a bijective correspondence between $T \cap \mathbf{Q}^2$ and $\mathbf{Q}$ given by $F : T \cap \mathbf{Q}^2 \to \mathbf{Q}$ and $G : \mathbf{Q} \to T \cap \mathbf{Q}^2$. Thus every element in $T \cap \mathbf{Q}^2$ is of the form

$$G\left(\tfrac{s}{t}\right) = \left(\tfrac{2(s/t)}{(s/t)^2+1}, \tfrac{(s/t)^2-1}{(s/t)^2+1}\right) = \left(\tfrac{2st}{s^2+t^2}, \tfrac{s^2-t^2}{s^2+t^2}\right)$$

for some $s, t \in \mathbf{Z}$ with $t \neq 0$ and $\gcd(s, t) = 1$.

After doing some additional work (which involves considering the case in which $s + t$ is even and the case in which $s + t$ is odd, and in the former case replacing s and t by $s' = \frac{s+t}{2}$ and $t' = \frac{s-t}{2}$) one can verify that every Pythagorean triple (x, y, z), after possibly interchanging x and y, is of the form $r\left(2st \, , \, s^2 - t^2 \, , \, s^2 + t^2\right)$ for some $r, s, t \in \mathbf{Z}$ with $\gcd(s, t) = 1$ and $s + t$ odd.

5.16 Definition: Let n be a positive integer. We say that n is a **prime number** when $n \geq 2$ and n has no factor $a \in \mathbf{Z}$ with $1 < a < n$. We say that n is **composite** when $n \geq 2$ and n is not prime, that is when n does have a factor $a \in \mathbf{Z}$ with $1 < a < n$.

5.17 Theorem: *(Basic Properties of Primes) Let p be a prime number.*

(1) For all $a \in \mathbf{Z}$ we have $\gcd(a, p) \in \{1, p\}$ with $\gcd(a, p) = p$ if and only if $p|a$.

(2) For all $a, b \in \mathbf{Z}$, if $p|ab$ then either $p|a$ or $p|b$.

Proof: The proof is left as an exercise. Part (2) follows from Part (5) of Theorem 5.10.

5.18 Theorem: *Every integer $n \geq 2$ has a prime factor. Every composite integer $n \geq 2$ has a prime factor p with $p \leq \sqrt{n}$.*

Proof: Let $n \geq 2$. Suppose, inductively, that every integer k with $2 \leq k < n$ has a prime factor. If n is prime, then n is a prime factor of itself, so n has a prime factor. Suppose that n is composite. Let a be a factor of n with $1 < a < n$. By the induction hypothesis, a has a prime factor. Let p be a prime factor of a. Since $p|a$ and $a|n$ we have $p|n$, and so p is a prime factor of n. It follows, by induction, that every integer $n \geq 2$ has a prime factor.

Now suppose that n is composite. Write $n = ab$ where $a, b \in \mathbf{Z}$ with $1 < a \leq b < n$. Note that $a \leq \sqrt{n}$ because if we had $a > \sqrt{n}$ then we would also have $b \geq a > \sqrt{n}$ so that $n = ab > \sqrt{n}\sqrt{n} = n$ which is impossible. Let p be a prime factor of a. Since $p|a$ and $a|n$ we have $p|n$ so that p is a prime factor of n. Since $p|a$ and $a \leq \sqrt{n}$ we have $p \leq a \leq \sqrt{n}$.

5.19 Note: Given an integer $n \geq 2$, we can list all primes p with $p \leq n$ using the following procedure, which is called the **Sieve of Eratosthenes**. We begin by listing all the integers from 1 to n, and we cross off the number 1 (1 is a unit; it is not a prime). We circle the smallest remaining number p_1 (namely $p_1 = 2$, which is prime) then we cross off all other multiples of p_1 (which are composite). We circle the smallest remaining number p_2 (namely $p_2 = 3$, which is prime) then we cross off all other multiples of p_2 (which are all composite). At the k^{th} step of the procedure, when we circle the smallest remaining number p_k, it must be prime because if p_k was composite then it would have a prime factor p_i with $p_i < p_k$, but we have already found all primes $p_i < p_k$ and we have already crossed off all their multiples. We continue the procedure until we have circled a prime p_ℓ with $p_\ell \geq \sqrt{n}$ and crossed off its multiples. At this stage we circle all of the remaining numbers in the list because they are all prime. Indeed, if a remaining number m was composite then it would have a prime factor p with $p \leq \sqrt{m} \leq \sqrt{n}$, but we have already found all primes p with $\leq \sqrt{n}$ and crossed off all their multiples.

5.20 Exercise: Use the Sieve of Eratosthenes to list all primes p with $p \leq 100$.

5.21 Theorem: *(Euclid) There exist infinitely many prime numbers.*

Proof: Suppose, for a contradiction, that there exist finitely many prime numbers. Let $p_1, p_2, \cdots, p_\ell$ be all of the prime numbers. Consider the number $n = p_1 p_2 \cdots p_\ell + 1$. By Theorem 5.11, the number n has a prime factor and so $p_k|n$ for some index k. But p_k is not a factor of n because when we write $n = qp_k + r$ as in the Division Algorithm, we find that the remainder is $r = 1 \neq 0$ (and the quotient is $q = \prod_{i \neq k} p_i$).

5.22 Example: Note that there exist arbitrarily large gaps between consecutive prime numbers because, given a positive integer $m \geq 2$, we have $2\big|(m!+2)$, $3\big|(m!+3)$, $4\big|(m!+4)$ and so on, so the consecutive numbers $m!+2$, $m!+3$, $m!+4$, $\cdots$, $m!+m$ are all composite.

5.23 Remark: Here are a few facts about prime numbers which are difficult to prove.

(1) Bertrand's Postulate: for every integer $n \geq 1$ there exists a prime p with $n \leq p \leq 2n$.

(2) Dirichlet's Theorem: for all positive integers a, b with $\gcd(a, b) = 1$, there exist infinitely many primes of the form $p = a + kb$ for some $k \in \mathbf{N}$.

(3) The Prime Number Theorem: for $x \in \mathbf{R}$, let $\pi(x)$ be the number of primes p with $p \leq x$. Then $\lim_{x\to\infty} \dfrac{\pi(x)}{x/\ln x} = 1$.

5.24 Remark: Here are a few statements about prime numbers which are conjectured to be true, but for which no proof has, as yet, been found.

(1) Legendre's Conjecture: for every $n \in \mathbf{Z}^+$ there exists a prime p with $n^2 \leq p \leq (n+1)^2$.

(2) Goldbach's Conjecture: every even integer $n \geq 4$ is the sum of two prime numbers.

(3) Twin Primes Conjecture: there exist infinitely many p for which p and $p+2$ are prime.

(4) The $n^2 + 1$ Conjecture: there exist infinitely many primes $p = n^2 + 1$ with $n \in \mathbf{Z}^+$.

(5) Mersenne Primes Conjecture: there exist infinitely many primes $p = 2^k - 1$ with $k \in \mathbf{Z}^+$.

(6) Fermat Primes Conjecture: there exist finitely many primes $p = 2^k + 1$ with $k \in \mathbf{N}$.

5.25 Theorem: *(The Unique Factorization Theorem) Every integer $n \geq 2$ can be written uniquely in the form $n = \prod_{k=1}^{\ell} p_k = p_1 p_2 \cdots p_\ell$ where $\ell \in \mathbf{Z}^+$ and the p_k are primes with $p_1 \leq p_2 \leq \cdots \leq p_\ell$.*

Proof: First we prove the existence of such a factorization. Let n be an integer with $n \geq 2$ and suppose, inductively, that every integer k with $2 \leq k < n$ can be written in the required form. If n is prime then we can write $n = \prod_{k=1}^{\ell} p_k = p_1$ with $\ell = 1$ and $p_1 = n$. Suppose that n is composite. Write $n = ab$ where $a, b \in \mathbf{Z}$ with $1 < a < n$ and $1 < b < n$. By the induction hypothesis, we can write $a = q_1 q_2 \cdots q_\ell$ and $b = r_1 r_2 \cdots r_m$ where $\ell, m \in \mathbf{Z}^+$ and the p_i and q_i are primes with $p_1 \leq p_2 \leq \cdots \leq p_\ell$ and $q_1 \leq q_2 \leq \cdots \leq q_m$. Then $n = q_1 q_2 \cdots q_\ell r_1 r_2 \cdots r_m = p_1 p_2 \cdots p_{\ell+m}$ where the ordered $(\ell+m)$-tuple $(p_1, p_2, \cdots, p_{\ell+m})$ is obtained from the ordered $(\ell+m)$-tuple $(q_1, q_2, \cdots, q_\ell, r_1, r_2, \cdots, r_m)$ by rearranging the terms into non-decreasing order.

Let us prove uniqueness. Suppose that $n = p_1 p_2 \cdots p_\ell = q_1 q_2 \cdots q_m$ where $\ell, m \in \mathbf{Z}^+$ and the p_i and q_j are primes with $p_1 \leq p_2 \leq \cdots \leq p_\ell$ and $q_1 \leq q_2 \leq \cdots \leq q_m$. We need to prove that $\ell = m$ and that $p_i = q_i$ for every index i. Since $n = p_1 p_2 \cdots p_\ell$ we see that $p_1 | n$ and so $p \big| q_1 q_2 \cdots q_m$. By applying Part (2) of Theorem 5.12 repeatedly, it follows that $p_1 \big| q_i$ for some index i. Since $p_1 | q_i$ and q_i is prime, we must have $p_1 \in \{\pm 1, \pm q_i\}$. Since p_1 is prime, we have $p_1 > 1$. Since $p_1 > 1$ and $p_1 \in \{\pm 1, \pm q_i\}$ it follows that $p_1 = q_i$. A similar argument shows that $q_1 = p_j$ for some index j. Since $p_1 = q_i \geq q_1 = p_j \geq p_1$, it follows that $p_1 = q_1$.

Since $p_1 p_2 \cdots p_\ell = q_1 q_2 \cdots q_m$ and $p_1 = q_1$, we can divide both sides by p_1 to get $p_2 p_3 \cdots p_\ell = q_2 q_3 \cdots q_m$. By repeating the above argument, we can show that $p_2 = q_2$, then we can divide both sides by $p_2 = q_2$ to get $p_3 \cdots p_\ell = q_3 \cdots q_m$ and so on.

If we had $\ell \neq m$, say $\ell < m$, repeating the above procedure would eventually yield $p_\ell = q_\ell q_{\ell+1} \cdots q_m$ with $p_\ell = q_\ell$ and then $1 = q_{\ell+1} \cdots q_m$ which is not possible since each $q_i > 1$. Thus we must have $\ell = m$ and repeating the above procedure gives $p_i = q_i$ for all indices i, as required.

5.26 Note: Here are two alternate ways of expressing the above theorem.

(1) Every integer $n \geq 2$ can be written uniquely in the form $n = \prod_{i=1}^{\ell} {p_i}^{m_i} = {p_1}^{m_1} \cdots {p_\ell}^{m_\ell}$ where $\ell \in \mathbf{Z}^+$ and the p_i are distinct primes with $p_1 < p_2 < \cdots < p_\ell$ and each $m_i \in \mathbf{Z}^+$.

(2) Given distinct primes $p_1, p_2, \cdots, p_\ell$, every $n \in \mathbf{Z}^+$ whose prime factors are included in $\{p_1, \cdots, p_\ell\}$ can be written uniquely in the form $n = \prod_{i=1}^{\ell} {p_i}^{m_i} = {p_1}^{m_1} \cdots {p_\ell}^{m_\ell}$ with $m_i \in \mathbf{N}$.

5.27 Theorem: *(Unique Factorization and Divisors) Let $n = {p_1}^{m_1}{p_2}^{m_2} \cdots {p_\ell}^{m_\ell}$ where $\ell \in \mathbf{Z}^+$, the p_i are distinct primes, and each $m_i \in \mathbf{N}$. Then the positive divisors of n are the numbers of the form $a = {p_1}^{j_1}{p_2}^{j_2} \cdots {p_\ell}^{j_l}$ where each $j_i \in \mathbf{Z}$ with $0 \leq j_i \leq m_i$.*

Proof: Suppose that $n = {p_1}^{m_1}{p_2}^{m_2} \cdots {p_\ell}^{m_\ell}$ and $a = {p_1}^{j_1}{p_2}^{j_2} \cdots {p_\ell}^{j_\ell}$ where $p_1, p_2, \cdots, p_\ell$ are distinct primes and $0 \leq j_i \leq m_i$ for all indices i. Let $b = {p_1}^{k_1}{p_2}^{k_2} \cdots {p_\ell}^{k_\ell}$ where $k_i = m_i - j_i$ (note that $k_i \geq 0$ since $j_i \leq m_i$). Then

$$ab = ({p_1}^{j_1} \cdots {p_\ell}^{j\ell})({p_1}^{k_1} \cdots {p_\ell}^{k_\ell}) = {p_1}^{j_1+k_1} \cdots {p_\ell}^{j_\ell+k_\ell} = {p_1}^{m_1} \cdots {p_\ell}^{m_\ell} = n$$

and so $a|n$.

Conversely, suppose that $n = {p_1}^{m_1}{p_2}^{m_2} \cdots {p_\ell}^{m_\ell}$, as above, and let a be a positive divisor of n. Let p be any prime factor of a. Since $p|a$ and $a|n$ we have $p|n$. Since $p|n$ and $n = {p_1}^{m_1}{p_2}^{m_2} \cdots {p_\ell}^{m_\ell}$ we have $p|p_i$ for some index i. Since p and p_i are both prime and $p|p_i$, we have $p = p_i$. This proves that every prime factor of a is among the primes $p_1, p_2, \cdots, p_\ell$. It follows that a can be written in the form $a = {p_1}^{j_1}{p_2}^{j_2} \cdots {p_\ell}^{j_\ell}$ with each $j_i \in \mathbf{N}$. It remains to show that $j_i \leq m_i$.

Since $a|n$ we can choose $b \in \mathbf{Z}$ so that $n = ab$. Since n and a are positive, so is b. Since b is a positive factor of n, the above argument shows that every prime factor of b is among the primes $p_1, p_2, \cdots, p_\ell$ and so we can write $b = {p_1}^{k_1}{p_2}^{k_2} \cdots {p_\ell}^{k_\ell}$ for some $k_i \in \mathbf{N}$. Since $n = ab$ we have

$${p_1}^{m_1}{p_2}^{m_2} \cdots {p_\ell}^{m_\ell} = n = ab = ({p_1}^{j_1} \cdots {p_\ell}^{j_\ell})({p_1}^{k_1} \cdots {p_\ell}^{k_\ell}) = {p_1}^{j_1+k_1} \cdots {p_\ell}^{j_\ell+k_\ell}.$$

By the uniqueness of prime factorization, it follows that $m_i = j_i + k_i$ for all indices i. Since $k_i \geq 0$ it follows that $j_i = m_i - k_i \leq m_i$, as required.

5.28 Definition: For $a, b \in \mathbf{Z}$, a **common multiple** of a and b is an integer m such that $a|m$ and $b|m$. When a and b are nonzero, we define $\text{lcm}(a, b)$ to be the smallest positive common multiple of a and b. For convenience, we also define $\text{lcm}(a, 0) = \text{lcm}(0, a) = 0$ for $a \in \mathbf{Z}$.

5.29 Theorem: *Let $a = \prod_{i=1}^{\ell} {p_i}^{j_i}$ and $b = \prod_{i=1}^{\ell} {p_i}^{k_i}$ where $\ell \in \mathbf{Z}^+$, the p_i are distinct primes, and $j_i, k_i \in \mathbf{N}$. Then*

(1) $\gcd(a, b) = \prod_{i=1}^{\ell} {p_i}^{\min\{j_i, k_i\}}$,

(2) $\text{lcm}(a, b) = \prod_{i=1}^{\ell} {p_i}^{\max\{j_i, k_i\}}$, *and*

(3) $\gcd(a, b) \cdot \text{lcm}(a, b) = ab$.

Proof: The proof is left as an exercise.

5.30 Definition: For a prime p and a positive integer n, the **exponent** of p in (the prime factorization of) n, denoted by $e(p,n)$, is defined as follows. We write n in the form $n = {p_1}^{m_1}{p_2}^{m_2}\cdots{p_\ell}^{m_\ell}$ where the p_i are distinct primes and each $m_i \in \mathbf{N}$, then we define $e(p,n) = m_i$ if $p = p_i$ and we define $e(p,n) = 0$ if $p \neq p_i$ for any index i.

5.31 Exercise: Show that $e(p,n!) = \lfloor \frac{n}{p} \rfloor + \lfloor \frac{n}{p^2} \rfloor + \lfloor \frac{n}{p^3} \rfloor + \cdots$ and that $\lfloor \frac{n}{p^{k+1}} \rfloor = \left\lfloor \lfloor \frac{n}{p^k} \rfloor \Big/ p \right\rfloor$.

5.32 Example: Since $e(5,100!) = \lfloor \frac{100}{5} \rfloor + \lfloor \frac{100}{25} \rfloor + \lfloor \frac{100}{125} \rfloor + \cdots = 20 + 4 + 0 = 24$ and $e(2,100!) > 24$, it follows that the number 100! ends with exactly 24 zeros in its decimal representation.

5.33 Definition: For a positive integer n, we write $\tau(n)$ to denote the number of positive divisors of n, we write $\sigma(n)$ to denote the sum of the positive divisors of n, and we write $\rho(n)$ to denote the product of the positive divisors of n.

5.34 Exercise: Let $n = \prod_{i=1}^{\ell} {p_i}^{k_i}$ where $p_1, p_2, \cdots, p_\ell$ are distinct primes and each $k_i \in \mathbf{N}$. Show that $\tau(n) = \prod_{i=1}^{\ell} (k_i + 1)$, $\sigma(n) = \prod_{i=1}^{\ell} \frac{{p_i}^{k_i+1}-1}{p_i-1}$ and $\rho(n) = n^{\tau(n)/2}$.

.

Chapter 6. Congruences and Modular Arithmetic

6.1 Definition: Let $n \in \mathbf{Z}^+$. For $a, b \in \mathbf{Z}$ we say that a is equal (or **congruent**) to b **modulo** n, and we write $a = b \bmod n$, when $n \big| (a - b)$ or, equivalently, when $a = b + kn$ for some $k \in \mathbf{Z}$.

6.2 Theorem: *Let $n \in \mathbf{Z}^+$. For $a, b \in \mathbf{Z}$ we have $a = b \bmod n$ if and only if a and b have the same remainder when divided by n. In particular, for every $a \in \mathbf{Z}$ there is a unique $r \in \mathbf{Z}$ with $a = r \bmod n$ and $0 \leq r < n$.*

Proof: Let $a, b \in \mathbf{Z}$. Use the Division Algorithm to write $a = qn + r$ with $0 \leq r < n$ and $b = pn + s$ with $0 \leq s < n$. We need to show that $a = b \bmod n$ if and only if $r = s$. Suppose that $a = b \bmod n$, say $a = b + kn$ where $k \in \mathbf{Z}$. Then since $a = qn + r$ and $a = b + kn = (pn + s) + kn = (p + k)n + s$ with $0 \leq r < n$ and $0 \leq s < n$, it follows that $q = p + s$ and $r = s$ by the uniqueness part of the Division Algorithm. Conversely, suppose that $r = s$. Then we have $0 = r - s = (a - qn) - (b - pn)$ so that $a = b + (q - p)n$, and hence $a = b \bmod n$.

6.3 Example: Find 117 mod 35.

Solution: We are being asked to find the unique integer r with $0 \leq r < n$ such that $117 = r \bmod 35$ or, in other words, to find the remainder r when 117 is divided by 35. Since $117 = 3 \cdot 35 + 12$ we have $117 = 12 \bmod 35$.

6.4 Definition: An **equivalence relation** on a set S is a binary relation $\sim$ on S such that

E1. $\sim$ is reflexive: for every $a \in S$ we have $a \sim a$,
E2. $\sim$ is symmetric: for all $a, b \in S$, if $a \sim b$ then $b \sim a$, and
E3. $\sim$ is transitive: for all $a, b, c \in S$, if $a \sim b$ and $b \sim c$ then $a \sim c$.

When $\sim$ is an equivalence relation on S and $a \in S$, the **equivalence class** of a in S is the set

$$[a] = \big\{ x \in S \big| x \sim a \big\}.$$

6.5 Theorem: *Let $n \in \mathbf{Z}^+$. Then congruence modulo n is an equivalence relation on $\mathbf{Z}$.*

Proof: Let $a \in \mathbf{Z}$. Since $a = a + 0 \cdot n$ we have $a = a \bmod n$. Thus congruence modulo n satisfies Property E1. Let $a, b \in \mathbf{Z}$ and suppose that $a = b \bmod n$, say $a = b + kn$ with $k \in \mathbf{Z}$. Then $b = a + (-k)n$ so we have $b = a \bmod n$. Thus congruence modulo n satisfies Property E2. Let $a, b, c \in \mathbf{Z}$ and suppose that $a = b \bmod n$ and $b = c \bmod n$. Since $a = b \bmod n$ we can choose $k \in \mathbf{Z}$ so that $a = b + kn$. Since $b = c \bmod n$ we can choose $\ell \in \mathbf{Z}$ so that $b = c + \ell n$. Then $a = b + kn = (c + \ell n) + kn = c + (k + \ell)n$ and so $a = c \bmod n$. Thus congruence modulo n satisfies Property E3.

6.6 Definition: A **partition** of a set S is a set $\mathcal{P}$ of nonempty disjoint subsets of S whose union is S. This means that

P1. for all $A \in \mathcal{P}$ we have $\emptyset \neq A \subseteq S$,
P2. for all $A, B \in \mathcal{P}$, if $A \neq B$ then $A \cap B = \emptyset$, and
P3. for every $a \in S$ we have $a \in A$ for some $A \in \mathcal{P}$.

6.7 Example: $\mathcal{P} = \big\{ \{1, 3, 5\}, \{2\}, \{4, 6\} \big\}$ is a partition of $S = \{1, 2, 3, 4, 5, 6\}$.

6.8 Theorem: *Let $\sim$ be an equivalence relation on a set S. Then $\mathcal{P} = \big\{[a]\big|a \in S\big\}$ is a partition of S.*

Proof: For $a \in S$, it is clear from the definition of $[a]$ that $[a] \subseteq S$, and we have $[a] \neq \emptyset$ because $a \sim a$ so $a \in [a]$. This shows that $\mathcal{P}$ satisfies P1.

Let $a, b \in S$. We claim that $a \sim b$ if and only if $[a] = [b]$. Suppose that $a \sim b$. Let $x \in S$. Suppose that $x \in [a]$. Then $x \sim a$ by the definition of $[a]$. Since $x \sim a$ and $a \sim b$ we have $x \sim b$ since $\sim$ is transitive. Since $x \sim b$ we have $x \in [b]$. This shows that $[a] \subset [b]$. Since $a \sim b$ implies that $b \sim a$ by symmetry, a similar argument shows that $[b] \subseteq [a]$. Thus we have $[a] = [b]$. Conversely, suppose that $[a] = [b]$. Then since $a \sim a$ we have $a \in [a]$. Since $a \in [a]$ and $[a] = [b]$, we have $a \in [b]$. Since $a \in [b]$, we have $a \sim b$. Thus $a \sim b$ if and only if $[a] = [b]$, as claimed.

Let $a, b \in S$. We claim that if $[a] \neq [b]$ then $[a] \cap [b] = \emptyset$. Suppose that $[a] \cap [b] \neq \emptyset$. Choose $c \in [a] \cap [b]$. Since $c \in [a]$ so that $c \sim a$ we have $[c] = [a]$ (by the above claim). Since $c \in [b]$ so that $c \sim b$ we have $[c] = [b]$. Thus $[a] = [c] = [b]$, as required. This completes the proof that $\mathcal{P}$ satisfies P2.

Finally, note that $\mathcal{P}$ satisfies P3 because given $a \in S$ we have $a \in [a] \in P$.

6.9 Definition: Let $\sim$ be an equivalence relation on a set S. The **quotient** of the set S by the relation $\sim$, denoted by $S/\!\sim$, is the partition $\mathcal{P}$ of the above theorem, that is

$$S/\!\sim = \big\{[a]\big|a \in S\big\}.$$

6.10 Remark: In Appendix 1, the above quotient construction is used to define $\mathbf{Z}$ from $\mathbf{N}$ and to define $\mathbf{Q}$ from $\mathbf{Z}$.

6.11 Definition: Let $n \in \mathbf{Z}^+$. Let $\sim$ be the equivalence relation on $\mathbf{Z}$ defined for $a, b \in \mathbf{Z}$ by $a \sim b \Longleftrightarrow a = b \bmod n$, and write $[a] = \{x \in \mathbf{Z} | x \sim a\} = \{x \in \mathbf{Z} | x = a \bmod n\}$. The set of **integers modulo n**, denoted by $\mathbf{Z}_n$, is defined to be the quotient set

$$\mathbf{Z}_n = \mathbf{Z}/\!\sim = \big\{[a]\big|a \in \mathbf{Z}\big\}.$$

Since every $a \in \mathbf{Z}$ is congruent modulo n to a unique $r \in \mathbf{Z}$ with $0 \leq r < n$, we have

$$\mathbf{Z}_n = \big\{[0], [1], [2], \cdots, [n-1]\big\}$$

and the elements listed in the above set are distinct so that $\mathbf{Z}_n$ is an n-element set.

6.12 Example: We have

$$\mathbf{Z}_3 = \big\{[0], [1], [2]\big\} = \big\{\{\cdots, -3, 0, 3, 6, \cdots\}, \{\cdots, -2, 1, 4, 7, \cdots\}, \{\cdots, -1, 2, 5, 8, \cdots\}\big\}.$$

6.13 Theorem: *(Addition and Multiplication Modulo n) Let $n \in \mathbf{Z}^+$. For $a, b, c, d \in \mathbf{Z}$, if $a = c \bmod n$ and $b = d \bmod n$ then $a + b = c + d \bmod n$ and $ab = cd \bmod n$. It follows that we can define addition and multiplication operations on $\mathbf{Z}_n$ by defining*

$$[a] + [b] = [a + b] \quad \text{and} \quad [a]\,[b] = [ab]$$

for all $a, b \in \mathbf{Z}$. When $n \geq 2$, the set $\mathbf{Z}_n$ is a commutative ring using these operations with zero and identity elements $[0]$ and $[1]$.

Proof: Let $a, b, c, d \in \mathbf{Z}$. Suppose that $a = c \bmod n$ and $b = d \bmod n$. Since $a = c \bmod n$ we can choose $k \in \mathbf{Z}$ so that $a = c + kn$. Since $b = d \bmod n$ we can choose $\ell \in \mathbf{Z}$ so that $b = d + \ell n$. Then $a + b = (c + kn) + (d + \ell n) = (c + d) + (k + \ell)n$ so that $a + b = c + d \bmod n$, and $ab = (c + kn)(d + \ell n) = cd + c\ell n + knd + kn\ell n = cd + (kd + \ell c + k\ell n)n$ so that $ab = cd \bmod n$.

It follows that we can define addition and multiplication operations in $\mathbf{Z}_n$ by defining $[a] + [b] = [a + b]$ and $[a]\,[b] = [ab]$ for all $a, b \in \mathbf{Z}$. It is easy to verify that these operations satisfy all of the Axioms R1 - R8 which define a commutative ring. As a sample proof, we shall verify that one half of the distributivity Axiom R7 is satisfied. Let $a, b, c \in \mathbf{Z}$. Then

$$\begin{aligned}
[a]\big([b] + [c]\big) &= [a]\,[b + c]\ \text{, by the definition of addition in } \mathbf{Z}_n \\
&= \big[a(b + c)\big]\ \text{, by the definition of multiplication in } \mathbf{Z}_n, \\
&= [ab + ac]\ \text{, by distributivity in } \mathbf{Z}. \\
&= [ab] + [ac]\ \text{, by the definition of addition in } \mathbf{Z}_n, \\
&= [a]\,[b] + [a]\,[c]\ \text{, by the definition of multiplication in } \mathbf{Z}_n.
\end{aligned}$$

6.14 Note: When no confusion arises, we shall often omit the square brackets from our notation so that for $a \in \mathbf{Z}$ we write $[a] \in \mathbf{Z}_n$ simply as $a \in \mathbf{Z}_n$. Using this notation, for $a, b \in \mathbf{Z}$ we have $a = b$ in $\mathbf{Z}_n$ if and only if $a = b \bmod n$ in $\mathbf{Z}$.

6.15 Example: Addition and multiplication in $\mathbf{Z}_6$ are given by the following tables.

+	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	4
5	5	0	1	2	3	4

×	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	0	4	2
5	0	5	4	3	2	1

6.16 Example: Find $251 \cdot 329 + (41)^2 \bmod 16$.

Solution: Since $251 = 15 \cdot 16 + 11$ and $329 = 20 \cdot 16 + 9$ and $41 = 2 \cdot 16 + 9$, working in $\mathbf{Z}_{16}$ we have $251 = 11$ and $329 = 41 = 9$ so that

$$251 \cdot 329 + (41)^2 = 11 \cdot 9 + 9^2 = (11 + 9) \cdot 9 = 20 \cdot 9 = 4 \cdot 9 = 36 = 4.$$

Thus $251 \cdot 329 + (41)^2 = 4 \bmod 16$.

6.17 Example: Show that for all $a \in \mathbf{Z}$, if $a = 3 \bmod 4$ then a is not equal to the sum of 2 perfect squares.

Solution: In $\mathbf{Z}_4$ we have $0^2 = 0$, $1^2 = 1$, $2^2 = 4 = 0$ and $3^2 = 9 = 1$ so that $x^2 \in \{0, 1\}$ for all $x \in \mathbf{Z}_4$. It follows that for all $x, y \in \mathbf{Z}_4$ we have $x^2 + y^2 \in \{0+0, 0+1, 1+0, 1+1\} = \{0, 1, 2\}$ so that $x^2 + y^2 \neq 3$. Equivalently, for all $x, y \in \mathbf{Z}$ we have $x^2 + y^2 \neq 3 \bmod 4$.

6.18 Example: Show that there do not exist integers x and y such that $3x^2 + 4 = y^3$.

Solution: In $\mathbf{Z}_9$ we have

x	0	1	2	3	4	5	6	7	8
x^2	0	1	4	0	7	7	0	4	1
x^3	0	1	8	0	1	8	0	1	8
$3x^2$	0	3	3	0	3	3	0	3	3
$3x^2+4$	4	7	7	4	7	7	4	7	7

From the table we see that for all $x, y \in \mathbf{Z}_9$ we have $3x^2 + 4 \in \{4, 7\}$ and $y^3 \in \{0, 1, 8\}$ and so $3x^2 + 4 \neq y^3$. It follows that for all $x, y \in \mathbf{Z}$ we have $3x^2 + 4 \neq y^3$.

6.19 Example: There are several well known tests for divisibility which can be easily explained using modular arithmetic. Suppose that a positive integer n is written in decimal form as $n = d_\ell \cdots d_1 d_0$ where each d_i is a decimal digit, that is $d_i \in \{0, 1, \cdots, 9\}$. This means that

$$n = \sum_{k=0}^{\ell} 10^i d_i.$$

Since $2 \big| 10$ we have $10 = 0 \bmod 2$. It follows that in $\mathbf{Z}_2$ we have $10 = 0$ so $n = \sum_{i=0}^{\ell} 10^i d_i = d_0$. Thus in $\mathbf{Z}$, we have $2|n \Longleftrightarrow n = 0 \bmod 2 \Longleftrightarrow d_0 = 0 \bmod 2 \Longleftrightarrow 2|d_0$. In other words,

2 divides n if and only if 2 divides the final digit of n.

More generally for $k \in \mathbf{Z}$ with $1 \leq k \leq \ell$, since $2^k \big| 10^k$ it follows that in $\mathbf{Z}_{2^k}$ we have $10^k = 0$, hence $10^i = 0$ for all $i \geq k$, and so $n = \sum_{i=0}^{\ell} 10^i d_i = \sum_{i=0}^{k-1} 10^i d_i$. Thus in $\mathbf{Z}$, we have $2^k \big| n$ if and only if $2^k \Big| \sum_{i=0}^{k-1} 10^i d_i$. In other words,

2^k divides n if and only if 2^k divides the tailing k-digit number of n.

Similarly, since $5^k \big| 10^k$ it follows that

5^k divides n if and only if 5^k divides the tailing k-digit number of n.

Since $10 = 1 \bmod 3$ it follows that in $\mathbf{Z}_3$ we have $10 = 1$ so that $n = \sum_{i=1}^{\ell} 10^i d_i = \sum_{i=0}^{\ell} d_i$. Thus in $\mathbf{Z}$, $3|n \Longleftrightarrow n = 0 \bmod 3 \Longleftrightarrow \sum_{i=0}^{\ell} d_i = 0 \bmod 3 \Longleftrightarrow 3 \Big| \sum_{i=0}^{\ell}$. In other words, 3 divides n if and only if 3 divides the sum of the digits of n. Similarly, since $10 = 1 \bmod 9$,

9 divides n if and only if 9 divides the sum of the digits of n.

Since $10 = -1 \bmod 11$, in $\mathbf{Z}_{11}$ we have $10 = -1$ so that $n = \sum_{i=0}^{\ell} 10^i d_i = \sum_{i=0}^{\ell} (-1)^i d_i$. Thus in $\mathbf{Z}$, $11 \big| n \Longleftrightarrow 11 \Big| \sum_{i=0}^{\ell} (-1)^i d_i$. In other words,

11 divides n if and only if 11 divides the alternating sum of the digits of n.

6.20 Exercise: Use the divisibility tests described in the above example to find the prime factorization of the number 28880280. Also, consider the problem of factoring the number 28880281.

6.21 Remark: For $a, b \in \mathbf{Z}$ and $n \in \mathbf{Z}^+$ note that if $a = b \bmod n$ so that $[a] = [b] \in \mathbf{Z}_n$ then we have $\gcd(a, n) = \gcd(b, n)$ and so it makes sense to define $\gcd([a], n) = \gcd(a, n)$.

6.22 Theorem: *(Inverses Modulo n) Let $n \in \mathbf{Z}$ with $n \geq 2$. For $a \in \mathbf{Z}$, $[a]$ is a unit in $\mathbf{Z}_n$ if and only if $\gcd(a, n) = 1$ in $\mathbf{Z}$.*

Proof: Let $a \in \mathbf{Z}$ and let $d = gcd(a, n)$. Suppose that $[a]$ is a unit in $\mathbf{Z}_n$. Choose $b \in \mathbf{Z}$ so that $[a]\,[b] = [1] \in \mathbf{Z}_n$. Then $[ab] = [1] \in \mathbf{Z}_n$ and so $ab = 1 \bmod n$ in $\mathbf{Z}$. Since $ab = 1 \bmod n$ we can choose k so that $ab = 1 + kn$. Then we have $ab - kn = 1$. Since $d|a$ and $d|n$ it follows that $d\big|(ax + ny)$ for all $x, y \in \mathbf{Z}$ so in particular $d\big|(ab - kn)$, that is $d|1$. Since $d|1$ and $d \geq 0$, we must have $d = 1$.

Conversely, suppose that $d = 1$. By the Euclidean Algorithm with Back-Substitution, we can choose $s, t \in \mathbf{Z}$ so that $as + nt = 1$. Then we have $as = 1 - nt$ so that $as = 1 \bmod n$. Thus in $\mathbf{Z}_n$, we have $[as] = [1]$ so that $[a][s] = [1]$. Thus $[a]$ is a unit with $[a]^{-1} = [s]$.

6.23 Example: Determine whether 125 is a unit in $\mathbf{Z}_{471}$ and if so find 125^{-1}.

Solution: The Euclidean Algorithm gives

$$471 = 3 \cdot 125 + 96\ ,\ 125 = 1 \cdot 96 + 29\ ,\ 96 = 3 \cdot 29 + 9\ ,\ 29 = 3 \cdot 9 + 2\ ,\ 9 = 4 \cdot 2 + 1$$

and so $d = \gcd(125, 471) = 1$ and it follows that 125 is a unit in $\mathbf{Z}_{471}$. Back-Substitution gives the sequence

$$1\ ,\ -4\ ,\ 13\ ,\ -43\ ,\ 56\ ,\ -211$$

so we have $125(-211) + 471(56) = 1$. It follows that in $\mathbf{Z}_{471}$ we have $125^{-1} = -211 = 260$.

6.24 Example: Solve the pair of equations $3x + 4y = 7$ (1) and $11x + 15y = 8$ (2) for $x, y \in \mathbf{Z}_{20}$.

Solution: We work in $\mathbf{Z}_{20}$. Since $3 \cdot 7 = 21 = 1$ we have $3^{-1} = 7$. Multiply both sides of Equation (1) by 7 to get $x + 8y = 9$, that is $x = 9 - 8y$ (3). Substitute $x = 9 - 8y$ into Equation (2) to get $11(9 - 8y) + 15y = 8$, that is $19 - 8y + 15y = 8$ or equivalently $7y = 9$ (4). Multiply both sides of Equation (4) by $7^{-1} = 3$ to get $y = 7$. Put $y = 7$ into Equation (3) to get $x = 9 - 8 \cdot 7 = 9 - 16 = 13$. Thus the only solution is $(x, y) = (13, 7)$.

6.25 Definition: A **group** is a set G with an element $e \in G$ and a binary operation $* : G \times G \to G$, where for $a, b \in G$ we write $*(a, b)$ as $a * b$ or simply as ab, such that

G1. $*$ is associative: for all $a, b, c \in G$ we have $(ab)c = a(bc)$,
G2. e is an identity element: for all $a \in G$ we have $ae = ea = a$, and
G3. every $a \in G$ has an inverse: for every $a \in G$ there exists $b \in G$ such that $ab = ba = e$.

A group G is called **abelian** when

G4. $*$ is commutative: for all $a, b \in G$ we have $ab = ba$.

6.26 Definition: When R is a ring under the operations $+$ and $\times$, the set R is also a group under the operation $+$ with identity element 0. The group R under $+$ is called the **additive group** of R. The set R is not a group under the operation $\times$ because not every element $a \in R$ has an inverse under $\times$ (in particular, the element 0 has no inverse). The set of all invertible elements in R, however, is a group under multiplication, and we denote it by R^*, so we have

$$R^* = \big\{a \in R \big| a \text{ is a unit}\big\}.$$

The group R^* is called the **group of units** of R.

6.27 Example: When F is a field, every nonzero element in F is invertible so we have $F^* = F \setminus \{0\}$. In $\mathbf{Z}$, the only invertible elements are ± 1 and so $\mathbf{Z}^* = \{1, -1\}$.

6.28 Definition: For $n \in \mathbf{Z}$ with $n \geq 2$, the group of units of $\mathbf{Z}_n$ is called the **group of units modulo** n and is denoted by U_n. Thus

$$U_n = \big\{ a \in \mathbf{Z}_n \big| \gcd(a, n) = 1 \big\}.$$

For convenience, we also let U_1 be the trivial group $U_1 = \mathbf{Z}_1 = \{1\}$. For a set S, let $|S|$ denote the cardinality of S, so that in particular when S is a finite set, $|S|$ denotes the number of elements in S. We define the **Euler phi function**, also called the **Euler totient function**, $\varphi : \mathbf{Z}^+ \to \mathbf{Z}^+$ by

$$\varphi(n) = \big|U_n\big|$$

so that $\varphi(n)$ is equal to the number of elements $a \in \{1, 2, \cdots, n\}$ such that $\gcd(a, n) = 1$.

6.29 Example: Since $U_{20} = \{1, 3, 7, 9, 11, 13, 17, 19\}$ we have $\varphi(20) = 8$.

6.30 Example: When p is a prime number and $k \in \mathbf{Z}^+$ notice that

$$U_{p^k} = \{1, 2, 3, \cdots, p^k\} \setminus \{p, 2p, 3p, \cdots, p^k\}$$

and so

$$\varphi(p^k) = p^k - p^{k-1}.$$

6.31 Theorem: *(Fermat's Little Theorem) Let p be a prime number. Then*
(1) For all $a \in \mathbf{Z}$ with $\gcd(a, p) = 1$ we have $a^{p-1} = 1 \bmod p$.
(2) For all $a \in \mathbf{Z}$ we have $a^p = a \bmod p$.

Proof: To prove Part (1), let $a \in \mathbf{Z}$ with $\gcd(a, p) = 1$. Then we have $a \in U_p$. Define $F : U_p \to U_p$ by $F(x) = ax$ (note that when a and x are units in a ring, the product ax is also a unit with $(ax)^{-1} = x^{-1}a^{-1}$, so the map F is well-defined). Notice that F is bijective with inverse $G : U_p \to U_p$ given by $G(x) = a^{-1}x$. Since F is bijective, it follows that the list of elements $1a, 2a, 3a, \cdots, (p-1)a$ is a permutation (that is a a re-orderring) of the list $1, 2, 3, \cdots, p-1$. Thus in U_p we have

$$1a \cdot 2a \cdot 3a \cdot \ldots \cdot (p-1)a = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (p-1)$$
$$(p-1)!\, a^{p-1} = (p-1)!$$

Multiply both sides by the inverse of $(p-1)!$ in U_p to get $a^{p-1} = 1$ in U_p, as required.

To prove Part (2), let $a \in \mathbf{Z}$ be arbitrary. If $\gcd(a, p) = 1$ then by Part (1) we have $a^{p-1} = 1 \bmod p$ and so we can multiply by a to get $a^p = a \bmod p$. If $\gcd(a, p) \neq 1$ then since p is prime it follows that $p|n$ and so we have $a = 0 \bmod p$ hence $a^p = 0^p = 0 = a \bmod p$. In either case, we have $a^p = a \bmod p$, as required.

6.32 Example: If today is Tuesday, then what day will it be in 2^{100} days?

Solution: By Fermat's Little Theorem we have $2^6 = 1 \bmod 7$. It follows that the list of powers of 2 repeats every 6 terms in $\mathbf{Z}_7$. Since $100 = 16 \cdot 6 + 4$ so that $100 = 4 \bmod 6$, it follows that $2^{100} = 2^4 = 16 = 2 \bmod 7$. Thus in 2^{100} days it will be Thursday.

6.33 Example: Show that $2^{70} + 3^{70}$ is not prime.

Solution: In $\mathbf{Z}_2$ we have $2^{70} + 3^{70} = 0^{70} + 1^{70} = 1 \neq 0$. In $\mathbf{Z}_3$, we have $2^{70} + 3^{70} = (-1)^{70} + 0^{70} = 1 \neq 0$. In $\mathbf{Z}_5$, by Fermat's Little Theorem the list of powers of 2 and 3 repeats every 4 terms, and $70 = 2 \bmod 4$, so we have $2^{70} + 3^{70} = 2^2 + 3^2 = 4 + 9 = 3 \neq 0$. In $\mathbf{Z}_7$, the list of powers of 2 and 3 repeats every 6 terms, and $70 = 4 \bmod 6$, so we have $2^{70} + 3^{70} = 2^4 + 3^4 = 4^2 + 9^2 = 4^2 + 2^2 = 2 + 4 = 6 \neq 0$. In $\mathbf{Z}_{11}$, the list of powers of 2 and 3 repeats every 10 terms, and $70 = 0 \bmod 10$, so we have $2^{70} + 3^{70} = 2^{=} + 3^0 = 1 + 1 = 2 \neq 0$. In $\mathbf{Z}_{13}$, the list of powers of 2 and 3 repeats every 12 terms, and $70 = 10 \bmod 12$, so we have $2^{70} + 3^{70} = 2^{10} + 3^{10} = 2^4 \cdot 2^4 \cdot 2^2 + 3^3 \cdot 3^3 \cdot 3^1 = 3 \cdot 3 \cdot 4 + 1 \cdot 1 \cdot 3 = 10 + 3 = 0$. Since $2^{70} + 3^{70} = 0 \in \mathbf{Z}_{13}$ it follows that $13 \big| (2^{70} + 3^{70})$ in $\mathbf{Z}$, and so $2^{70} + 3^{70}$ is not prime.

6.34 Theorem: *(Euler-Fermat) Let $n \in \mathbf{Z}^+$. For all $a \in \mathbf{Z}$ with $\gcd(a, n) = 1$ we have $a^{\varphi(n)} = 1 \bmod n$.*

Proof: Let $a \in \mathbf{Z}$ with $\gcd(a, n) = 1$. Then we have $a \in U_n$. Let $\varphi = \varphi(n)$ and let $x_1, x_2, \cdots, x_\varphi$ be a list of all the elements in U_n. Define $F : U_n \to U_n$ by $F(x) = ax$. Then F is bijective with inverse $G : U_n \to U_n$ given by $G(x) = a^{-1}x$. Since F is bijective, it follows that the list $ax_1, ax_2, \cdots, ax_\varphi$ is a permutation of the list $x_1, x_2, \cdots, x_\varphi$, and so in U_n we have

$$ax_1 \cdot ax_2 \cdot \ldots \cdot ax_\varphi = x_1 \cdot x_2 \cdot \cdots \cdot x_\varphi$$

$$\big(\prod_{i=1}^{\varphi} x_i \big) a^\varphi = \prod_{i=1}^{\varphi} x_i$$

Multiply both sides by the inverse of $\prod_{i=1}^{\varphi} x_i$ in U_n to get $a^\varphi = 1$ in U_n, as required.

6.35 Remark: For any finite abelian group G, the above proof is valid and it shows that $a^{|G|} = e$ for all $a \in G$. The same result holds even in non-abelian finite groups, but a different proof is required.

6.36 Theorem: *(The Linear Congruence Theorem) Let $n \in \mathbf{Z}^+$, let $a, b \in \mathbf{Z}$, and let $d = \gcd(a, n)$. Consider the congruence $ax = b \bmod n$.*

(1) The congruence has a solution $x \in \mathbf{Z}$ if and only if $d|b$, and
(2) if $x = u$ is one solution to the congruence, then the general solution is

$$x = u \bmod \tfrac{n}{d}.$$

Proof: Suppose that the congruence $ax = b \bmod n$ has a solution. Let $x = u$ be a solution so we have $au = b \bmod n$. Since $au = b \bmod n$ we can choose $k \in \mathbf{Z}$ so that $au = b + kn$, that is $au - nk = b$. Since $d|a$ and $d|n$ it follows that $d|(ax + ny)$ for all $x, y \in \mathbf{Z}$, and so in particular $d \big| (au - nk)$, hence $d|b$. Conversely, suppose that $d|b$. By the Linear Diophanitine Equation Theorem, the equation $ax + ny = b$ has a solution. Choose $u, v \in \mathbf{Z}$ so that $au + nv = b$. Then since $au = b - nv$ we have $au = b \bmod n$ and so the congruence $ax = b \bmod n$ has a solution (namely $x = u$).

Suppose that $x = u$ is a solution to the given congruence, so we have $au = b \bmod n$. We need to show that for every $k \in \mathbf{Z}$ if we let $x = u + k\frac{n}{d}$ then we have $ax = b \bmod n$ and, conversely, that for every $x \in \mathbf{Z}$ such that $ax = b \bmod n$ there exists $k \in \mathbf{Z}$ such that $x = u + k\frac{n}{d}$. Let $k \in \mathbf{Z}$ and let $x = u + k\frac{n}{d}$. Then $ax = a\big(u + k\frac{n}{d}\big) = au + \frac{ka}{d}\, n$. Since $ax = au + \frac{ka}{d}\, n$ and $d|a$ so that $\frac{ka}{d} \in \mathbf{Z}$, it follows that $ax = au \bmod n$. Since $ax = au \bmod n$ and $au = b \bmod n$ we have $ax = b \bmod n$, as required.

Conversely, let $x \in \mathbf{Z}$ and suppose that $ax = b \bmod n$. Since $ax = b \bmod n$ and $au = b \bmod n$ we have $ax = au \bmod n$. Since $ax = au \bmod n$ we can choose $\ell \in \mathbf{Z}$ so that $ax = au + \ell n$. Then we have $a(x-u) = \ell n$ and so $\frac{a}{d}(x-u) = \frac{n}{d}\ell$. Since $\frac{n}{d}\big|\frac{a}{d}(x-u)$ and $\gcd\left(\frac{a}{d}, \frac{n}{d}\right) = 1$, it follows that $\frac{n}{d}\big|(x-u)$. Thus we can choose $k \in \mathbf{Z}$ so that $x - u = k\frac{n}{d}$ and then we have $x = u + k\frac{n}{d}$, as required.

6.37 Example: Solve $221x = 595 \bmod 323$.

Solution: The Euclidean Algorithm gives

$$323 = 1 \cdot 221 + 102 \ , \ 221 = 2 \cdot 102 + 17 \ , \ 102 = 6 \cdot 17 + 0$$

and so $\gcd(221, 323) = 17$. Note that $\frac{595}{17} = 35$, so the congruence has a solution. Back-Substitution gives the sequence

$$1 \ , \ -2 \ , \ 3$$

so we have $221 \cdot 3 - 323 \cdot 2 = 17$. Multiply by 35 to get $221 \cdot 105 - 323 \cdot 70 = 595$. Thus one solution to the given congruence is $x = 105$. Since $\frac{323}{17} = 19$ and $105 = 5 \cdot 19 + 10$, the general solution is given by $x = 105 = 10 \bmod 19$.

6.38 Theorem: *(The Chinese Remainder Theorem) Let $n, m \in \mathbf{Z}^+$ and let $a, b \in \mathbf{Z}$. Consider the pair of congruences*

$$x = a \bmod n,$$
$$x = b \bmod m.$$

(1) The pair of congruences has a solution $x \in \mathbf{Z}$ if and only if $\gcd(n, m)\big|(b - a)$, and
(2) if $x = u$ is one solution, then the general solution is $x = u \bmod \operatorname{lcm}(n, m)$.

Proof: Suppose that the given pair of congruences has a solution and let $d = \gcd(n, m)$. Let $x = u$ be a solution, so we have $u = a \bmod n$ and $u = b \bmod m$. Since $u = a \bmod n$ we can choose $k \in \mathbf{Z}$ so that $u = a + kn$. Since $u = b \bmod m$ we can choose $\ell \in \mathbf{Z}$ so that $u = b + \ell m$. Since $u = a + kn = b + \ell n$ we have $b - a = nk - m\ell$. Since $d|n$ and $d|m$ it follows that $d\big|(nx + my)$ for all $x, y \in \mathbf{Z}$ so in particular $d\big|(nk - m\ell)$, hence $d|(b - a)$. Conversely, suppose that $d|(b - a)$. By the Linear Diophantine Equation Theorem, the equation $nx + my = b - a$ has a solution. Choose $k, \ell \in \mathbf{Z}$ so that $nk - m\ell = b - a$. Then we have $a + nk = b + m\ell$. Let $u = a + nk = b + m\ell$. Since $u = a + nk$ we have $u = a \bmod n$ and since $u = b + m\ell$ we have $u = b \bmod m$. Thus $x = u$ is a solution to the pair of congruence.

Now suppose that $u = a \bmod n$ and $u = b \bmod m$. Let $\ell = \operatorname{lcm}(n, m)$. Let $k \in \mathbf{Z}$ be arbitrary and let $x = u + k\ell$. Since $x - u = k\ell$ we have $\ell\big|(x - u)$. Since $n|\ell$ and $\ell\big|(x - u)$ we have $n\big|(x - u)$ so that $x = u \bmod n$. Since $x = u \bmod n$ and $u = a \bmod n$ we have $x = a \bmod n$. Similarly $x = b \bmod m$.

Conversely, let $x \in \mathbf{Z}$ and suppose that $x = a \bmod n$ and $x = b \bmod m$. Since $x = a \bmod n$ and $u = a \bmod n$ we have $x = u \bmod n$ so that $n\big|(x - u)$. Since $x = b \bmod m$ and $u = b \bmod m$ we have $x = u \bmod m$ so that $m\big|(x - u)$. Since $n\big|(x - u)$ and $m\big|(x - u)$ and $\ell = \operatorname{lcm}(n, m)$, it follows that $\ell\big|(x - u)$ so that $x = u \bmod \ell$.

6.39 Example: Solve the pair of congruences $x = 2 \bmod 15$ and $x = 13 \bmod 28$.

Solution: We want to find $k, \ell \in \mathbf{Z}$ such that $x = 2+15k = 13+28\ell$. We need $15k-28\ell = 11$. The Euclidean Algorithm gives

$$28 = 1 \cdot 15 + 13 \ , \ 15 = 1 \cdot 13 + 2 \ , \ 13 = 6 \cdot 2 + 1$$

so that $\gcd(15, 28) = 1$ and Back-Substitution gives the sequence

$$1 \ , \ -6 \ , \ 7 \ , \ -13$$

so that $(15)(-13) + (28)(7) = 1$. Multiplying by 11 gives $(15)(-143) + (28)(77) = 11$, so one solution to the equation $15k - 28\ell = 11$ is given by $(k, l) = (-143, 77)$. It follows that one solution to the pair of congruences is given by $u = 2 + 15k = 2 - 15 \cdot 143 = -2143$. Since $\operatorname{lcm}(15, 28) = 15 \cdot 28 = 420$, and $-2143 = -6 \cdot 420 + 377$, the general solution to the pair of congruences is $x = -2143 = 377 \bmod 420$.

6.40 Exercise: Solve the congruence $x^3 + 2x = 18 \bmod 35$.

6.41 Exercise: Find the last 2 digits of $14^{14^{14}}$ in its decimal representation.

6.42 Theorem: *(The Generalized Chinese Remainder Theorem) Let $\ell \in \mathbf{Z}^+$, let $n_i \in \mathbf{Z}^+$ and $a_i \in \mathbf{Z}$ for all indices i with $1 \leq i \leq \ell$. Consider the system of ℓ congruences $x = a_i \bmod n_i$ for all indices i with $1 \leq i \leq \ell$.*

(1) The system has a solution x if and only if $\gcd(n_i, n_j) \big| (a_i - a_j)$ for all i, j, and
(2) if $x = u$ is one solution then the general solution is $x = u \bmod \operatorname{lcm}(n_1, n_2, \cdots, n_\ell)$.

Proof: The proof is left as an exercise.

6.43 Exercise: Solve the system $x = 17 \bmod 25$, $x = 14 \bmod 18$ and $x = 22 \bmod 40$.

6.44 Theorem: *Let $n = \prod_{i=1}^{\ell} {p_i}^{k_k}$ where $\ell \in \mathbf{Z}^+$ and the p_i are distinct primes and each $k_i \in \mathbf{Z}^+$. Then*

$$\varphi(n) = \prod_{i=1}^{\ell} \varphi\big({p_i}^{k_i}\big).$$

Proof: I may include a proof later.

6.45 Example: When p and q are distinct primes, we have $\varphi(pq) = (p-1)(q-1)$.

.

Chapter 7. Cryptography

7.1 Definition: Cryptography is the study of secret codes. When we convert a message from a normal language, say English, to a secret code, we say that we **encrypt** (or **encipher**) the message, and the coded word is called the **ciphertext**. When we convert the ciphertext back into normal language, we say that we **decipher** (or **decrypt**) the ciphertext to obtain the original message.

7.2 Example: One of the simplest encryption methods is a **Caesar cipher**. Suppose Alice wants to send a secret message to Bob using a Caesar cipher. Alice and Bob agree in advance on a number n between 1 and 25. Alice encrypts the message by replacing each letter in the message by the letter which follows it by n positions (modulo 26) in the English alphabet. For example, if $n = 4$ then the letter P would be replaced by the letter T (which follows P by 4 positions), and the message PONY would be replaced by the ciphertext TSRB. Bob can easily decrypt the ciphertext by replacing each letter by the letter which precedes it by n positions.

7.3 Example: A slightly more secure encryption method is a **substitution cipher**. Suppose that Alice wants to send a secret message to Bob using a substitution cipher. Alice and Bob agree in advance on a permutation p of the letters of the English alphabet. Alice enciphers the message by replacing each letter by the letter which corresponds to it under the permutation p. For example, if the permutation p is given as follows

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
V G S C F U Q L A P I D X N W T H Y O J K Z B E R M

then the letter H would be replaced by the letter L and the message HORSE would be replaced by the ciphertext LWYOF.

7.4 Definition: A far more secure encryption system, which is commonly used by modern computers, is the **RSA scheme**. The letters R, S and A stand for Rivest, Shamir and Adleman. who first described this encryption system. The RSA scheme is a **public key** encryption system, which means that when a person, say Alice, wishes to receive a secret message, she makes her encryption rules publicly known so that anyone can encipher a message and send it to Alice and yet, although everyone knows the encryption rules, only Alice knows the decryption rules and can decipher the ciphertext.

Suppose that Alice wishes to receive a secret message using the RSA scheme. Alice chooses two large prime numbers p and q (in practice, p and q would have over 100 decimal digits) and calculates $n = pq$ and $\varphi = \varphi(n) = (p-1)(q-1)$. Then Alice chooses a positive integer $e < \varphi$ with $\gcd(e, \varphi) = 1$ and calculates $d = e^{-1} \mod \varphi$. The number e is called the **encryption key** and the number d is called the **decryption key**. Then Alice makes the numbers n and e publicly known. Suppose that Bob wishes to send a message to Alice. Bob converts his message to a positive integer m with $m < n$ (if his message is too long then he breaks it into shorter messages). Bob calculates the ciphertext $c = m^e \mod n$ which he sends to Alice. Note that since $ed = 1 \mod \varphi$, we have $c^d = (m^e)^d = m^{ed} = m^1 = m \mod n$ by the Euler Fermat Theorem, and so Alice can recover the original message m by calculating $m = c^d \mod n$.

7.5 Note: Alice can save some time if, instead of calculating $\varphi = (p-1)(q-1)$ and $d = e^{-1} \bmod \varphi$, she instead calculates $\psi = \text{lcm}(p-1, q-1)$ and $d = e^{-1} \bmod \psi$. Verify that when $c = m^e \bmod n$ we have $c^d = (c^e)^d = c^{ed} = c^1 = m \bmod n$.

7.6 Note: The reason that the RSA scheme is practical and secure is that there do exist efficient (polynomial time) algorithms which can be used to find p, q, n, φ, e and d and to calculate $c = m^e \mod n$ and $m = c^d \mod n$, but there is no known efficient algorithm which can be used to determine m from n, e and c. In particular, there do exist efficient algorithms which can be used to determine whether a given positive integer n is prime, but there is no known efficient algorithm which can determine a prime factor of n in the case that n is composite.

There do, of course, exist inefficient algorithms which can determine a prime factor of n. For example, we can use the Sieve of Eratosthenes to list all primes p with $1 < p \leq \sqrt{n}$ and then test each such prime p to determine whether it is a factor of n. But when the prime factors of n are over a hundred digits long, this algorithm is too slow (if a computer could list 10^{10} prime numbers each second then it would take about 10^{80} years to list all the prime numbers p with $p < 10^{100}$).

7.7 Example: The calculation of $d = e^{-1} \mod \varphi$ can be performed using the Euclidean Algorithm, which is efficient.

7.8 Example: When n, e and m are all large, we can calculate $c = m^e \mod n$ efficiently as follows. Express e in base 2, say $e = \sum_{i=1}^{\ell} 2^{k_i}$ with $0 \leq k_1 < k_2 < k_3 < \cdots$, calculate the residues $m^1, m^2, m^4, m^8, \cdots, m^{2^{k_\ell}} \mod n$, then calculate $c = m^e = \prod_{i=1}^{\ell} m^{2^{k_i}} \mod n$. This algorithm is known as the **Square and Multiply Algorithm**.

7.9 Example: Alice wishes to receive a message. She chooses $p = 13$ and $q = 17$ and calculates $n = pq = 221$. She also chooses $e = 35$ and makes the numbers n and e public. Bob wishes to secretly send Alice the letter T. Bob converts the letter T to the number $m = 20$ (since T is the 20^{th} letter in the English alphabet) and sends the cyphertext $c = m^e \bmod n$. As an exercise, calculate $c = m^e \bmod n$ and calculate $\psi = \text{lcm}(p-1, q-1)$ and $d = e^{-1} \bmod \psi$, then directly calculate $c^d \bmod n$ to verify that $c^d = m \bmod n$.

7.10 Definition: Let us describe a simple test for primality which is called the **Fermat Primality Test**. Suppose that we are given an integer $n > 2$. Choose an integer a with $1 < a < n$. By Fermat's Little Theorem, if n is prime then we must have $\gcd(a, n) = 1$ and $a^{n-1} = 1 \mod n$, so we use the Square and Multiply Algorithm to calculate $a^{n-1} \mod n$. If $a^{n-1} \neq 1 \bmod n$ then we can conclude that n is composite while if $a^{n-1} = 1 \mod n$ then we can conclude that n is probably prime.

7.11 Example: Unfortunately, given $n, a \in \mathbf{Z}^+$ with $1 < a < n$, if $a^{n-1} = 1 \bmod n$ then it does not necessarily follow that n is prime. For example, verify that $2^{340} = 1 \bmod 341$ but $341 = 11 \cdot 31$. As another example, verify that $3^{90} = 1 \bmod 91$ but $91 = 7 \cdot 13$.

7.12 Definition: Let $n, a \in \mathbf{Z}^+$ with n composite and $1 < a < n$. If $a^{n-1} \neq 1 \bmod n$ then we say that a is a **Fermat witness** for the compositeness of n. If $a^{n-1} = 1 \bmod n$ then we say that a is a **Fermat liar** and that n is a **Fermat pseudoprime** to base a.

7.13 Note: We can improve the reliability of the above test simply by repeating it. Given $n \in \mathbf{Z}^+$, we choose a finite set S of integers a with $1 < a < n$. For each $a \in S$ we calculate $a^{n-1} \bmod n$. If we find some $a \in S$ such that $a^{n-1} \neq 1 \bmod n$ then we know that n is composite. If we find that for every $a \in S$ we have $a^{n-1} = 1 \bmod n$ then we can conclude that n is probably prime.

7.14 Example: Unfortunately, if if $a^{n-1} = 1 \bmod n$ for every a with $1 < a < n$ and $\gcd(a, n) = 1$ then it does not necessarily follow that n is prime. For example, show that when $n = 3 \cdot 11 \cdot 17 = 561$ we have $a^{n-1} = 1 \bmod n$ for all $a \in \mathbf{Z}$ with $\gcd(a, n) = 1$.

7.15 Definition: For $n \in \mathbf{Z}^+$ we say that n is a **Carmichael number** when n is composite and $a^{n-1} = 1 \bmod n$ for every $a \in \mathbf{Z}$ with $\gcd(a, n) = 1$.

7.16 Theorem: *(Carmichael Numbers) Let $n \in \mathbf{Z}^+$.*
(1) If $n = p_1 p_2 \cdots p_l$ where $\ell \geq 2$ and the p_i are distinct primes which satisfy $(p_i - 1)\big|(n-1)$ for all indices i, then n is a Carmichael number .
(2) If $n = p_1 p_2 \cdots p_l$ where $\ell \geq 2$ and the p_i are distinct primes which satisfy $(p_i - 1)\big|(n-1)$ for all indices i (so that n is a Carmichael number, by Part (1)) then n is odd and $\ell \geq 3$.

Proof: Suppose that $n = p_1 p_2 \cdots p_l$ where the p_i are distinct primes with $(p_i - 1)\big|(n-1)$. Let $a \in \mathbf{Z}^+$ with $\gcd(a, n) = 1$. Fix an index i. Since $\gcd(a, n) = 1$ we have $p_i \not| \, a$ and so $a^{p_i - 1} = 1 \mod p_i$ by Fermat's Little Theorem. Since $a^{p_i - 1} = 1 \mod p_i$ and $(p_i - 1)\big|(n-1)$, we also have $a^{n-1} = 1 \mod p_i$. Since $a^{n-1} = 1 \mod p_i$ for every index i, it follows from the Chinese Remainder Theorem that $a^{n-1} = 1 \mod n$. Thus n is a Carmichael number, so we have proven Part (1).

Let us prove Part (2). Since $l \geq 2$, at least one of the primes p_i is odd, say p_k is odd. Since $p_k - 1$ is even and $(p_k - 1)\big|(n-1)$, it follows that $(n-1)$ is even and so n is odd.

Suppose, for a contradiction, that n is a Carmichael number of the form $n = pq$ where p and q are primes with $p < q$ and we have $(p-1)\big|(n-1)$ and $(q-1)\big|(n-1)$. Note that $n - 1 = pq - 1 = p(q-1) + (p-1)$. Since $(q-1)\big|(n-1)$ we have $(q-1)\big|(n-1) - p(q-1)$, that is $(p-1)\big|(p-1)$. But this implies that $q \leq p$ giving the desired contradiction.

7.17 Exercise: Find distinct primes p and q such that $145\,p$ and $145\,q$ are both Carmichael numbers.

7.18 Theorem: *(The Miller-Rabin Test Theorem) Let n be an odd prime number and let $a \in \mathbf{Z}$ with $\gcd(a, n) = 1$. Write $n - 1 = 2^s d$ where $s, d \in \mathbf{Z}^+$ with d odd. Then*

$$\text{either } \; a^d = 1 \bmod n \;\; \text{ or } \;\; a^{2^r d} = -1 \text{ for some } 0 \leq r < n.$$

Proof: First we remark that since n is prime, $\mathbf{Z}_n$ is a field, so for all $x \in \mathbf{Z}_n$ we have

$$x^2 = 1 \Longleftrightarrow x^2 - 1 = 0 \Longleftrightarrow (x-1)(x+1) = 0 \Longleftrightarrow x = \pm 1\,.$$

By Fermat's Little Theorem, we have $a^{n-1} = 1 \bmod n$, that is $a^{2^s d} = 1 \bmod n$. By the above remark $\big($using $x = a^{2^{s-1} d}\big)$ it follows that $a^{2^{s-1} d} = \pm 1 \bmod n$. If $a^{2^{s-1} d} \neq -1$ then $a^{2^{s-1} d} = 1$ so, by the above remark again, it follows that $a^{2^{s-2}} d = \pm 1$. Similarly, if $a^{2^{s-1} d} \neq -1$ and $a^{2^{s-2} d} \neq -1$ then $a^{2^{s-2} d} = 1$ and hence $a^{2^{s-3}} d = \pm 1$ and so on. Repeating the above argument we find that if $a^{2^{s-1} d} \neq -1\,,\, a^{2^{s-2} d} \neq -1\,,\, \cdots\,,\, a^{2^2 d} \neq -1$ and $a^{2d} \neq -1$ then $a^{2d} = 1$ and hence $a^d = \pm 1$.

7.19 Definition: Using the above theorem we obtain the following test for primality, called the **Miller-Rabin Primality Test**. Given an odd integer $n \in \mathbf{Z}^+$ write $n-1 = 2^s d$ and choose an integer a with $1 < a < n$. By the above theorem, if $a^d \neq 1 \bmod n$ and $a^{2^r d} \neq -1 \bmod n$ for all $0 \leq r < n$ then we can conclude that n is composite. If, on the other hand, we find that either $a^d = 1 \bmod n$ or $a^{2^r d} = -1 \bmod n$ for some $0 \leq r < s$ then we can conclude that n is probably prime.

7.20 Example: Unfortunately, given $n = 1 + 2^s d$ where $s, d \in \mathbf{Z}^+$ with d odd, and given $a \in \mathbf{Z}$ with $1 < a < n$, even if it is true that either $a^d = 1 \bmod n$ or $a^{2^r d} = -1$ for some $0 \leq r < s$, it does not necessarily follow that n is prime. For example, verify that when $n = 221 = 13 \cdot 17$ and $a = 174$ we have $s = 2$ and $d = 55$ and $a^{2d} = -1 \bmod n$.

7.21 Definition: Let $n, a \in \mathbf{Z}^+$ where n is an odd composite number and $1 < a < n$. Write $n - 1 = 2^s d$ where $s, d \in \mathbf{Z}^+$ with d odd. If $a^d \neq 1$ and $a^{2^r d} \neq -1$ for all $0 \leq r < s$ then we say that a is a **Miller-Rabin witness** (or a **strong witness**) for the compositeness of n. If either $a^d = 1$ or $a^{2^r d} = -1$ for some $0 \leq r < n$ then we say that a is a **Rabin-Miller liar** (or a **strong liar**) and that n is a **Rabin-Miller pseudoprime** (or a **strong pseudoprime**).

7.22 Note: As with the Fermat primality test, we can make the Miller-Rabin test more reliable simply by repeating it. Given an odd positive integer n, write $n - 1 = 2^s d$ with $s, d \in \mathbf{Z}^+$ and d odd. Choose a finite set S of integers a with $1 < a < n$. For each $a \in S$, calculate $a^{2^r d} \bmod n$ for $0 \leq r < s$. If we find some $a \in S$ for which $a^d \neq 1 \bmod n$ and $a^{2^r d} \neq -1$ for all $0 \leq r < s$ then we know that n is composite. If, on the other hand, we find that for every $a \in S$, either $a^d = 1 \bmod n$ or $a^{2^r d} = -1 \bmod n$ for some $0 \leq r < s$ then we can conclude that n is probably prime.

7.23 Note: Recall that repeating the Fermat primality test does not make the test become completely reliable because of the existence of Carmichael numbers. The situation is different with the Miller-Rabin primality test. It has been proven that for every composite positive integer n, at least $\frac{3}{4}$ of the numbers a with $1 < a < n$ are strong witnesses for the compositeness of n. It follows that, given an odd composite number n, if we choose m integers a with $1 < a < n$, the probability that none of the numbers a is a strong witness is at most $\frac{1}{4^m}$.

.

Chapter 8. Complex Numbers

8.1 Definition: A **complex number** is a vector in $\mathbf{R}^2$. The **complex plane**, denoted by $\mathbf{C}$, is the set of complex numbers:

$$\mathbf{C} = \mathbf{R}^2 = \big\{(x,y)\big|\, x \in \mathbf{R}, y \in \mathbf{R}\big\}\,.$$

In $\mathbf{C}$ we write $0 = (0,0)$, $1 = (1,0)$, $i = (0,1)$, and for $x, y \in \mathbf{R}$ we write $x = (x,0)$, $iy = yi = (o,y)$ and

$$x + iy = x + yi = (x,y).$$

If $z = x+iy$ with $x, y \in \mathbf{R}$ then x is called the **real** part of z and y is called the **imaginary** part of z, and we write

$$\operatorname{Re} z = x \text{ , and } \operatorname{Im} z = y\,.$$

8.2 Definition: We define the **sum** of two complex numbers to be the usual vector sum:

$$(a+ib) + (c+id) = (a+c) + i(b+d)\,,$$

where $a, b \in \mathbf{R}$. We define the **product** of two complex numbers by setting $i^2 = -1$ and by requiring the product to be commutative and associative and distributive over the sum:

$$(a+ib)(c+id) = ac + iad + ibc + i^2bd = (ac - bd) + i(ad + bc)\,.$$

8.3 Example: Let $z = 2 + i$ and $w = 1 + 3i$. Find $z + w$ and zw.

Solution: $z+w = (2+i)+(1+3i) = (2+1)+i(1+3) = 3+4i$, and $zw = (2+i)(1+3i) = 2 + 6i + i - 3 = -1 + 7i$.

8.4 Theorem: *The set of complex numbers is a field.*

Proof: We shall only verify that each non-zero complex number has an inverse. Let $z = a + ib$ where $a, b \in \mathbf{R}$. Suppose that $z \neq 0$ so $a^2 + b^2 \neq 0$. For $x, y \in \mathbf{R}$ we have

$$\begin{aligned}(a+ib)(x+iy) = 1 &\Longleftrightarrow (ax - by) + (ay + bx)\,i = 1 + 0\,i \\ &\Longleftrightarrow \big(ax - by = 1 \text{ and } bx + ay = 0\big).\end{aligned}$$

We solve the pair of equations $ax - by = 1$ (1) and $bx + ay = 0$ (2). Multiply equation (1) by a and add b times Equation (2) to get $(a^2 + b^2)x = a$, so we need $x = \frac{a}{a^2+b^2}$. Multiply Equation (2) by a and subtract b times Equation (1) to get $(a^2 + b^2)y = -b$ so we need $y = \frac{-b}{a^2+b^2}$. Verify that when $x = \frac{a}{a^2+b^2}$ and $y = \frac{-b}{a^2+b^2}$ we do indeed have $(a+ib)(x+iy) = 1$. This shows that $(a+ib)^{-1}$ does exist and is given by

$$(a+ib)^{-1} = \tfrac{a}{a^2+b^2} - i\tfrac{b}{a^2+b^2}\,.$$

8.5 Example: Find $\dfrac{(4-i)-(1-2i)}{1+2i}$.

Solution: $\dfrac{(4-i)-(1-2i)}{1+2i} = \dfrac{3+i}{1+2i} = (3+i)(1+2i)^{-1} = (3+i)(\frac{1}{5} - \frac{2}{5}i) = 1 - i.$

8.6 Definition: If $z = x + iy$ with $x, y \in \mathbf{R}$ then we define the **conjugate** of z to be

$$\overline{z} = x - iy\,.$$

and we define the **length** (or **magnitude**) of z to be

$$|z| = \sqrt{x^2 + y^2}\,.$$

8.7 Note: For z and w in $\mathbf{C}$ the following identities are all easy to verify.

$$\begin{aligned}
&\overline{\overline{z}} = z \\
&z + \overline{z} = 2\,\mathrm{Re}\,z\ , \quad z - \overline{z} = 2i\,\mathrm{Im}\,z \\
&z\overline{z} = |z|^2\ , \quad |\overline{z}| = |z| \\
&\overline{z+w} = \overline{z} + \overline{w}\ , \quad \overline{zw} = \overline{z}\,\overline{w}\ , \quad |zw| = |z||w|
\end{aligned}$$

8.8 Note: We do *not* have inequalities between complex numbers. We can *only* write $a < b$ or $a \leq b$ in the case that a and b are both *real* numbers. But there are several inequalities between real numbers which concern complex numbers. For $z \in \mathbf{C}$ and $w \in C$,

$$\begin{aligned}
&|\mathrm{Re}\,(z)| \leq |z|\ , \quad |\mathrm{Im}\,(z)| \leq |z| \\
&|z + w| \leq |z| + |w|\ , \quad \text{this is called the } \textbf{triangle inequality} \\
&|z + w| \geq \big||z| - |w|\big|
\end{aligned}$$

The first two inequalities follow from the fact that $|z|^2 = |\mathrm{Re}\,(z)|^2 + |\mathrm{Im}\,(z)|^2$. We can then prove the triangle inequality as follows: $|z+w|^2 = (z+w)(\overline{z}+\overline{w}) = |z|^2+|w|^2+(w\overline{z}+z\overline{w}) = |z|^2 + |w|^2 + 2\mathrm{Re}\,(z\overline{w}) \leq |z|^2 + |w|^2 + 2|z\overline{w}| = |z|^2 + |w|^2 + 2|z||w| = (|z| + |w|)^2$. The last inequality follows from the triangle inequality since $|z| = |z + w - w| \leq |z + w| + |w|$ and $|w| = |z + w - z| \leq |z + w| + |z|$. (Alternatively, the last two inequalities can be proven using the Law of Cosines).

8.9 Example: Given complex numbers a and b, describe the set $\big\{z \in \mathbf{C}\big||z-a| < |z-b|\big\}$.

Solution: Geometrically, this is the set of all z such that z is closer to a than to b, so it is the **half-plane** which contains a and lies on one side of the perpendicular bisector of the line segment ab.

8.10 Example: Given a complex number a, describe the set $\big\{z \in \mathbf{C}\big|1 < |z - a| < 2\big\}$.

Solution: $\big\{z\big||z - a| = 1\big\}$ is the circle centred at a of radius 1 and $\big\{z\big||z - a| = 2\big\}$ is the circle centred at a of radius 2, and $\big\{z \in \mathbf{C}\big|1 < |z - a| < 2\big\}$ is the region between these two circles. Such a region is called an **annulus**.

8.11 Example: Show that every non-zero complex number has exactly two complex square roots, and find a formula for the two square roots of $z = x + iy$.

Solution: Let $z = x + iy$ where $x, y \in \mathbf{R}$ with x and y not both zero. We need to solve $w^2 = z$ for $w \in \mathbf{C}$. Write $w = u + iv$ with $u, v \in \mathbf{R}$. We have

$$\begin{aligned} w^2 = z &\iff (u+iv)^2 = x + iy \iff (u^2 - v^2) + i(2uv) = x + iy \\ &\iff \left(u^2 - v^2 = x \text{ and } 2uv = y\right). \end{aligned}$$

To solve this pair of equations for u, square both sides of the second equation to get $4u^2v^2 = y^2$, then multiply the first equation by $4u^2$ to get $4u^4 - 4u^2v^2 = 4x\,u^2$, that is $4u^4 - 4x\,u^2 - y^2 = 0$. By the quadratic formula,

$$u^2 = \frac{4x \pm \sqrt{16x^2 + 16y^2}}{8} = \frac{x \pm \sqrt{x^2 + y^2}}{2}\,.$$

In the case that $y \neq 0$, we must use the + sign so that the right side is non-negative, so we obtain

$$u = \pm\sqrt{\frac{x + \sqrt{x^2 + y^2}}{2}}\,.$$

A similar calculation gives

$$v = \pm\sqrt{\frac{-x + \sqrt{x^2 + y^2}}{2}}\,.$$

All four choices of sign will satisfy the equation $u^2 - v^2 = x$, but to satisfy $2uv = y$ notice that when $y > 0$, u and v have the same sign, and when $y < 0$, u and v have the opposite sign. It remains only to consider the case that $y = 0$, and we leave this case as an exercise. The final result is that

$$w = \begin{cases} \pm\left(\sqrt{\dfrac{x + \sqrt{x^2+y^2}}{2}} + i\sqrt{\dfrac{-x + \sqrt{x^2+y^2}}{2}}\right) & \text{, if } y > 0, \\ \pm\left(\sqrt{\dfrac{x + \sqrt{x^2+y^2}}{2}} - i\sqrt{\dfrac{-x + \sqrt{x^2+y^2}}{2}}\right) & \text{, if } y < 0, \\ \pm\sqrt{x} \quad \text{, if } y = 0 \text{ and } x > 0, \\ \pm i\sqrt{|x|}\ \text{, if } y = 0 \text{ and } x < 0\,. \end{cases}$$

8.12 Note: When working with real numbers, for $0 < x \in \mathbf{R}$ it is customary to write $\sqrt{x}$ or $x^{1/2}$ to denote the unique positive square root of x. When working with complex numbers, for $0 \neq z \in \mathbf{C}$ we sometimes write $\sqrt{z}$ or $z^{1/2}$ to denote one of the two square roots of z, and we sometimes write $\sqrt{z}$ or $z^{1/2}$ to denote both square roots of z.

8.13 Example: Find $\sqrt{3 - 4i}$.

Solution: Using the formula derived in the previous example, we have

$$\sqrt{3-4i} = \pm\left(\sqrt{\tfrac{3+\sqrt{3^2+4^2}}{2}} - i\sqrt{\tfrac{-3+\sqrt{3^2+4^2}}{2}}\right) = \pm\left(\sqrt{\tfrac{3+5}{2}} - i\sqrt{\tfrac{-3+5}{2}}\right) = \pm(2 - i)\,.$$

8.14 Note: The Quadratic Formula can be used for complex numbers. Indeed for $a, b, c, z \in \mathbf{C}$ with $a \neq 0$ we have

$$az^2 + bz + c = 0 \Longleftrightarrow z^2 + \frac{b}{a}\, z + \frac{c}{a} = 0 \Longleftrightarrow z^2 + \frac{b}{2a}\, z + \left(\frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2 + \frac{c}{a} = 0$$
$$\Longleftrightarrow \left(z + \frac{b}{2a}\right)^2 = \left(\frac{b}{2a}\right)^2 - \frac{c}{a} = \frac{b^2 - 4ac}{4a^2} \Longleftrightarrow z + \frac{b}{2a} = \frac{\sqrt{b^2 - 4ac}}{2a}$$
$$\Longleftrightarrow z = \frac{-b + \sqrt{b^2 - 4ac}}{2a}\,,$$

where $\sqrt{b^2 - 4ac}$ is being used to denote both square roots in the case that $b^2 - 4ac \neq 0$.

8.15 Example: Solve $i\, z^2 - (2 + 3i)z + 5(1 + i) = 0$.

Solution: By the Quadratic Formula, we have

$$z = \frac{(2 + 3i) + \sqrt{(2 + 3i)^2 - 20i(1 + i)}}{2i} = \frac{(2 + 3i) + \sqrt{-5 + 12i + 20 - 20i}}{2i}$$
$$= \frac{(2 + 3i) + \sqrt{15 - 8i}}{2i}$$

and by the formula for square roots we have

$$\sqrt{15 - 8i} = \pm\left(\sqrt{\tfrac{15+\sqrt{15^2+8^2}}{2}} - i\sqrt{\tfrac{-15+\sqrt{15^2+8^2}}{2}}\right) = \pm\left(\sqrt{\tfrac{15+17}{2}} - i\sqrt{\tfrac{-15+17}{2}}\right) = \pm\,(4 - i)$$

and so

$$z = \frac{(2 + 3i) \pm (4 - i)}{2i} = \frac{6 + 2i}{2i} \text{ or } \frac{-2 + 4i}{2i} = 1 - 3i \text{ or } 2 + i\,.$$

8.16 Definition: If $z \neq 0$, we define the **angle** (or **argument**) of z to be the angle $\theta(z)$ from the positive x-axis counterclockwise to z. In other words, $\theta(z)$ is the angle such that

$$z = |z|\big(\cos\theta(z) + i \sin\theta(z)\big)\,.$$

8.17 Note: We can think of the angle $\theta(z)$ in several different ways. We can require, for example, that $0 \leq \theta(z) < 2\pi$ so that the angle is uniquely determined. Or we can allow $\theta(z)$ to be any real number, in which case the angle will be unique up to a multiple of 2π. Then again, we can think of $\theta(z)$ as the infinite set of real numbers $\theta(z) = \{\theta_0 + 2\pi k | k \in z\}$, that is we can regard $\theta(z)$ as an element of $\mathbf{R}/2\pi$, the set of real numbers modulo 2π.

8.18 Notation: For $\theta \in \mathbf{R}$ (or for $\theta \in \mathbf{R}/2\pi$) we shall write

$$e^{i\theta} = \cos\theta + i \sin\theta\,.$$

8.19 Note: If $z \neq 0$ and we have $x = \mathrm{Re}\,(z)$, $y = \mathrm{Im}\,(z)$, $r = |z|$ and $\theta = \theta(z)$ then

$$x = r\cos\theta\ , \quad y = r\sin\theta$$
$$r = \sqrt{x^2 + y^2}\ , \quad \tan\theta = \frac{y}{x},\ \text{if } x \neq 0$$
$$z = re^{i\theta}\ , \quad \overline{z} = r\, e^{-i\,\theta}\ , \quad z^{-1} = \frac{1}{r}e^{-i\,\theta}$$

We say that $x + i\,y$ is the **cartesian** form of z and $re^{i\,\theta}$ is the **polar** form.

8.20 Example: Let $z = -3 - 4i$. Express z in polar form.

Solution: We have $|z| = 5$ and $\tan\theta(z) = \frac{4}{3}$. Since $\theta(z)$ is in the third quadrant, we have $\theta(z) = \pi + \tan^{-1}\frac{4}{3}$. So $z = 5e^{i(\pi + \tan^{-1}(4/3))}$.

8.21 Example: Let $z = 10e^{i\,\tan^{-1}3}$. Express z in cartesian form.

Solution: $z = 10\left(\cos(\tan^{-1}3) + i\,\sin(\tan^{-1}3)\right) = 10\left(\frac{1}{\sqrt{10}} + i\frac{3}{\sqrt{10}}\right) = \sqrt{10} + 3\sqrt{10}\,i.$

8.22 Example: Find a formula for multiplication in polar coordinates.

Solution: For $z = re^{i\alpha}$ and $w = e^{i\beta}$ we have $zw = rs(\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta) = \big((\cos\alpha\cos\beta - \sin\alpha\sin\beta) + i(\sin\alpha\cos\beta + \cos\alpha\sin\beta)\big) = rs\big(\cos(\alpha+\beta) + i\sin(\alpha+\beta)\big)$ and so we obtain the formula

$$re^{i\alpha}se^{i\beta} = rs\,e^{i(\alpha+\beta)}\,.$$

8.23 Note: An immediate consequence of the above example is that

$$(r\,e^{i\,\theta})^n = r^n e^{i\,n\theta}$$

for $r, \theta \in \mathbf{R}$ and for $n \in \mathbf{Z}$. This result is known as **De Moivre's Law**.

8.24 Example: Find $(1+i)^{10}$.

Solution: This can be done in cartesian coordinates using the binomial theorem (which holds for complex numbers), but it is easier in polar coordinates. We have $1+i = \sqrt{2}e^{i\,\pi/4}$ so $(1+i)^{10} = (\sqrt{2}e^{i\,\pi/4})^{10} = (\sqrt{2})^{10}e^{i\,10\pi/4} = 32e^{i\,\pi/2} = 32i.$

8.25 Example: Find a formula for the n^{th} roots of a complex number. In other words, given $z = re^{i\theta}$, solve $w^n = z$.

Solution: Let $w = se^{i\alpha}$. We have $w^n = z \Longleftrightarrow (se^{i\alpha})^n = re^{i\theta} \Longleftrightarrow s^n e^{i\,n\alpha} = re^{i\theta} \Longleftrightarrow s^n = r$ and $n\alpha = \theta + 2\pi k$ for some $k \in \mathbf{Z} \Longleftrightarrow s = \sqrt[n]{r}$ and $\alpha = \dfrac{\theta + 2\pi k}{n}$ for some $k \in \mathbf{Z}$. Notice that when $z \neq 0$ there are exactly n solutions obtained by taking $0 \leq k < n$. So we obtain the formula

$$(r\,e^{i\,\theta})^{1/n} = \sqrt[n]{r}\,e^{i\,(\theta+2\pi k)/n} \quad , \qquad k \in \{0, 1, \ldots, n-1\}\,.$$

In particular, $(r\,e^{i\,\theta})^{1/2} = \pm\sqrt{r}\,e^{i\,\theta/2}$. For $0 < a \in \mathbf{R}$ we have $z^2 = a \Longleftrightarrow z = \pm\sqrt{a}$, and for $0 > a \in \mathbf{R}$ we have $z^2 = a \Longleftrightarrow z = \pm\sqrt{|a|}\,i$.

8.26 Note: When working with complex numbers, for $0 \neq z \in \mathbf{C}$ and for $0 < n \in \mathbf{Z}$, we sometimes write $\sqrt[n]{z}$ or $w^{1/n}$ to denote one of the n solutions to $w^n = z$, and we sometimes write $\sqrt[n]{z}$ or $z^{1/n}$ to denote the set of all n^{th} roots.

8.27 Note: For $z, w \in \mathbf{C}$, the rule

$$(zw)^{1/n} = z^{1/n}w^{1/n}$$

does hold provided that $z^{1/n}$ is used to denote the set of all n^{th} roots, but it does not always hold when $z^{1/n}$ is used to denote one of the n^{th} roots. Consider the following amusing "proof" that $1 = -1$:

$$1 = \sqrt{1} = \sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1} = i^2 = -1\,.$$

8.28 Example: Find $\sqrt[3]{-2+2i}$.

Solution: Note that $-2 + 2i = 2\sqrt{2}\,e^{i\,3\pi/4}$, and so the formula for n^{th} roots gives

$$\begin{aligned}\sqrt[3]{-2+2i} &= \sqrt[3]{2\sqrt{2}\,e^{3\pi/4}} \\ &= \sqrt{2}\,e^{i\left(\pi/4 + \frac{2\pi}{3}\,k\right)}, k \in \{0,1,2\} \\ &= \sqrt{2}\,e^{i\,\pi/3}, \sqrt{2}\,e^{i\,11\pi/12}, \sqrt{2}\,e^{i\,19\pi/12}\,.\end{aligned}$$

8.29 Note: The remaining examples in this chapter illustrate situations in which we can use complex numbers as a tool to help solve certain problems which only involve real numbers.

8.30 Example: Let $x_0 = 1$ and $x_1 = 1$, and for $n \geq 2$ let $x_n = 2x_{n-1} - 5x_{n-2}$. Find a closed-form formula for x_n.

Solution: The characteristic polynomial for the recursion is $z^2 - 2z + 5 = 0$ which has (complex) roots $z = \frac{2 \pm \sqrt{4-20}}{2} = 1 \pm 2i$. By the Linear Recursion Theorem (Theorem 2.47)

$$x_n = A(1+2i)^n + B(1-2i)^n$$

for some constants A and B. To get $x_0 = 1$ and $x_1 = 1$, we need $A + B = 1$ and $A(1+2i) + B(1-2i) = 1$. Solving these two equations gives $A = B = \frac{1}{2}$, so we have

$$x_n = \tfrac{1}{2}\left((1+2i)^n + (1-2i)^n\right) = \tfrac{1}{2}\left(\left(\sqrt{5}\, e^{i\,\theta}\right)^n + \left(\sqrt{5}\, e^{-i\,\theta}\right)^n\right) = \tfrac{(\sqrt{5})^n}{2}\left(e^{i\,n\theta} + e^{-i\,n\theta}\right)$$
$$= \tfrac{\left(\sqrt{5}\right)^n}{2}\left(2\cos n\theta\right) = \left(\sqrt{5}\right)^n \cos n\theta$$

where $\theta = \theta(1+2i) = \tan^{-1} 2$. Thus we obtain

$$x_n = (\sqrt{5})^n \cos\left(n \tan^{-1} 2\right) .$$

8.31 Example: Find $\displaystyle\sum_{i=0}^{n} \binom{3n}{3i}$.

Solution: Let $\alpha = e^{i\,2\pi/3}$. Note that $1 + \alpha + \alpha^2 = 1 + \left(-\frac{1}{2} + \frac{\sqrt{3}}{2}\, i\right) + \left(-\frac{1}{2} - \frac{\sqrt{3}}{2} i\right) = 0$. By the Binomial Theorem we have

$$(1+1)^{3n} = \binom{3n}{0} + \binom{3n}{1} + \binom{3n}{2} + \binom{3n}{3} + \binom{3n}{4} + \cdots + \binom{3n}{3n}$$
$$(1+\alpha)^{3n} = \binom{3n}{0} + \binom{3n}{1}\alpha + \binom{3n}{2}\alpha^2 + \binom{3n}{3} + \binom{3n}{4}\alpha + \cdots + \binom{3n}{3n}$$
$$(1+\alpha^2)^{3n} = \binom{3n}{0} + \binom{3n}{1}\alpha^2 + \binom{3n}{2}\alpha + \binom{3n}{3} + \binom{3n}{4}\alpha^2 + \cdots + \binom{3n}{3n}$$

Adding these three equations gives $(1+1)^{3n} + (1+\alpha)^{3n} + (1+\alpha^2)^{3n} = 3\displaystyle\sum_{i=0}^{n} \binom{3n}{3i}$. Note that $1 + \alpha = 1 - \frac{1}{2} + \frac{\sqrt{3}}{2} i = \frac{1}{2} + \frac{\sqrt{3}}{2} i = e^{i\,\pi/3}$ and similarly $1 + \alpha^2 = e^{-i\,\pi/3}$, and so

$$\sum_{i=0}^{n} \binom{3n}{3i} = \tfrac{1}{3}\left((1+1)^{3n} + (1+\alpha)^{3n} + (1+\alpha^2)^{3n}\right) = \tfrac{1}{3}\left(2^{3n} + \left(e^{i\,\pi/3}\right)^{3n} + \left(e^{-i\,\pi/3}\right)^{3n}\right)$$
$$= \tfrac{1}{3}\left(2^{3n} + e^{i\,n\pi} + e^{-i\,n\pi}\right) = \frac{2^{3n} + 2(-1)^n}{3} .$$

8.32 Note: The Fundamental Theorem of Algebra states that every non-constant polynomial over $\mathbf{C}$ has a root in $\mathbf{C}$. It follows that every such polynomial factors into linear factors over $\mathbf{C}$. If a polynomial $f(x)$ has real coefficients, and α is a complex root of f so that $f(\alpha) = 0$, then we have $f(\overline{\alpha}) = \overline{f(\alpha)} = 0$ so that $\overline{\alpha}$ is also a root of f. Notice that in this case

$$(x - \alpha)(x - \overline{\alpha}) = x^2 - (\alpha + \overline{\alpha})x + \alpha\overline{\alpha} = x^2 - 2\mathrm{Re}\,(\alpha) + |\alpha|^2 ,$$

which has real coefficients. It follows that every non-constant polynomial over $\mathbf{R}$ factors into linear and quadratic factors over $\mathbf{R}$.

8.33 Example: Let $f(x) = x^4 + 2x^2 + 4$. Solve $f(z) = 0$ for $z \in \mathbf{C}$, factor $f(z)$ over the complex number, and then factor $f(x)$ over the real numbers.

Solution: By the quadratic formula, $f(z) = 0$ when $z^2 = -1 \pm \sqrt{3}\,i$ or in polar coordinates $z = 2e^{\pm i\,2\pi/3}$. Thus the roots of f are $z = \pm\sqrt{2}e^{\pm i\,\pi/3}$, and so f factors over $\mathbf{C}$ as

$$z^4 + 2z^2 + 4 = \left(z - \sqrt{2}e^{i\,\pi/3}\right)\left(z - \sqrt{2}e^{-i\,\pi/3}\right)\left(z + \sqrt{2}e^{i\,\pi/3}\right)\left(z + \sqrt{2}e^{-i\,\pi/3}\right).$$

Since $\left(z - \sqrt{2}e^{i\,\pi/3}\right)\left(z - \sqrt{2}e^{-i\,\pi/3}\right) = z^2 - \sqrt{2}\,z + 2$ and $\left(z + \sqrt{2}e^{i\,\pi/3}\right)\left(z + \sqrt{2}e^{-i\,\pi/3}\right) = z^2 + \sqrt{2}\,z + 2$, we see that over $\mathbf{R}$, f factors as

$$f(x) = (x^2 - \sqrt{2}\,x + 2)(x^2 + \sqrt{2}\,x + 2).$$

8.34 Note: Historically, complex numbers first arose in the study of cubic equations. An equation of the form $ax^3+bx^2+cx+d = 0$, where $a, b, c, d \in \mathbf{C}$ with $a \neq 0$ can be solved as follows. First, divide by a to obtain an equation of the form $x^3+Bx^2+Cx+D = 0$. Next, make the substitution $x = y-\frac{B}{3}$ and rewrite the equation in the form $y^3+py+q = 0$. Then make the substitution $y = z - \frac{p}{3z}$ to convert the equation to the form $z^3 + q - \frac{p^3}{27}z^{-3} = 0$. Finally, multiply by z^3 to obtain $z^6+qz^3-\frac{p^3}{27}$ and solve for z^3 using the Quadratic Formula.

8.35 Example: Let $f(x) = x^3 + 3x^2 + 4x + 1$. Note that $f'(x) = 3x^2 + 6x + 4 = 3(x + 1)^2 + 1 > 0$, so f is increasing and hence has exactly one real root. Find the real root of f.

Solution: Let $x = y-1$. Then $x^3+3x^2+4x+1 = (y-1)^3+3(y-1)^2+4(y-1)+1 = y^3+y-1$. Let $y = z-\frac{1}{3z}$. Then $y^3+y-1 = (z-\frac{1}{3}z^{-1})^3+(z-\frac{1}{3}z^{-1})-1 = z^3-1-\frac{1}{27}z^{-3}$. We solve $z^6 - z^3 - \frac{1}{27} = 0$ using the quadratic formula, and obtain $z^3 = \frac{1\pm\sqrt{\frac{31}{27}}}{2}$. If $z = \sqrt[3]{\frac{1+\sqrt{\frac{31}{27}}}{2}}$ then $r\,z^{-1} = -\frac{1}{3}\sqrt[3]{\frac{2}{1+\sqrt{\frac{31}{27}}}} = -\frac{1}{3}\sqrt[3]{\frac{2(1-\sqrt{\frac{31}{27}})}{1-\frac{31}{27}}} = \sqrt[3]{\frac{1-\sqrt{\frac{31}{27}}}{2}}$. Similarly, if $z = \sqrt[3]{\frac{1-\sqrt{\frac{31}{27}}}{2}}$ then $r\,z^{-1} = \sqrt[3]{\frac{1+\sqrt{\frac{31}{27}}}{2}}$. In either case we have $y = z + rz^{-1} = \sqrt[3]{\frac{1+\sqrt{\frac{31}{27}}}{2}} + \sqrt[3]{\frac{1-\sqrt{\frac{31}{27}}}{2}}$, and $x = y-1 = \sqrt[3]{\frac{\sqrt{\frac{31}{27}}+1}{2}} - \sqrt[3]{\frac{\sqrt{\frac{31}{27}}-1}{2}} - 1$. (We did not use complex numbers in this example).

8.36 Example: Find the three real roots of $f(x) = x^3 - 3x + 1$.

Solution: Let $x = z + z^{-1}$ so that $f(x) = (z + z^{-1})^3 - 3(z + z^{-1}) + 1 = z^3 + 1 + z^{-3}$. Multiply by z^3 and solve $z^6 + z^3 + 1 = 0$ to get $z^3 = \frac{-1\pm\sqrt{3}\,i}{2} = e^{\pm i\,2\pi/3}$. If $z^3 = e^{i\,2\pi/3}$ then $z = e^{i\,2\pi/9}$, $e^{i\,8\pi/9}$ or $e^{i\,14\pi/9}$ and so $x = z + z^{-1} = z + \overline{z} = 2\mathrm{Re}\,(z) = 2\cos(\frac{2\pi}{9})$, $2\cos(\frac{8\pi}{9})$ or $2\cos(\frac{14\pi}{9})$. If $z^3 = e^{-i\,2\pi/3}$ then we obtain the same values for x. Thus the three real roots are $2\cos(40°)$, $-2\cos(20°)$ and $2\cos(80°)$.

.

Chapter 9. Cardinality

9.1 Definition: Let X and Y be sets and let $f : X \to Y$. Recall that the **domain** of f and the **range** of f are the sets

$$\text{Domain}(f) = X \ , \ \text{Range}(f) = f(X) = \big\{ f(x) \big| x \in X \big\} .$$

For $A \subseteq X$, the **image** of A under f is the set

$$f(A) = \big\{ f(x) \big| x \in A \big\}.$$

For $B \subseteq Y$, the **inverse image** of B under f is the set

$$f^{-1}(B) = \big\{ x \in X \big| f(x) \in B \big\}.$$

9.2 Definition: Let X, Y and Z be sets, let $f : X \to Y$ and let $g : Y \to Z$. We define the **composite** function $g \circ f : X \to Z$ by $(g \circ f)(x) = g\big(f(x)\big)$ for all $x \in X$.

9.3 Definition: We say that f is **injective** (or **one-to-one**, written as $1:1$) when for every $y \in Y$ there exists at most one $x \in X$ such that $f(x) = y$. Equivalently, f is injective when for all $x_1, x_2 \in X$, if $f(x_1) = f(x_2)$ then $x_1 = x_2$. We say that f is **surjective** (or **onto**) when for every $y \in Y$ there exists at least one $x \in X$ such that $f(x) = y$. Equivalently, f is surjective when $\text{Range}(f) = Y$. We say that f is **bijective** (or **invertible**) when f is both injective and surjective, that is when for every $y \in Y$ there exists exactly one $x \in X$ such that $f(x) = y$. When f is bijective, we define the **inverse** of f to be the function $f^{-1} : Y \to X$ such that for all $y \in Y$, $f^{-1}(y)$ is equal to the unique element $x \in X$ such that $f(x) = y$. Note that when f is bijective so is f^{-1}, and in this case we have $(f^{-1})^{-1} = f$.

9.4 Theorem: *Let $f : X \to Y$ and let $g : Y \to Z$. Then*

(1) if f and g are both injective then so is $g \circ f$,
(2) if f and g are both surjective then so is $g \circ f$, and
(3) if f and g are both invertible then so is $g \circ f$, and in this case $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Proof: To prove Part (1), suppose that f and g are both injective. Let $x_1, x_2 \in X$. If $g(f(x_1)) = g(f(x_2))$ then since g is injective we have $f(x_1) = f(x_2)$, and then since f is injective we have $x_1 = x_2$. Thus $g \circ f$ is injective.

To prove Part (2), suppose that f and g are surjective. Given $z \in Z$, since g is surjective we can choose $y \in Y$ so that $g(y) = z$, then since f is surjective we can choose $x \in X$ so that $f(x) = y$, and then we have $g(f(x)) = g(y) = z$. Thus $g \circ f$ is surjective.

Finally, note that Part (3) follows from Parts (1) and (2).

9.5 Definition: For a set X, we define the **identity function** on X to be the function $I_X : X \to X$ given by $I_X(x) = x$ for all $x \in X$. Note that for $f : X \to Y$ we have $f \circ I_X = f$ and $I_Y \circ f = f$.

9.6 Definition: Let X and Y be sets and let $f : X \to Y$. A **left inverse** of f is a function $g : Y \to X$ such that $g \circ f = I_X$. Equivalently, a function $g : Y \to X$ is a left inverse of f when $g\big(f(x)\big) = x$ for all $x \in X$. A **right inverse** of f is a function $h : Y \to X$ such that $f \circ h = I_Y$. Equivalently, a function $h : Y \to X$ is a right inverse of f when $f\big(h(y)\big) = y$ for all $y \in Y$.

9.7 Theorem: *Let X and Y be nonempty sets and let $f : X \to Y$. Then*

(1) f is injective if and only if f has a left inverse,
(2) f is surjective if and only if f has a right inverse, and
(3) f is bijective if and only if f has a left inverse g and a right inverse h, and in this case we have $g = h = f^{-1}$.

Proof: To prove Part (1), suppose first that f is injective. Since $X \neq \emptyset$ we can choose $a \in X$ and then define $g : Y \to X$ as follows: if $y \in \text{Range}(f)$ then (using the fact that f is 1:1) we define $g(y)$ to be the unique element $x_y \in X$ with $f(x_y) = y$, and if $y \notin \text{Range}(f)$ then we define $g(y) = a$. Then for every $x \in X$ we have $y = f(x) \in \text{Range}(f)$, so $g(y) = x_y = x$, that is $g\big(f(x)\big) = x$. Conversely, if f has a left inverse, say g, then f is 1:1 since for all $x_1, x_2 \in X$, if $f(x_1) = f(x_2)$ then $x_1 = g\big(f(x_1)\big) = g\big(f(x_2)\big) = x_2$.

To prove Part (2), suppose first that f is onto. For each $y \in Y$, choose $x_y \in X$ with $f(x_y) = y$, then define $g : X \to Y$ by $g(y) = x_y$ (we need the Axiom of Choice for this). Then g is a right inverse of f since for every $y \in Y$ we have $f\big(g(y)\big) = f(x_y) = y$. Conversely, if f has a right inverse, say g, then f is onto since given any $y \in Y$ we can choose $x = g(y)$ and then we have $f(x) = f\big(g(y)\big) = y$.

To prove Part (3), suppose first that f is bijective. The inverse function $f^{-1} : Y \to X$ is a left inverse for f because given $x \in X$ we can let $y = f(x)$ and then $f^{-1}(y) = x$ so that $f^{-1}\big(f(x)\big) = f^{-1}(y) = x$. Similarly, f^{-1} is a right inverse for f because given $y \in Y$ we can let x be the unique element in X with $y = f(x)$ and then we have $x = f^{-1}(y)$ so that $f\big(f^{-1}(y)\big) = f(x) = y$. Conversely, suppose that g is a left inverse for f and h is a right inverse for f. Since f has a left inverse, it is injective by Part (1). Since f has a right inverse, it is surjective by Part (2). Since f is injective and surjective, it is bijective. As shown above, the inverse function f^{-1} is both a left inverse and a right inverse. Finally, note that $g = f^{-1} = h$ because for all $y \in Y$ we have

$$g(y) = g\big(f\big(f^{-1}(y)\big)\big) = f^{-1}(y) = f^{-1}\big(f\big(h(y)\big)\big) = h(y)\,.$$

9.8 Corollary: *Let X and Y be sets. Then there exists an injective map $f : X \to Y$ if and only if there exists a surjective map $g : Y \to X$.*

Proof: Suppose $f : X \to Y$ is an injective map. Then f has a left inverse. Let g be a left inverse of f. Since $g \circ f = I_X$, we see that f is a right inverse of g. Since g has a right inverse, g is surjective. Thus there is a surjective map $g : Y \to X$. Similarly, if $g : Y \to X$ is surjective, then it has a right inverse $f : X \to Y$ which is injective.

9.9 Definition: Let A and B be sets. We say that A and B have the **same cardinality**, and we write $|A| = |B|$, when there exists a bijective map $f : A \to B$ (or equivalently when there exists a bijective map $g : Y \to X$). We say that the cardinality of A is **less than or equal to** the cardinality of B, and we write $|A| \leq |B|$, when there exists an injective map $f : A \to B$ (or equivalently when there exists a surjective map $g : Y \to X$). We say that the cardinality of A is **less than** the cardinality of B, and we write $|A| < |B|$, when $|A| \leq |B|$ and $|A| \neq |B|$, (that is when there exists an injective map $f : A \to B$ but there does not exist a bijective map $g : A \to B$). We also write $|A| \geq |B|$ when $|B| \leq |A|$ and $|A| > |B|$ when $|B| < |A|$.

9.10 Example: The map $f : \mathbf{N} \to 2\mathbf{N}$ given by $f(k) = 2k$ is bijective, so $|2\mathbf{N}| = |N|$. The map $g : \mathbf{N} \to \mathbf{Z}$ given by $g(2k) = k$ and $g(2k+1) = -k-1$ for $k \in \mathbf{N}$ is bijective, so we have $|\mathbf{Z}| = |\mathbf{N}|$. The map $h : \mathbf{N} \times \mathbf{N} \to \mathbf{N}$ given by $h(k,l) = 2^k(2l+1) - 1$ is bijective, so we have $|\mathbf{N} \times \mathbf{N}| = |\mathbf{N}|$.

9.11 Theorem: *For all sets A, B and C,*

(1) $|A| = |A|$,
(2) if $|A| = |B|$ *then* $|B| = |A|$,
(3) if $|A| = |B|$ *and* $|B| = |C|$ *then* $|A| = |C|$,
(4) $|A| \leq |B|$ *if and only if* ($|A| = |B|$ *or* $|A| < |B|$), *and*
(5) if $|A| \leq |B|$ *and* $|B| \leq |C|$ *then* $|A| \leq |C|$.

Proof: Part (1) holds because the identity function $I_A : A \to A$ is bijective. Part (2) holds because if $f : A \to B$ is bijective then so is $f^{-1} : B \to A$. Part (3) holds because if $f : A \to B$ and $g : B \to C$ are bijective then so is the composite $g \circ f : A \to C$. The rest of the proof is left as an exercise.

9.12 Definition: Let A be a set. For each $n \in \mathbf{N}$, let $S_n = \{0, 1, 2, \cdots, n-1\}$. For $n \in \mathbf{N}$, we say that the cardinality of A is equal to n, or that A **has** n **elements**, and we write $|A| = n$, when $|A| = |S_n|$. We say that A is **finite** when $|A| = n$ for some $n \in \mathbf{N}$. We say that A is **infinite** when A is not finite. We say that A is **countable** when $|A| = |\mathbf{N}|$.

9.13 Note: When a set A is finite with $|A| = n$, and when $f : A \to S_n$ is a bijection, if we let $a_k = f^{-1}(k)$ for each $k \in S_n$ then we have $A = \{a_0, a_1, \cdots, a_{k-1}\}$ with the elements a_k distinct. Conversely, if $A = \{a_0, a_1, \cdots, a_{k-1}\}$ with the elements a_k all distinct, then we define a bijection $f : A \to S_n$ by $f(a_k) = k$. Thus we see that A is finite with $|A| = n$ if and only if A is of the form $A = \{a_0, a_1, \cdots, a_{n-1}\}$ with the elements a_k all distinct. Similarly, a set A is countable if and only if A is of the form $A = \{a_0, a_1, a_2, \cdots\}$ with the elements a_k all distinct.

9.14 Note: For $n \in \mathbf{N}$, if A is a finite set with $|A| = n + 1$ and $a \in A$ then $|A \setminus \{a\}| = n$. Indeed, if $A = \{a_0, a_1, \cdots, a_n\}$ with the elements a_i distinct, and if $a = a_k$ so that we have $A \setminus \{a\} = \{a_0, a_1, \cdots, a_{k-1}, a_{k+1}, \cdots, a_n\}$, then we can define a bijection $f : S_n \to A \setminus \{a\}$ by $f(i) = a_i$ for $0 \leq i < k$ and $f(i) = a_{i+1}$ for $k \leq i < n$.

9.15 Theorem: *Let A be a set. Then the following are equivalent.*

(1) A is infinite.
(2) A contains a countable subset.
(3) $|\mathbf{N}| \leq |A|$
(4) There exists a map $f : A \to A$ which is injective but not surjective.

Proof: To prove that (1) implies (2), suppose that A is infinite. Since $A \neq \emptyset$ we can choose an element $a_0 \in A$. Since $A \neq \{a_0\}$ we can choose an element $a_1 \in A \setminus \{a_0\}$. Since $A \neq \{a_0, a_1\}$ we can choose $a_3 \in A \setminus \{a_0, a_1\}$. Continue this procedure: having chosen distinct elements $a_0, a_1, \cdots, a_{n-1} \in A$, since $A \neq \{a_0, a_1, \cdots, a_{n-1}\}$ we can choose $a_n \in A \setminus \{a_0, a_1, \cdots, a_{n-1}\}$. In this way, we obtain a countable set $\{a_0, a_1, a_2, \cdots\} \subseteq A$.

Next we show that (2) is equivalent to (3). Suppose that A contains a countable subset, say $\{a_0, a_1, a_2, \cdots\} \subseteq A$ with the element a_i distinct. Since the a_i are distinct, the map $f : \mathbf{N} \to A$ given by $f(k) = a_k$ is injective, and so we have $|\mathbf{N}| \leq |A|$. Conversely, suppose that $|\mathbf{N}| \leq |A|$, and chose an injective map $f : \mathbf{N} \to A$. Considered as a map from $\mathbf{N}$ to $f(\mathbf{N})$, f is bijective, so we have $|\mathbf{N}| = |f(\mathbf{N})|$ hence $f(\mathbf{N})$ is a countable subset of A.

Next, let us show that (2) implies (4). Suppose that A has a countable subset, say $\{a_0, a_1, a_2, \cdots\} \subseteq A$ with the element a_i distinct. Define $f : A \to A$ by $f(a_k) = a_{k+1}$ for all $k \in \mathbf{N}$ and by $f(b) = b$ for all $b \in A \setminus \{a_0, a_1, a_2, \cdots\}$. Then f is injective but not surjective (the element a_0 is not in the range of f).

Finally, to prove that (4) implies (1) we shall prove that if A is finite then every injective map $f : A \to A$ is surjective. We prove this by induction on the cardinality of A. The only set A with $|A| = 0$ is the set $A = \emptyset$, and then the only function $f : A \to A$ is the empty function, which is surjective. Since that base case may appear too trivial, let us consider the next case. Let $n = 1$ and let A be a set with $|A| = 1$, say $A = \{a\}$. The only function $f : A \to A$ is the function given by $f(a) = a$, which is surjective. Let $n \geq 1$ and suppose, inductively, that for every set A with $|A| = n$, every injective map $f : A \to A$ is surjective. Let B be a set with $|B| = n + 1$ and let $g : B \to B$ be injective. Suppose, for a contradiction, that g is not surjective. Choose an element $b \in B$ which is not in the range of g so that we have $g : B \to B \setminus \{b\}$. Let $A = B \setminus \{b\}$ and let $f : A \to A$ be given by $f(x) = g(x)$ for all $x \in A$. Since $g : B \to A$ is injective and $f(x) = g(x)$ for all $x \in A$, f is also injective. Again since g is injective, there is no element $x \in B \setminus \{b\}$ with $g(x) = g(b)$, so there is no element $x \in A$ with $f(x) = g(b)$, and so f is not surjective. Since $|A| = n$ (by the above note), this contradicts the induction hypothesis. Thus f must be surjective. By the Principle of Induction, for every $n \in \mathbf{N}$ and for every set A with $|A| = n$, every injective function $f : A \to A$ is surjective.

9.16 Corollary: *Let A and B be sets.*

(1) If A is countable then A is infinite.
(2) When $|A| \leq |B|$, if B is finite then so is A (equivalently if A is infinite then so is B).
(3) If $|A| = n$ and $|B| = m$ then $|A| = |B|$ if and only if $n = m$.
(4) If $|A| = n$ and $|B| = m$ then $|A| \leq |B|$ if and only if $n \leq m$.
(5) When one of the two sets A and B is finite, if $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$.

Proof: Part (1) is immediate: if A is countable then A contains a countable subset (itself), so A is infinite, by Theorem 4.15.

To prove Part (2), suppose that $|A| \leq |B$ and that $|A|$ is infinite. Since A is infinite, we have $|\mathbf{N}| \leq |A|$ (by Theorem 4.15). Since $|\mathbf{N}| \leq |A|$ and $|A| \leq |B|$ we have $|\mathbf{N}| \leq |B|$ (by Theorem 4.11). Since $|\mathbf{N}| \leq |B|$, B is infinite (by Theorem 4.15 again).

To Prove Part (3), suppose that $|A| = n$ and $|B| = m$. If $n = m$ then we have $S_n = S_m$ and so $|A| = |S_n| = |S_m| = |B|$. Conversely, suppose that $|A| = |B|$. Suppose, for a contradiction, that $n \neq m$, say $n > m$, and note that $S_m \subsetneqq S_n$. Since $|A| = |B|$ we have $|S_n| = |A| = |B| = |S_m|$ so we can choose a bijection $f : S_n \to S_m$. Since $S_m \subsetneqq S_n$, we can consider f as a function $f : S_n \to S_n$ which is injective but not surjective. This contradicts Theorem 4.16, and so we must have $n = m$. This proves Part (3).

To prove Part (4), we again suppose that $|A| = n$ and $|B| = m$. If $n \leq m$ then $S_n \subseteq S_m$ so the inclusion map $I : S_n \to S_m$ is injective and we have $|A| = |S_n| \leq |S_m| = |B|$. Conversely, suppose that $|A| \leq |B|$ and suppose, for a contradiction, that $n > m$. Since $|A| \leq |B|$ we have $|S_n| = |A| \leq |B| = |S_m|$ so we can choose an injective map $f : S_n \to S_m$. Since $n > m$ we have $S_m \subsetneqq S_n$ so we can consider f as a map $f : S_n \to S_n$, and this map is injective but not surjective. This contradicts Theorem 2.16, and so $n \leq m$.

Finally, to prove Part (5) we suppose that one of the two sets A and B is finite, and that $|A| \leq |B|$ and $|B| \leq |A|$. If A is finite then, since $|B| \leq |A|$, Part (2) implies that B is finite. If B is finite then, since $|A| \leq |B|$, Part (2) implies that A is finite. Thus, in either case, we see that A and B are both finite. Since A and B are both finite with $|A| \leq |B|$ and $|B| \leq |A|$, we must have $|A| = |B|$ by Parts (3) and (4).

9.17 Theorem: *Let A be a set. Then $|A| \leq |\mathbf{N}|$ if and only if A is finite or countable.*

Proof: First we claim that every subset of $\mathbf{N}$ is either finite or countable. Let $A \subseteq \mathbf{N}$ and suppose that A is not finite. Since $A \neq \emptyset$, we can set $a_0 = \min A$ (using the Well-Ordering Property of $\mathbf{N}$). Note that $\{0, 1, \cdots, a_0\} \cap A = \{a_0\}$. Since $A \neq \{a_0\}$ (so the set $A \setminus \{a_0\}$ is nonempty) we can set $a_1 = \min A \setminus \{a_0\}$. Then we have $a_0 < a_1$ and $\{0, 1, 2, \cdots, a_1\} \cap A = \{a_0, a_1\}$. Since $A \neq \{a_0, a_1\}$ we can set $a_2 = \min A \setminus \{a_0, a_1\}$. Then we have $a_0 < a_1 < a_2$ and $\{0, 1, 2, \cdots, a_3\} \cap A = \{a_0, a_1, a_2\}$. We continue the procedure: having chosen $a_0, a_1, \cdots, a_{n-1} \in A$ with $a_0 < a_1 < \cdots < a_{n-1}$ such that $A \cap \{0, 1, \cdots, a_{n-1}\} = \{a_0, a_1, \cdots, a_{n-1}\}$, since $A \neq \{a_0, a_1, \cdots, a_{n-1}\}$ we can set $a_n = \min A \setminus \{a_0, a_1, \cdots, a_{n-1}\}$, and then we have $a_0 < a_1 < \cdots < a_{n-1} < a_n$ and $A\{0, 1, 2, \cdots, a_n\} \cap A = \{a_0, a_1, \cdots, a_n\}$. In this way, we obtain a countable set $\{a_0, a_1, a_2, \cdots\} \subseteq A$ with $a_0 < a_1 < a_2 < \cdots$ with the property that for all $m \in \mathbf{N}$, $\{0, 1, 2, \cdots, a_m\} \cap A = \{a_0, a_1, \cdots, a_m\}$. Since $0 \leq a_0 < a_1 < a_2 < \cdots$, it follows (by induction) that $a_k \geq k$ for all $k \in \mathbf{N}$. It follows in turn that $A \subseteq \{a_0, a_1, a_2 \cdots\}$ because given $m \in A$, since $m \leq a_m$ we have

$$m \in \{0, 1, 2, \cdots, m\} \cap A \subseteq \{0, 1, 2, \cdots, a_m\} \cap A = \{a_0, a_1, \cdots, a_m\}.$$

Thus $A = \{a_0, a_1, a_2, \cdots\}$ and the elements a_i are distinct, so A is countable. This proves our claim that every subset of $\mathbf{N}$ is either finite or countable.

Now suppose that $|A| \leq |N|$ and choose an injective map $f : A \to \mathbf{N}$. Since f is injective, when we consider it as a map $f : A \to f(A)$, it is bijective, and so $|A| = |f(A)|$. Since $f(A) \subseteq \mathbf{N}$, the previous paragraph shows that $f(A)$ is either finite or countable. If $f(A)$ is finite with $|f(A)| = n$ then $|A| = |f(A)| = |S_n|$, and if $f(A)$ is countable then we have $|A| = |f(A)| = |\mathbf{N}|$. Thus A is finite or countable.

9.18 Theorem: *Let A be a set. Then*

(1) $|A| < |\mathbf{N}|$ if and only if A is finite,
(2) $|\mathbf{N}| < |A|$ if and only if A is neither finite nor countable, and
(3) if $|A| \leq |\mathbf{N}|$ and $|\mathbf{N}| \leq |A|$ then $|A| = |\mathbf{N}|$.

Proof: Part (1) follows from Theorem 4.15 because

$$\begin{aligned} |A| < |\mathbf{N}| &\iff (|A| \leq |\mathbf{N}| \text{ and } |A| \neq |\mathbf{N}|) \\ &\iff (A \text{ is finite or countable and } A \text{ is not countable}) \\ &\iff A \text{ is finite} \end{aligned}$$

and Part (2) follows from Theorem 4.17 because

$$\begin{aligned} |\mathbf{N}| < |A| &\iff (|\mathbf{N}| \leq |A| \text{ and } |\mathbf{N}| \neq |A|) \\ &\iff (A \text{ is not finite and } A \text{ is not countable.}) \end{aligned}$$

To prove Part (3), suppose that $|A| \leq |\mathbf{N}|$ and $|\mathbf{N}| \leq |A|$. Since $|A| \leq |\mathbf{N}|$, we know that A is finite or countable by Theorem 4.17. Since $|\mathbf{N}| \leq |A|$, we know that that A is infinite by Theorem 4.15. Since A is finite or countable and A is not finite, it follows that A is countable. Thus $|A| = |\mathbf{N}|$.

9.19 Definition: Let A be a set. When A is countable we write $|A| = \aleph_0$. When A is finite we write $|A| < \aleph_0$. When A is infinite we write $|A| \geq \aleph_0$. When A is either finite or countable we write $|A| \leq \aleph_0$ and we say that A is **at most countable**. when A is neither finite nor countable we write $|A| > \aleph_0$ and we say that A is **uncountable**.

9.20 Theorem:

(1) If A and B are countable sets, then so is $A \times B$.
(2) If A and B are countable sets, then so is $A \cup B$.
(3) If $A_0, A_1, A_2, \cdots$ are countable sets, then so is $\bigcup_{k=0}^{\infty} A_k$.
(4) $\mathbf{Q}$ is countable.

Proof: To prove Parts (1) and (2), let $A = \{a_0, a_1, a_2, \cdots\}$ with the a_i distinct and let $B = \{b_0, b_1, b_2, \cdots\}$ with the b_i distinct. Since every positive integer can be written uniquely in the form $2^k(2l+1)$ with $k, l \in \mathbf{N}$, the map $f : A \times B \to \mathbf{N}$ given by $f(a_k, b_l) = 2^k(2l+1)-1$ is bijective, and so $|A \times B| = |\mathbf{N}|$. This proves Part (1). Since the map $g : \mathbf{N} \to A \cup B$ given by $g(k) = a_k$ is injective, we have $|\mathbf{N}| \leq |A \cup B|$. Since the map $h : \mathbf{N} \to A \cup B$ given by $h(2k) = a_k$ and $h(2k+1) = b_k$ is surjective, we have $|A \cup B| \leq |N|$. Since $|\mathbf{N}| \leq |A \cup B|$ and $|A \cup B| \leq |\mathbf{N}|$, we have $|A \cup B| = |\mathbf{N}|$ by Part (3) of Theorem 4.18. This proves (2).

To prove Part (3), for each $k \in \mathbf{N}$, let $A_k = \{a_{k0}, a_{k1}, a_{k2}, \cdots\}$ with the a_{ki} distinct. Since the map $f : \mathbf{N} \to \bigcup_{k=0}^{\infty} A_k$ given by $f(k) = a_{0,k}$ is injective, $|\mathbf{N}| \leq \left|\bigcup_{k=0}^{\infty} A_k\right|$. Since $\mathbf{N} \times \mathbf{N}$ is countable by Part (1), and since the map $g : \mathbf{N} \times \mathbf{N} \to \bigcup_{k=0}^{\infty} A_k$ given by $g(k, l) = a_{k,l}$ is surjective, we have $\left|\bigcup_{k=0}^{\infty} A_k\right| \leq |\mathbf{N} \times \mathbf{N}| = |\mathbf{N}|$. By Part (3) of Theorem 4.18, we have $\left|\bigcup_{k=0}^{\infty} A_k\right| = |\mathbf{N}|$, as required.

Finally, we prove Part (4). Since the map $f : \mathbf{N} \to \mathbf{Q}$ given by $f(k) = k$ is injective, we have $|\mathbf{N}| \leq |\mathbf{Q}|$. Since the map $g : \mathbf{Q} \to \mathbf{Z} \times \mathbf{Z}$, given by $g\left(\frac{a}{b}\right) = (a, b)$ for all $a, b \in \mathbf{Z}$ with $b > 0$ and $\gcd(a, b) = 1$, is injective, and since $\mathbf{Z} \times \mathbf{Z}$ is countable, we have $|\mathbf{Q}| \leq |\mathbf{Z} \times \mathbf{Z}| = |\mathbf{N}|$. Since $|\mathbf{N}| \leq |\mathbf{Q}|$ and $|\mathbf{Q}| \leq |\mathbf{N}|$, we have $|\mathbf{Q}| = |\mathbf{N}|$, as required.

9.21 Definition: For a set A, let $\mathcal{P}(A)$ denote the **power set** of A, that is the set of all subsets of A, and let 2^A denote the set of all functions from A to $S_2 = \{0, 1\}$.

9.22 Theorem:

(1) For every set A, $\left|\mathcal{P}(A)\right| = \left|2^A\right|$.
(2) For every set A, $|A| < \left|\mathcal{P}(A)\right|$.
(3) $\mathbf{R}$ is uncountable.

Proof: Let A be any set. Define a map $g : \mathcal{P}(A) \to 2^A$ as follows. Given $S \in \mathcal{P}(A)$, that is given $S \subseteq A$, we define $g(S) \in 2^A$ to be the map $g(S) : A \to \{0, 1\}$ given by

$$g(S)(a) = \begin{cases} 1 \text{ if } a \in S, \\ 0 \text{ if } a \notin S. \end{cases}$$

Define a map $h : 2^A \to \mathcal{P}(A)$ as follows. Given $f \in 2^A$, that is given a map $f : A \to \{0, 1\}$, we define $h(f) \in \mathcal{P}(A)$ to be the subset

$$h(f) = \left\{a \in A \middle| f(a) = 1\right\} \subseteq A.$$

The maps g and h are the inverses of each other because for every $S \subseteq A$ and every $f : A \to \{0, 1\}$ we have

$$f = g(S) \Longleftrightarrow \forall a \in A \;\; f(a) = g(S)(a) \Longleftrightarrow \forall a \in A \;\; f(a) = \begin{cases} 1 \text{ if } a \in S, \\ 0 \text{ if } a \notin S, \end{cases}$$

$$\Longleftrightarrow \forall a \in A \; \big(f(a) = 1 \Longleftrightarrow a \in S\big) \Longleftrightarrow \left\{a \in A \middle| f(a) = 1\right\} = S \Longleftrightarrow h(f) = S.$$

This completes the proof of Part (1).

Let us prove Part (2). Again we let A be any set. Since the the map $f : A \to \mathcal{P}(A)$ given by $f(a) = \{a\}$ is injective, we have $|A| \leq \left|\mathcal{P}(A)\right|$. We need to show that $|A| \neq \left|\mathcal{P}(A)\right|$.

Let $g : A \to \mathcal{P}(A)$ be any map. Let $S = \{a \in A | a \notin g(a)\}$. Note that S cannot be in the range of g because if we could choose $a \in A$ so that $g(a) = S$ then, by the definition of S, we would have $a \in S \Longleftrightarrow a \notin g(a) \Longleftrightarrow a \notin S$ which is not possible. Since S is not in the range of g, the map g is not surjective. Since g was an arbitrary map from A to $\mathcal{P}(A)$, it follows that there is no surjective map from A to $\mathcal{P}(A)$. Thus there is no bijective map from A to $\mathcal{P}(A)$ and so we have $|A| \neq \big|\mathcal{P}(A)\big|$, as desired.

Finally, we shall prove that $\mathbf{R}$ is uncountable using the fact (which we did not prove) that every real number has a unique decimal expansion which does not end with an infinite string of 9's. We define a map $g : 2^{\mathbf{N}} \to \mathbf{R}$ as follows. Given $f \in 2^{\mathbf{N}}$, that is given a map $f : \mathbf{N} \to \{0, 1\}$, we define $g(f)$ to be the real number $g(f) \in [0, 1)$ with the decimal expansion $g(f) = 0.f(0)f(1)f(2)f(3)\cdots$ (for those who have seen infinite series, this is the number $g(f) = \sum_{k=0}^{\infty} f(k)10^{-k-1}\big)$. By the uniqueness of decimal expansions, the map g is injective, so we have $\big|2^{\mathbf{N}}\big| \leq |\mathbf{R}|$. Thus $|\mathbf{N}| < \big|\mathcal{P}(\mathbf{N})\big| = \big|2^{\mathbf{N}}\big| \leq |\mathbf{R}|$, and so $\mathbf{R}$ is uncountable, by Part (2) of Theorem 4.18.

9.23 Theorem: *(Cantor - Schroeder - Bernstein) Let A and B be sets. Suppose that $|A| \leq |B|$ and $|B| \leq |A|$. Then $|A| = |B|$*

Proof: We sketch a proof. Choose injective functions $f : A \to B$ and $g : B \to A$. Since the functions $f : A \to f(A)$, $g : B \to g(B)$ and $f : g(B) \to f\big(g(B)\big)$ are bijective we have $|A| = |f(A)|$ and $|B| = |g(B)| = \big|f(g(B))\big|$. Also note that $f\big(g(B)\big) \subseteq f(A) \subseteq B$. Let $X = f\big(g(B)\big)$, $Y = f(A)$ and $Z = B$. Then we have $X \subseteq Y \subseteq Z$ and we have $|X| = |Z|$ and we need to show that $|Y| = |Z|$. The composite $h = f \circ g : Z \to X$ is a bijection. Define sets Z_n and Y_n for $n \in \mathbf{N}$ recursively by

$$Z_0 = Z,\ Z_n = h(Z_{n-1}) \ \text{ and } \ Y_0 = Y\ ,\ Y_n = h(Y_{n-1}).$$

Since $Y_0 = Y$, $Z_0 = Z$, $Z_1 = h(Z_0) = h(Z) = X$ and $X \subseteq Y \subseteq Z$, we have

$$Z_1 \subseteq Y_0 \subseteq Z_0.$$

Also note that for $1 \leq n \in \mathbf{N}$,

$$Z_n \subseteq Y_{n-1} \subseteq Z_{n-1} \Longrightarrow h(Z_n) \subseteq h(Y_{n-1}) \subseteq h(Z_{n-1}) \Longrightarrow Z_{n+1} \subseteq Y_n \subseteq Z_n.$$

By the Induction Principle, it follows that $Z_n \subseteq Y_{n-1} \subseteq Z_{n-1}$ for all $n \geq 1$, so we have

$$Z_0 \supseteq Y_0 \supseteq Z_1 \supseteq Y_1 \supseteq Z_2 \supseteq Y_2 \supseteq \cdots$$

Let $U_n = Z_n \setminus Y_n$, $U = \bigcup_{n=1}^{\infty} U_n$ and $V = Z \setminus U$. Define $H : Z \to Y$ by

$$H(x) = \begin{cases} h(x) \text{ if } x \in U, \\ x \quad \text{if } x \in V. \end{cases}$$

Verify that H is bijective.

9.24 Example: Show that $|\mathbf{R}| = \left|2^{\mathbf{N}}\right|$.

Solution: $g : 2^{\mathbf{N}} \to \mathbf{R}$ as follows: for $f \in 2^{\mathbf{N}}$ we let $g(f)$ be the real number $g(f) \in [0, 1)$ with decimal expansion $g(f) = 0.f(0)f(1)f(2)\cdots$. Then g is injective so $\left|2^{\mathbf{N}}\right| \leq |\mathbf{R}|$. Define $h : 2^{\mathbf{N}} \to [0, 1)$ as follows: for $f \in 2^{\mathbf{N}}$ let $h(f)$ be the real number $h(f) \in [0, 1]$ with binary expansion $h(f) = 0.f(0)f(1)f(2)\cdots$. Then h is surjective so we have $\left|[0, 1]\right| \leq \left|2^{\mathbf{N}}\right|$. The map $k : \mathbf{R} \to [0, 1]$ given by $k(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1} x$ is injective so we have $|\mathbf{R}| \leq \left|[0, 1]\right|$. Since $|\mathbf{R}| \leq \left|[0, 1]\right| \leq \left|2^{\mathbf{N}}\right|$ and $\left|2^{\mathbf{N}}\right| \leq |\mathbf{R}|$, we have $|\mathbf{R}| = \left|2^{\mathbf{N}}\right|$ by the Cantor-Schroeder-Bernstein Theorem.

9.25 Theorem: *Let A and B be finite sets, let A^B be the set of all functions $f : A \to B$, and let $\mathcal{P}(A)$ be the power set of A (that is the set of all subsets of A). Then*

(1) if A and B are disjoint then $|A \cup B| = |A| \cup |B|$,
(2) $|A \times B| = |A| \cdot |B|$,
(3) $\left|A^B\right| = |A|^{|B|}$, and
(3) $|\mathcal{P}(A)| = 2^{|A|}$.

Proof: The proof is left as an exercise.

.

Chapter 10. Factorization in Rings

10.1 Definition: Let R be a commutative ring. Let $a, b \in R$. We say that a **divides** b (or a is a **divisor** or **factor** of b, or b is a **multiple** of a), and we write $a|b$, when $b = ar$ for some $r \in R$. We say that a and b are **associates**, and we write $a \sim b$, when $a|b$ and $b|a$.

10.2 Theorem: *Let R be a commutative ring. Let $a, b \in R$. Then*

(1) $a \sim b$ if and only if a and b have the same multiples and divisors,
(2) $a \sim 0$ if and only if $a = 0$,
(3) $a \sim 1$ if and only if a is a unit.
(4) if R is an integral domain then $a \sim b$ if and only if $b = au$ for some unit $u \in R$.

Proof: The proof is left as an exercise.

10.3 Definition: Let R be a commutative ring. Let $a \in R$ be a non-zero non-unit. We say that a is **reducible** when $a = bc$ for some non-units $b, c \in R$, and otherwise we say that a is **irreducible**. Note that if a is irreducible then the divisors of a are the units and the associates of a. We say that a is **prime** when for all $b, c \in R$, if $a|bc$ then either $a|b$ or $a|c$.

10.4 Theorem: *Let R be a commutative ring. Let $a, b \in R$ with $a \sim b$. Then*

(1) a is reducible if and only if b is reducible,
(2) a is irreducible if and only if b is irreducible,
(3) a is prime if and only if b is prime.

Proof: The proof is left as an exercise.

10.5 Theorem: *Let R be an integral domain. Then every prime element in R is also irreducible.*

Proof: The proof is left as an exercise.

10.6 Exercise: Find all primes and irreducible elements in $\mathbf{Z}_{12}$.

10.7 Exercise: Use the method of the Sieve of Eratosthenes to find several irreducible elements in $\mathbf{Z}[\sqrt{3}\,i]$ and also some irreducible elements which are not prime.

10.8 Definition: A **Euclidean domain** (or ED) is an integral domain R together with a function $N : R \to \mathbf{N}$, called a **Euclidean norm**, such that

N1. for all $a \in R$ we have $N(a) = 0 \Longleftrightarrow a = 0$,
N2. for all $a \in R$ we have $N(a) = 1 \Longleftrightarrow a$ is a unit,
N3. for all nonzer nonunits $a, b, c \in R$, if $a = bc$ then $N(b) < N(a)$ and $N(c) < N(a)$, and
N4. for all $a, b \in R$ with $b \neq 0$ there exist $q, r \in R$ such that $a = qb + r$ and $N(r) < N(b)$.

10.9 Definition: A **unique factorization domain** (or UFD) is an integral domain R with the property that for every nonzero non-unit $a \in R$ we have

(1) $a = p_1 p_2 \cdots p_\ell$ for some $\ell \in \mathbf{Z}^+$ and some irreducible elements $p_i \in R$, and
(2) if $a = p_1 p_2 \cdots p_\ell = q_1 q_2 \cdots q_m$ where $\ell, m \in \mathbf{Z}^+$ and each p_i and q_j is irreducible, then $m = \ell$ and for some bijection $\sigma : \{1, 2, \cdots, \ell\} \to \{1, 2, \cdots, \ell\}$ we have $a_i \sim b_{\sigma(i)}$ for all i.

10.10 Theorem: *Every Euclidean domain is a unique factorization domain*

Proof: The proof is almost identical to the proof of unique factorization in $\mathbf{Z}$.

10.11 Example: $\mathbf{Z}$ is a Euclidean domain using the Euclidean norm $N : \mathbf{Z} \to \mathbf{N}$ given by $N(a) = |a|$.

10.12 Example: Every field is a Euclidean domain, using the function $N : R \to \mathbf{N}$ given by $N(0) = 0$ and $N(a) = 1$ for all $a \neq 0$.

10.13 Example: If F is a field then $F[x]$ is a Euclidean domain with Euclidean norm $N(f) = \deg(f) + 1$ (where we follow the convention that $\deg(0) = -1$ so that $N(0) = 0$). The fact that N satisfies Property N4 follows from the Division Algorithm proven below.

10.14 Exercise: Show that for each $d \in \{-2, -1, 2, 3\}$ the ring $\mathbf{Z}\big[\sqrt{d}\,\big]$ is a Euclidean domain with Euclidean norm given by $N(a + b\sqrt{d}) = \big|a^2 - db^2\big|$.

10.15 Exercise: Show that the rings $\mathbf{Z}\big[\sqrt{3}\,i\big]$ and $\mathbf{Z}\big[\sqrt{5}\,\big]$ are not unique factorization domains.

10.16 Note: Here are a few remarks about polynomials. Recall that $R[x]$ denotes the ring of polynomials with coefficients in the ring R, and R^R denotes the ring of all functions $f : R \to R$.

(1) A polynomial $f \in R[x]$ determines a function $f \in R^R$. Given $f(x) = \sum_{i=0}^{n} a_i x^i \in R[x]$ we obtain the function $f : R \to R$ given by $f(x) = \sum_{i=0}^{n} a_i x^i$.

(2) Although we do not usually distinguish notationally between the polynomial $f \in R[x]$ and its corresponding function $f \in R^R$, they are not always identical. If the ring R is not commutative then multiplication of polynomials does not agree with multiplication of functions. For $f, g \in R[x]$ given by $f(x) = a + bx$ and $g(x) = c + dx$, in the ring $R[x]$ we have $(fg)(x) = (a + bx)(c + dx) = (ac) + (ad + bc)x + (bd)x^2$, but in the ring R^R we have $(fg)(x) = (a + bx)(c + dx) = ac + adx + bxc + bxdx$.

(3) Equality of polynomials does not agree with equality of functions. For $f, g \in R[x]$ given by $f(x) = \sum_{i=0}^{n} a_i x^i$ and $g(x) = \sum_{i=0}^{m} b_i x^i$ we have $f = g \in R[x]$ if and only if $a_i = b_i$ for all i (and if say $n < m$ then $b_i = a_i = 0$ for $i > n$), but $f = g \in R^R$ if and only if $f(x) = g(x)$ for all $x \in R$. These two notions of equality do not always agree. For example if R is finite then the ring $R[x]$ is infinite but the ring R^R is finite. Indeed if $|R| = n$ then $R[x]$ is countably infinite but $\big|R^R\big| = n^n$. For a more specific example, if $f(x) = x^p - x$ then we have $f \neq 0 \in \mathbf{Z}_p[x]$ (because its coefficients are not equal to zero) but $f = 0 \in \mathbf{Z}_p{}^{\mathbf{Z}_p}$ because, by Fermat's Little Theorem, we have $f(x) = 0$ for all $x \in \mathbf{Z}_p$.

(4) Recall that for $f(x) = \sum_{i=0}^{n} a_i x^i$ with each $a_i \in R$ and $a_n \neq 0$, the element $a_n \in R$ is called the leading coefficient of f, and the positive integer n is called the degree of $f(x)$, and we write $\deg(f) = n$. For convenience, we also define $\deg(0) = -1$. When R is an integral domain, it is easy to see that for $0 \neq f, g \in R[x]$ we have $\deg(fg) = \deg(f) + \deg(g)$. When R is not an integral domain, however, we only have $\deg(fg) \leq \deg(f) + \deg(g)$ because the product of the two leading coefficients can be equal to zero.

(5) When R is an integral domain, because we have $\deg(fg) = \deg(f) + \deg(g)$ for all $0 \neq f, g \in R[x]$, it is easy to see that the units in $R[x]$ are the constant polynomials $f(x) = c$ where c is a unit in R. In particular, when F is a field, the units in $F[x]$ are the elements $f \in F[x]$ with $\deg(f) = 0$.

10.17 Example: In the ring $\mathbf{Z}_4[x]$ we have $(1+2x)^2 = 1+4x+4x^2 = 1$, so $f(x) = (1+2x)$ is a unit in $\mathbf{Z}_4[x]$.

10.18 Theorem: *(Division Algorithm) Let R be a ring. Let $f, g \in R[x]$ and suppose that the leading coefficient of g is a unit in R. Then there exist unique polynomials $q, r \in R$ such that $f = qg + r$ and $\deg(r) < \deg(g)$.*

Proof: First we prove existence. If $\deg(f) < \deg(g)$ then we can take $q = 0$ and $r = f$. Suppose that $\deg(f \geq \deg(g)$, Say $f(x) = \sum_{i=0}^{n} a_i x^i$ with $a_i \in R$ and $a_n \neq 0$ and $g(x) = \sum_{i=0}^{m} b_i x^i$ with $b_i \in R$ and b_m is a unit. Note that the polynomial $a_n {b_m}^{-1} x^{n-m} g(x)$ has degree n and leading coefficient a_n. It follows that the polynomial $f(x) - a_n {b_m}^{-1} x^{n-m} g(x)$ has degree smaller than n (because the leading coefficients cancel). We can suppose, inductively, that there exist polynomials $p, r \in R[x]$ such that $f(x) - a_n {b_m}^{-1} x^{n-m} g(x) = p(x)g(x) + r(x)$ and $\deg(r) < \deg(g)$. Then we have $f = qg + r$ by taking $q(x) = a_n {b_m}^{-1} x^{n-m} - p(x)$.

Next we prove uniqueness. Suppose that $f = qg + r = pg + s$ where $q, p, r, s \in R[x]$ with $\deg(r) < \deg(g)$ and $\deg(s) < \deg(g)$. Then we have $(q-p)g = s - r$ and so $\deg\big((q-p)g\big) = \deg(s-r)$. Since the leading coefficient of g is a unit (hence not a zero divisor), it follows that $\deg\big((q-p)g\big) = \deg(q-p) + \deg(g)$. If we had $q - p \neq 0$ then we would have $\deg\big((q-p)g\big) \geq \deg(g)$ but $\deg(s-r) < \deg(g)$, giving a contradiction. Thus we must have $q - p = 0$. Since $q - p = 0$ we have $s - r = (q-p)g = 0$. Since $q - p = 0$ and $s - r = 0$ we have $q = p$ and $r = s$, proving uniqueness.

10.19 Corollary: *(The Remainder Theorem) Let R be a ring, let $f \in R[x]$, and let $a \in R$. When we divide $f(x)$ by $(x-a)$ to obtain the quotient $q(x)$ and remainder $r(x)$, the remainder is the constant polynomial $r(x) = f(a)$.*

Proof: Use the division algorithm to obtain $q, r \in R[x]$ such that $f = q(x)(x-a) + r(x)$ and $\deg(r) < \deg(x-a)$. Since $\deg(x-a) = 1$ we have $\deg(r) \in \{-1, 0\}$, and so r is a constant polynomial, say $r(x) = c$ with $c \in R$. Then we have $f(x) = q(x)(x-a) + c$. Put in $x = a$ to get $f(a) = q(a)(a-a) + c = q(a) \cdot 0 + c = c$.

10.20 Corollary: *(The Factor Theorem) Let R be a commutative ring, let $f \in R[x]$ and let $a \in R$. Then $f(a) = 0$ if and only if $(x-a)\big|f(x)$.*

Proof: Suppose that $f(a) = 0$. Choose $q, r \in R[x]$ such that $f(x) = q(x)(x-a) + r(x)$ and $\deg(r) < \deg(x-a)$. Then $r(x)$ is the constant polynomial $r(x) = f(a) = 0$ and so we have $f(x) = q(x)(x-a)$. Since $f(x) = (x-a)q(x)$ we have $(x-a)\big|f(x)$. Conversely, suppose that $(x-a)\big|f(a)$ and choose $p \in R[x]$ so that $f(x) = (x-a)p(x)$. Then $f(a) = (a-a)p(a) = 0 \cdot p(a) = 0$.

10.21 Definition: Let R be a commutative ring, let $f \in R[x]$, and let $a \in R$. We say that a is a **root** of f when $f(a) = 0$. When $f \neq 0$, we define the **multiplicity** of a as a root of f to be the largest $m = m(f,a) \in \mathbf{N}$ such that $(x-a)^m \big| f(x)$ (where we use the convention that $(x-a)^0 = 1$). Note that a is a root of f if and only if $m(f,a) \geq 1$.

10.22 Example: Let $f(x) = x^3 - 3x - 2 \in \mathbf{Q}[x]$. Since $f(x) = (x+1)^2(x-2) \in \mathbf{Q}[x]$, we have $m(f,2) = 1$ and $m(f,-1) = 2$.

10.23 Exercise: Let p be an odd prime and let $f(x) = x^p - a \in \mathbf{Z}_p[x]$. Find $m(f,a)$.

10.24 Theorem: *(The Roots Theorem) Let R be an integral domain, let $0 \neq f \in R[x]$ and let $n = \deg(f)$. Then*

(1) f has at most n distinct roots in R, and
(2) if $a_1, a_2, \cdots, a_\ell$ are all of the distinct roots of f in R and $m_i = m(f, a_i)$ for $1 \leq i \leq \ell$, then $(x - a_1)^{m_1}(x - a_2)^{m_2} \cdots (x - a_\ell)^{m_\ell} \big| f(x)$ and so $\sum_{i=1}^{\ell} m(f, a) \leq n$.

Proof: We prove Part (1) and leave the proof of Part (2) as an exercise. If $\deg(f) = 0$, then $f(x) = c$ for some $0 \in c \in R$, and so $f(x)$ has no roots. Let f be a polynomial with $\deg(f) = n \geq 1$ and suppose, inductively, that every polynomial $g \in R[x]$ with $\deg(g) = n - 1$ has at most $n - 1$ distinct roots. Suppose that a is a root of f in R. By the Factor Theorem, $(x - a) \big| f(x)$ so we can choose a polynomial $g \in R[x]$ so that $f(x) = (x - a)g(x)$. Note that $\deg(g) = n - 1$ so, by the induction hypothesis, g has at most $n - 1$ distinct roots. Let $b \in R$ be any root of f with $b \neq a$. Since $f(x) = (x - a)g(x)$ and $f(b) = 0$ we have $0 = f(b) = (b - a)g(b)$. Since $(b - a)g(b) = 0$ and $(b - a) \neq 0$ and R has no zero divisors, it follows that $g(b) = 0$. Thus b must be one of the roots of g. Since every root b of f with $b \neq a$ is equal to one of the roots of g, and since g has at most $n - 1$ distinct roots, it follows that f has at most n distinct roots, as required.

10.25 Note: Here are a few remarks about irreducible polynomials.

(1) When F is a field, we know that $F[x]$ is a unique factorization domain. For $f \in F[x]$ we know that $f = 0$ if and only if $\deg(f) = -1$, and f is a unit if and only if $\deg(f) = 0$, and for $0 \neq f, g \in F[x]$ we know that $\deg(fg) = \deg(f) + \deg(g)$. It follows that for $f \in F[x]$, if $\deg(f) = 1$ then f is irreducible. It also follows that for $f \in F[x]$, if $\deg(f) = 2$ or 3 then f is reducible in $F[x]$ if and only if f has a f has a root in F.

(2) When p is a fairly small prime number and n is a fairly small positive integer, it is easy to list all reducible and irreducible polynomials $f \in \mathbf{Z}_p[x]$ with $\deg(f) \leq n$. Note that it suffices to list monic polynomials (since for $f \in \mathbf{Z}_p[x]$ and $0 \neq c \in \mathbf{Z}_p[x]$ we have $f \sim cf$). We start by listing all monic polynomials of degree 1, that is all polynomials of the form $f(x) = x + a$ with $a \in \mathbf{Z}_p$, and noting that they are all irreducible. Having constructed all reducible and irreducible monic polynomials of all degrees less than n, we can construct all of the reducible monic polynomials of degree n by forming products of the reducible monic polynomials of smaller degree in all possible ways, and then all the remaining monic polynomials of degree n must be irreducible.

(3) For $f \in \mathbf{C}[x]$, we know that f is irreducible if and only if $\deg(f) = 1$. For $f \in \mathbf{R}[x]$, we know that f is irreducible polynomial if and only if either $\deg(f) = 1$ or $f(x) = ax^2 + bx + c$ for some $a, b, c \in \mathbf{R}$ with $a \neq 0$ and $b^2 - 4ac < 0$. For $R = \mathbf{Z}$ or $\mathbf{Q}$, it is a more challenging problem to determine which polynomials are irreducible in $R[x]$. The next few theorems are related to this problem.

10.26 Exercise: List all monic reducible and irreducible polynomials in $\mathbf{Z}_2[x]$ of degree less than 4, then determine the number of irreducible polynomials in $\mathbf{Z}_2[x]$ of degree 4.

10.27 Definition: Let $f \in \mathbf{Z}[x]$. The **content** of f, denoted by $c(f)$, is the greatest common divisor of the coefficients of f. We say that f is **primitive** when $c(f) = 1$.

10.28 Note: Let $a_0, a_1, \cdots, a_n \in \mathbf{Z}$, let $r \in \mathbf{Z}$ and let $d = \gcd(a_0, a_1, \cdots, a_n)$. Then $\gcd\left(ra_0, ra_1, \cdots, ra_n\right) = |r| \gcd(a_0, a_1, \cdots, a_n)$ and $\gcd\left(\frac{a_0}{d}, \frac{a_1}{d}, \cdots, \frac{a_n}{d}\right) = 1$. It follows that for $f(x) = \sum_{i=0}^{n} a_i x^i$ we have $c(rf) = |r|c(f)$ and that if we let $g(x) = \frac{1}{c(f)} f(x)$ then we have $g(x) \in \mathbf{Z}[x]$ and $c(g) = 1$.

10.29 Theorem: *(Gauss' Lemma)*
(1) For all $f, g \in \mathbf{Z}[x]$ we have $c(fg) = c(f)c(g)$.
(2) Let $0 \neq f \in \mathbf{Z}[x]$ and let $g(x) = \frac{1}{c(f)} f(x) \in \mathbf{Z}[x]$. Then f is irreducible in $\mathbf{Q}[x]$ if and only if g is irreducible in $\mathbf{Z}[x]$.

Proof: Let $f, g \in \mathbf{Z}[x]$. If $f = 0$ or $g = 0$ then we have $c(fg) = 0 = c(f)c(g)$. Suppose that $f \neq 0$ and $g \neq 0$. Let $h(x) = \frac{1}{c(f)} f(x)$ and $k(x) = \frac{1}{c(g)} g(x)$. Then we have $h, k \in \mathbf{Z}[x]$ with $c(h) = c(k) = 1$ and $fg = c(f)c(g)hk$ so that $c(fg) = c(f)c(g)c(hk)$. Thus to prove Part (1) it suffices to show that $c(hk) = 1$. Let $h(x) = \sum_{i-0}^{n} a_i x^i$ and $k(x) = \sum_{i=0}^{m} b_i x^i$ with $a_n \neq 0$ and $b_m \neq 0$. Suppose, for a contradiction, that $c(hk) \neq 1$. Let p be a prime factor of $c(hk)$. Then p divides all of the coefficients of $(hk)(x) = (a_0b_0)+(a_1b_0+a_0b_1)x+\cdots+(a_nb_m)x^{n+m}$. Since $c(h) = 1$, p does not divide all the coefficients of $h(x)$ so we can choose an index $r \geq 0$ so that $p|a_i$ for all $i < r$ and $p \not| a_r$. Since $c(k) = 1$ we can choose an index $s \geq 0$ so that $p|b_i$ for all $i < s$ and $p \not| b_s$. Since p divides every coefficient of $(hk)(x)$, it follows that in particular p divides the coefficient

$$c_{r+s} = a_0b_{r+s} + a_1b_{r+s-1} + \cdots + a_rb_s + \cdots + a_{r+s-1}b_1 + a_{r+s}.$$

Since $p|c_{r+s}$ and $p|a_i$ for all $i < r$ and $p|b_i$ for all $i < s$ it follows that $p|a_rb_s$. Since p is prime and $p|a_rb_s$ it follows that $p|a_r$ or $p|b_s$. But r and s were chosen so that $p \not| a_r$ and $p \not| b_s$ so we have obtained the desired contradiction. This proves Part (1).

To prove Part (2), let $0 \neq f(x) \in \mathbf{Z}[x]$ and let $g(x) = \frac{1}{c(f)} f(x)$, and note that $c(g) = 1$. Suppose that g is reducible in $\mathbf{Z}[x]$, say $g(x) = h(x)k(x)$ where $h(x)$ and $k(x)$ are non-units in $\mathbf{Z}[x]$. Since $c(h)c(k) = c(hk) = c(g) = 1$ it follows that $c(h) = c(k) = 1$. Note that $h(x)$ cannot be a constant polynomial since if we had $h(x) = r$ with $r \in \mathbf{Z}$, then we would have $|r| = c(h) = 1$ so that $h(x) = \pm 1$, but then h would be a unit. Similarly $k(x)$ cannot be a constant polynomial. Since $h(x)$ and $k(x)$ are nonconstant polynomials in $\mathbf{Z}[x]$, they are also nonconstant polynomials in $\mathbf{Q}[x]$. Since $f(x) = c(f)g(x) = c(f)h(x)k(x)$ and since $c(f)h(x)$ and $k(x)$ are both nonconstant polynomials (hence nonunits) in $\mathbf{Q}[x]$, it follows that $f(x)$ is reducible in $\mathbf{Q}[x]$.

Conversely, suppose that $f(x)$ is reducible in $\mathbf{Q}[x]$, say $f(x) = h(x)k(x)$ where h and k are nonzero, nonunits in $\mathbf{Q}[x]$. Since h and k are nonzero nonunits in $\mathbf{Q}[x]$, they are nonconstant polynomials. Let a be the least common multiple of the denominators of the coefficients of $h(x)$ and let b be the least common multiple of the coefficients of $k(x)$, and note that $ah(x) \in \mathbf{Z}[x]$ and $bk(x) \in \mathbf{Z}[x]$. Let $p(x) = \frac{1}{c(ah)} ah(x)$ and let $q(x) = \frac{1}{c(bk)} bk(x)$ and note that $p(x) \in \mathbf{Z}[x]$ and $q(x) \in \mathbf{Z}[x]$ with $c(p) = c(q) = 1$ and that $\deg(p) = \deg(h)$ and $\deg(q) = \deg(k)$. Since $f(x) = ah(x)\, bk(x) = c(ah)c(bk)p(x)q(x)$ we have $c(f) = c(ah)c(bk)c(pq) = c(ah)c(bk)$ and so $g(x) = \frac{1}{c(f)} f(x) = \frac{1}{c(ah)c(bk)} ah(x)\, bk(x) = p(x)q(x)$. Since $g(x) = p(x)q(x)$ where $p(x)$ and $q(x)$ are nonconstant polynomials in $\mathbf{Z}[x]$, we see that $g(x)$ is reducible in $\mathbf{Z}[x]$.

10.30 Corollary: *Let $0 \neq f(x) \in \mathbf{Z}[x]$. Then $f(x)$ is reducible in $\mathbf{Q}[x]$ if and only if $f(x)$ can be factored as a product of two nonconstant polynomials in $\mathbf{Z}[x]$.*

Proof: If $f(x)$ can be factored as $f(x) = g(x)h(x)$, where $g(x)$ and $h(x)$ are nonconstant polynomials in $\mathbf{Z}[x]$, then because $g(x)$ and $h(x)$ are also nonconstant polynomials in $\mathbf{Q}[x]$ (hence nonunits in $\mathbf{Q}[x]$), it follows immediately that f is reducible in $\mathbf{Q}[x]$. Suppose, conversely, that $f(x)$ is reducible in $\mathbf{Q}[x]$. Let $p(x) = \frac{1}{c(f)} f(x)$ and note that $p(x) \in \mathbf{Z}[x]$ with $c(p) = 1$. By Gauss' Lemma, $p(x)$ is reducible in $\mathbf{Z}[x]$. Choose nonunits $k, h \in \mathbf{Z}[x]$ such that $p = kh \in \mathbf{Z}[x]$. Since $c(k)c(h) = c(kh) = c(p) = 1$ we have $c(k) = c(h) = 1$. Since k and h are nonunits with $c(k) = c(h) = 1$, it follows that k and h are nonconstant polynomials (indeed, if $k(x)$ was constant with $k(x) = r$ then we would have $|r| = c(k) = 1$ so that $k(x) = r = \pm 1$, but then $k(x)$ would be a unit). Let $g(x) = c(f)k(x)$ and note that since $k(x)$ is nonconstant, so is $g(x)$. Then we have $f(x) = g(x)h(x)$, which is a product of two nonconstant polynomials in $\mathbf{Z}[x]$.

10.31 Example: Let $f(x) = 6x + 30 \in \mathbf{Z}[x]$. Note that $c(f) = 6$. Since $\deg(f) = 1$, it follows that f is irreducible in $\mathbf{Q}[x]$. But since $c(f) = 6$, it follows that f is reducible in $\mathbf{Z}[x]$, indeed in $\mathbf{Z}[x]$ we have $f(x) = 2 \cdot 3 \cdot (x + 5)$.

10.32 Theorem: *(Rational Roots) Let $f(x) = \sum_{i=0}^{n} c_i x^i$ where $n \in \mathbf{Z}^+$, each $c_i \in \mathbf{Z}$ and $c_n \neq 0$. Let $r, s \in \mathbf{Z}$ with $s \neq 0$ and $\gcd(r, s) = 1$. Then if $f\left(\frac{r}{s}\right) = 0$ then $r | c_0$ and $s | c_n$.*

Proof: Suppose that $f\left(\frac{r}{s}\right) = 0$, that is $c_0 + c_1 \frac{r}{s} + c_2 \frac{r^2}{s^2} + \cdots + c_n \frac{r^n}{s^n} = 0$. Multiply by s^n to get

$$0 = c_0 s^n + c_1 s^{n-1} r^1 + \cdots + c_{n-1} s^1 r^{n-1} + c_n r^n.$$

Thus we have

$$c_0 s^n = -r(c_1 s^{n-1} + \cdots + c_{n-1} s^1 r^{n-2} + c_n r^{n-1}) \text{ and}$$
$$c_n r^n = -s\big(c_0 s^{n-1} + c_1 s^{n-2} r^1 + \cdots + c_{n-1} r^{n-1}\big)$$

and it follows that $r \big| c_0 s^n$ and that $s \big| c_n r^n$. Since $\gcd(r, s) = 1$ we also have $\gcd(r, s^n) = 1$, and since $r \big| c_0 s^n$ it follows that $r | c_0$. Since $\gcd(s, r) = 1$ we also have $\gcd(s, r^n) = 1$, and since $s \big| c_n r^n$ it follows that $s | c_n$.

10.33 Exercise: Show that $\sqrt{1 + \sqrt{2}} \notin \mathbf{Q}$.

10.34 Theorem: *(Modular Reduction) Let $f(x) = \sum_{i=0}^{n} c_i x^i$ with $n \in \mathbf{Z}^+$, $c_i \in \mathbf{Z}$ and $c_n \neq 0$. Let p be a prime number with $p \not| c_n$. Let $\overline{f}(x) = \sum_{i=0}^{n} \overline{c_i}\, x^i \in \mathbf{Z}_p[x]$ where $\overline{c_i} = [c_i] \in \mathbf{Z}_p$. If $\overline{f}$ is irreducible in $\mathbf{Z}_p[x]$ then f is irreducible in $\mathbf{Q}[x]$.*

Proof: Suppose that $f(x)$ is reducible in $\mathbf{Q}[x]$. By the corollary to Gauss' Lemma, we can choose two nonconstant polynomials $g, h \in \mathbf{Z}[x]$ such that $f = gh \in \mathbf{Z}[x]$. Write $g(x) = \sum_{i=0}^{k} a_i x^k \in \mathbf{Z}[x]$ and $h(x) = \sum_{i=0}^{\ell} b_i x^i \in \mathbf{Z}[x]$ with $a_k \neq 0$, $b_\ell \neq 0$ and $k, \ell \geq 1$. Let $\overline{g} = \sum_{i=0}^{k} \overline{a}_i x^i \in \mathbf{Z}_p[x]$ and $\overline{h}(x) = \sum_{i=0}^{\ell} \overline{b}_i x^i \in \mathbf{Z}_p[x]$, and note that $\overline{f} = \overline{g}\,\overline{h} \in \mathbf{Z}_p[x]$. Since $c_n = a_k b_\ell$ and $p \not| c_n$ it follows that $p \not| a_k$ and $p \not| b_\ell$ in $\mathbf{Z}$ so $\overline{a}_k \neq 0$ and $\overline{b}_\ell \neq 0$ in $\mathbf{Z}_p$. Thus $\deg(\overline{g}) = \deg(g) = k$ and $\deg(\overline{h}) = \deg(h) = \ell$ so that $\overline{g}$ and $\overline{h}$ are nonconstant polynomials in $\mathbf{Z}_p[x]$, and so the polynomial $\overline{f} = \overline{g}\overline{h}$ is reducible in $\mathbf{Z}_p[x]$.

10.35 Exercise: Prove that $f(x) = x^5 + 2x + 4$ is irreducible in $\mathbf{Q}[x]$ by working in $\mathbf{Z}_3[x]$.

10.36 Theorem: *(Eisenstein's Criterion) Let* $f(x) = \sum_{i=0}^{n} c_i x^i$ *with* $n \in \mathbf{Z}^+$, $c_i \in \mathbf{Z}$ *and* $c_n \neq 0$. *Let* p *be a prime number such that* $p_i | c_i$ *for* $0 \leq i < n$ *and* $p \not| c_n$ *and* $p^2 \not| c_0$. *Then* f *is irreducible in* $\mathbf{Q}[x]$.

Proof: Suppose, for a contradiction, that $f(x)$ is reducible in $\mathbf{Q}[x]$. By the corollary to Gauss' Lemma, we can choose two nonconstant polynomials $g, h \in \mathbf{Z}[x]$ such that $f = gh \in \mathbf{Z}[x]$. Write $g(x) = \sum_{i=0}^{k} a_i x^k \in \mathbf{Z}[x]$ and $h(x) = \sum_{i=0}^{\ell} b_i x^i \in \mathbf{Z}[x]$ with $k, \ell \geq 1$ and $a_k \neq 0$, $b_\ell \neq 0$. Since $c_0 = a_0 b_0$ and $p | c_0$ but $p^2 \not| c_0$, it follows that p divides exactly one of the two numbers a_0 and b_0. Suppose that p divides a_0 but not b_0 (the case that p divides b_0 but not a_0 is similar). Since $p | c_1$, that is $p \big| (a_0 b_1 + a_1 b_0)$, and $p | a_0$ it follows that $p | a_1 b_0$, and since $p \not| b_0$ it follows that $p | a_1$. Since $p | c_2$, that is $p \big| (a_0 b_2 + a_1 b_1 + a_2 b_0)$ and $p | a_0$ and $p | a_1$, it follows that $p | a_2 b_0$, and since $p \not| b_0$ it then follows that $p | a_2$. Repeating this argument we find, inductively, that $p | a_i$ for all $i \geq 0$, and in particular we have $p | a_k$. Since $c_n = a_k b_\ell$ and $p | a_k$ it follows that $p | c_n$, giving the desired contradiction.

10.37 Example: Note that $f(x) = 5x^5 + 3x^4 - 18x^3 + 12x + 6$ is irreducible in $\mathbf{Q}[x]$ by Eisenstein's Criterion using $p = 3$.

10.38 Exercise: Let p be a prime number. Show that $f(x) = 1 + x + x^2 + \cdots + x^{p-1}$ is irreducible in $\mathbf{Q}[x]$,

.

Appendix 1: The Construction of the Real Numbers

Fist-Order Set Theory

1.1 Definition: In the language of **first-order set theory** we allow ourselves to use only symbols from the following **symbol set**

$$\Big\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, =, \in, \forall, \exists, (,)\Big\}$$

along with some variable symbols such as $x, y, z, u, v, w, \cdots$ or $x_1, x_2, x_3, \cdots$. The symbols in the symbol set are read as follows: $\neg$ is read as "not", $\wedge$ is read as "and", $\vee$ is read as "or", $\rightarrow$ is read as "implies", $\leftrightarrow$ is read as "if and only if", $\in$ is read as "is an element of", $\forall$" is read as "for all", $\exists$ is read as "there exists", and the symbols (and) are called parentheses.

1.2 Definition: A **formula** (in the formal symbolic language of **first-order set theory**) is a non-empty finite string of symbols, from the above list, which can be obtained using finitely many applications of the following three rules.

1. If x and y are variable symbols, then each of the following strings is a formula.

$$x = y \ , \quad x \in y$$

2. If F and G are formulas then each of the following strings is a formula.

$$\neg F \ , \quad (F \wedge G) \ , \quad (F \vee G) \ , \quad (F \rightarrow G) \ , \quad (F \leftrightarrow G)$$

3. If x is a variable symbol and F is a formula then each of the following is a formula.

$$\forall x \, F \ , \quad \exists x \, F$$

1.3 Definition: Let x be a variable symbol and let F be a formula. For each occurrence of the symbol x, which does not immediately follow a quantifier, in the formula F, we define whether the occurrence of x is **free** or **bound** inductively as follows.

1. If F is a formula of one of the forms $y = z$ or $y \in z$, where y and z are variable symbols (possibly equal to x), then every occurrence of x in F is free, and no occurrence is bound.

2. If F is a formula of one of the forms $\neg G$, $(G \wedge H)$, $(G \vee H)$, $(G \rightarrow H)$ or $(G \leftrightarrow H)$, where G and H are formulas, then each occurrence of the symbol x is either an occurrence in the formula G or an occurrence in the formula H, and each free (respectively, bound) occurrence of x in G remains free (respectivly, bound) in F, and similarly for each free (or bound) occurrence of x in H.

3. If F is a formula of one of the forms $\forall y \, G$ or $\exists y \, G$, where G is a formula and y is a variable symbol (possibly equal to x), then if y is different than x then each free (or bound) occurrence of x in G remains free (or bound) in the formula F, and if y is equal to x then every free occurrence of x in G becomes bound in the formula F, and every bound occurrence of x in G remains bound in the formula F.

1.4 Definition: When a quantifier symbol occurs in a given formula F, and is followed by the variable symbol x and then by the formula G, any free occurrence of x in G will become bound in the given formula F (by an application of part 3 of the above definition), and we shall say that that occurrence of x **is bound by** (that occurrence of) the quantifier symbol, or that (that occurrence of) the quantifier symbol **binds** that occurrence of x.

1.5 Definition: A **free variable** in a formula F is any variable symbol that has at least one free occurence in F. A formula F with no free variables is called a **statement**. When the free variables in F all lie in the set $\{x_1, x_2, \cdots, x_n\}$, we shall write F as $F(x_1, \cdots, x_n)$ and we shall say that F is a **statement about** the variables $x_1, x_2, \cdots, x_n$.

1.6 Example: In the following formula, determine which occurrences of the variable symbols are free and which are bound, and for each bound occurrence, indicate which quantifier binds it.

$$\forall x\, \exists y\big(\forall z(x \in y \to \exists y\ y = z) \wedge \forall x(\exists z\ z = u \vee z \in x)\big)$$

Solution: We indicate the free and bound occurrences and their binding quantifiers by placing integral labels under the relevant symbols: the free variables are given the label 0, each quantifier is given its own non-zero label, and each bound variable is given the same label as its binding quantifier:

$$\begin{array}{l}\forall x\, \exists y\big(\forall z(x \in y \to \exists y\ y = z) \wedge \forall x(\exists z\ z = u \vee z \in x)\big)\\ 1\ \ 2\ \ \ 3\ \ 1\ \ \ 2\ \ \ \ \ 4\ \ 4\ \ \ 3\ \ \ \ \ 5\ \ 6\ \ 6\ \ \ \ 0\ \ \ 0\ \ \ 5\end{array}$$

We remark that the free variables in this formula are z and u, so we say that it is a statement about z and u.

1.7 Example: Express the statement $x = \big\{y, \{z\}\big\}$ as a formal symbolic formula.

Solution: We can express the given statement in each of the following ways.

$$x = \big\{y, \{z\}\big\}$$
$$\forall u\, \big(u \in x \leftrightarrow u \in \big\{y, \{z\}\big\}\big)$$
$$\forall u\big(u \in x \leftrightarrow \big(u = y \vee u = \{z\}\big)\big)$$
$$\forall u\big(u \in x \leftrightarrow \big(u = y \vee \forall v(v \in u \leftrightarrow v = z)\big)\big)$$

The last expression is a formula.

1.8 Definition: Given a formula F and variable symbols x and y, we define $[F]_{x\mapsto y}$ as follows. When F is obtained using rule 1, the formula $[F]_{x\mapsto y}$ is obtained from F by replacing all occurrences of the symbol x by the symbol y. To deal with rule 2, we define $\big[\neg F\big]_{x\mapsto y} := \neg[F]_{x\mapsto y}$ and $\big[(F * G)\big]_{x\mapsto y} := \big([F]_{x\mapsto y} * [G]_{x\mapsto y}\big)$ for $* \in \{\wedge, \vee, \to, \leftrightarrow\}$. To deal with rule 3, we define $\big[\forall x\, F\big]_{x\mapsto y} := \forall x\, F$, and $\big[\forall y\, F\big]_{x\mapsto y} := \forall u\big[[F]_{y\mapsto u}\big]_{x\mapsto y}$ where u is the first variable which is not equal to x or y and which does not occur in F, and for a variable symbol z with $z \neq x$ and $z \neq y$ we define $\big[\forall z\, F\big]_{x\mapsto y} := \forall z[F]_{x\mapsto y}$.

1.9 Remark: Every mathematical **set** can be constructed using specific rules, which are known as the **ZFC axioms** of set theory, or the Zermelo-Fraenkel axioms of set theory, with the axiom of choice. We begin by listing the ZFC axioms, stating them informally.

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

Empty Set Axiom: There exists a set $\emptyset$ with no elements.

Separation Axiom: If u is a set and $F(x)$ is a statement about x, $\{x \in u | F(x)\}$ is a set.

Pair Axiom: If u and v are sets then $\{u, v\}$ is a set.

Union Axiom: If u is a set then $\bigcup u = \bigcup_{v \in u} v$ is a set.

Power Set Axiom: If u is a set then $\mathcal{P}(u) = \{v | v \subseteq u\}$ is a set.

Axiom of Infinity: If we define the natural numbers to be the sets $0 = \emptyset$, $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$ and so on, then $\mathbf{N} = \{0, 1, 2, 3, \cdots\}$ is a set.

Replacement Axiom: If u is a set and $F(x, y)$ is a statement about x and y with the property that $\forall x \, \exists! y \; F(x, y)$ then $\{y \mid \exists x \in u \; F(x, y)\}$ is a set.

Axiom of Choice: Given a set u of non-empty pairwise disjoint sets, there exists a set which contains exactly one element from each of the sets in u.

We now proceed to state each of the ZFC axioms formally (as a formula in first-order set theory) and give some indication as to how these axioms can be used as a rigorous framework for essentially all of mathematics.

1.10 Definition: The **Extension Axiom** is the formula

$$\forall u \, \forall v \left(u = v \leftrightarrow \forall x (x \in u \leftrightarrow x \in v) \right).$$

1.11 Definition: The **Empty Set Axiom** is the formula

$$\exists u \, \forall x \; \neg x \in u \, .$$

1.12 Theorem: *The empty set is unique.*

Proof: Suppose that u and v are both empty. Let x be arbitrary. Since u is empty, we have $\neg \, x \in u$ and hence $x \in u \rightarrow x \in v$. Similarly, since v is empty, we have $\neg \, x \in v$ and hence $x \in v \rightarrow x \in u$. Since $x \in u \rightarrow x \in v$ and $x \in v \rightarrow x \in u$, we have $x \in u \leftrightarrow x \in v$. Since x was arbitrary, we have $\forall x \, (x \in u \leftrightarrow x \in v)$. By the Axiom of Extension, $u = v$.

1.13 Definition: We denote the unique empty set by $\emptyset$.

1.14 Remark: In a formal and rigorous treatment of the foundations of mathematics, we would need to decide at this point how to interpret the use of the symbol $\emptyset$. One approach is to add the symbol $\emptyset$ to our list of symbols, modify our definition of a formula to allow the use of the new symbol $\emptyset$, and add the axiom $\forall x \; \neg x \in \emptyset$ to our list of axioms. Another option is to interpret the use of the symbol as a shorthand notation for an expression which can be expressed formally using the existing symbols, so that for example the expression $u = \emptyset$ would be shorthand for the formula $\forall x \; \neg x \in u$.

1.15 Definition: Given sets u and v, we say that u is a **subset** of v, and we write $u \subseteq v$, when every element of u also lies in v, that is when $\forall x\,(x \in u \to x \in v)$.

1.16 Definition: For any formula F with free variable x, the following formula is an axiom.

$$\forall u\,\exists v\,\forall x\big(x \in v \leftrightarrow (x \in u \wedge F)\big)$$

More generally, for any formula $F$with free variables $x, u_1, u_2, \cdots, u_n$, the following formula is an axiom.

$$\forall u \forall u_1 \cdots \forall u_n \exists v\,\forall x\big(x \in v \leftrightarrow \big(x \in u \wedge F\big)\big)$$

Any axiom of this form is called an **Axiom of Separation**.

1.17 Notation: Given sets $u, u_1, \cdots, u_n$ and given a formula F with free variables $x, u_1, \cdots, u_n$, by the appropriate Axiom of Separation, there exists a set v with the property that $\forall x\big(x \in v \leftrightarrow \big(x \in u \wedge F\big)$, and by the Extension Axiom, this set v is unique, and we denote it by

$$\big\{x \in u \,\big|\, F\big\}\,.$$

1.18 Note: It is important to realize that a Separation Axiom only allows us to construct a subset of a given set u, so for example we cannot use a Separation Axiom to show that the collection $S = \{x \big| \neg\, x \in x\}$, which is used to formulate Russel's paradox, is a set.

1.19 Definition: The **Pair Axiom** is the formula

$$\forall u\,\forall v\,\exists w\,\forall x\big(x \in w \leftrightarrow (x = u \vee x = v)\big)\,.$$

1.20 Notation: Given sets u and v, by the Pair Axiom there exists a set w with the property that $\forall x\big(x \in w \leftrightarrow (x = u \vee x = v)\big)$, and by the Extension Axiom, this set w is unique, and we denote it by

$$\{u, v\}$$

1.21 Example: With this axiom, we can construct some non-empty sets. For example, taking $u = v = \emptyset$ gives the set $\{\emptyset, \emptyset\} = \{\emptyset\}$ (note that $\{\emptyset\} \neq \emptyset$ by the Extension Axiom, since $\emptyset \in \{\emptyset\}$ but $\emptyset \notin \emptyset$). Then taking $u = \emptyset$ and $v = \{\emptyset\}$ gives the set $\big\{\emptyset, \{\emptyset\}\big\}$.

1.22 Definition: The **Union Axiom** is the formula

$$\forall u\,\exists w\,\forall x\big(x \in w \leftrightarrow \exists v\,(v \in u \wedge x \in v)\big)\,.$$

1.23 Definition: Given a set u, by the Union Axiom there exists a set w with the property that $\forall x\big(x \in w \leftrightarrow \exists v\,(v \in u \wedge x \in v)\big)$, and by the Extension Axiom this set w is unique. We call the set w the **union** of the elements in u, and we denote it by

$$\bigcup u = \bigcup_{v \in u} v\,.$$

Given two sets u and v, we define the **union** of u and v to be the set

$$u \cup v = \bigcup \{u, v\}\,.$$

Given three sets u, v and w, note that $\{z\} = \{z, z\}$ is a set and so $\{x, y, z\} = \{x, y\} \cup \{z\}$ is also a set. More generally, if $u_1, u_2, \cdots, u_n$ are sets then $\{u_1, u_2, \cdots, u_n\}$ is a set and we define the **union** of the sets $u_1, \cdots, u_n$ to be

$$u_1 \cup u_2 \cup \cdots \cup u_n = \bigcup_{k=1}^{n} u_k = \bigcup \{u_1, u_2, \cdots, u_n\}\,.$$

1.24 Definition: Given a non-empty set u, we define the **intersection** of the elements in u to be the set

$$\bigcap u = \left\{ x \in \bigcup u \,\middle|\, \forall v (v \in u \to x \in v) \right\}$$

Given two sets u and v, we define the **intersection** of u and v to be the set

$$u \cap v = \bigcap \{u, v\} ,$$

and more generally, given sets $u_1, u_2, \cdots, u_n$, we define the **intersection** of $u_1, u_2, \cdots, u_n$ to be the set

$$u_1 \cap u_2 \cap \cdots \cap u_n = \bigcap_{k=1}^{n} u_k = \bigcap \{u_1, u_2, \cdots, u_n\} .$$

1.25 Definition: The **Power Set Axiom** is the formula

$$\forall u \, \exists w \, \forall v \, (v \in w \leftrightarrow v \subseteq u) .$$

1.26 Definition: Given a set u, the set w with the property that $\forall v (v \in w \leftrightarrow v \subseteq u)$ (which exists by the Power Set Axiom and is unique by the Extension Axiom) is called the **power set** of u and is denoted by $\mathcal{P}(u)$, so we have

$$\mathcal{P}(u) = \{ v | v \subseteq u \} .$$

1.27 Example: Find the power set of the set $\{\emptyset, \{\emptyset\}\}$.

Solution: We have

$$\mathcal{P}\Big(\{\emptyset, \{\emptyset\}\}\Big) = \Big\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\Big\} .$$

1.28 Definition: Given two sets x and y, we define the **ordered pair** (x, y) to be the set

$$(x, y) = \big\{\{x\}, \{x, y\}\big\} .$$

Given two sets u and v, note that if $x \in u$ and $y \in v$ then we have $\{x\} \in \mathcal{P}(u \cup v)$ and $\{x, y\} \in \mathcal{P}(u \cup v)$ and so $(x, y) = \big\{\{x\}, \{x, y\}\big\} \in \mathcal{P}\big(\mathcal{P}(u \cup v)\big)$. We define the **product** $u \times v$ to be the set

$$u \times v = \big\{(x, y) \big| x \in u \wedge y \in v\big\} ,$$

that is

$$u \times v = \big\{ z \in \mathcal{P}\big(\mathcal{P}(u \cup v)\big) \,\big|\, \exists x \, \exists y \big((x \in u \wedge y \in v) \wedge z = (x, y)\big) \big\} .$$

1.29 Exercise: Find $\bigcup \Big(\{\emptyset\} \times \big\{\{\emptyset\}, \{\emptyset, \{\emptyset\}\}\big\} \Big)$.

1.30 Definition: We define

$$0 = \emptyset \ , \ 1 = \{0\} = 0 \cup \{0\} \ , \ 2 = \{0, 1\} = 1 \cup \{1\} \ , \ 3 = \{0, 1, 2\} = 2 \cup \{2\} ,$$

and so on. For a set x, we define the **successor** of x to be the set

$$x + 1 = x \cup \{x\} .$$

A set u is called **inductive** when it has the property that

$$\big(0 \in u \wedge \forall x \, (x \in u \to x + 1 \in u)\big) .$$

1.31 Definition: The **Axiom of Infinity** is the formula

$$\exists u \big(0 \in u \wedge \forall x \, (x \in u \to x + 1 \in u)\big) ,$$

so the Axiom of Infinity states that there exists an inductive set.

1.32 Theorem: *There exists a unique set w of the form*

$$w = \big\{x \,\big|\, x \in v \text{ for every inductive set } v\big\}.$$

Moreover, this set w is an inductive set.

Proof: By the axiom of infinity, there exists an inductive set, say u. Let w be the set

$$\begin{aligned} w &= \big\{x \in u \,\big|\, x \in v \text{ for every inductive set } v\big\} \\ &= \big\{x \in u \,\big|\, \forall v \left(\big(0 \in v \wedge \forall y \,(y \in v \to y+1 \in v)\big) \to x \in v\right)\big\}. \end{aligned}$$

We claim that this set w does not depend on the choice of u. To prove this, let u_1 and u_2 be two inductive sets and let

$$\begin{aligned} w_1 &= \big\{x \in u_1 \,\big|\, x \in v \text{ for every inductive set } v\big\} \\ w_2 &= \big\{x \in u_2 \,\big|\, x \in v \text{ for every inductive set } v\big\}. \end{aligned}$$

Then for any set x we have

$$\begin{aligned} x \in w_1 &\iff x \in u_1 \text{ and } x \in v \text{ for every inductive set } v \\ &\iff x \in v \text{ for every inductive set } v \text{ (since } u_1 \text{ is inductive)} \\ &\iff x \in u_2 \text{ and } x \in v \text{ for every inductive set } v \text{ (since } u_2 \text{ is inductive)} \\ &\iff x \in w_2\,. \end{aligned}$$

Thus $w_1 = w_2$, showing that w is unique. We leave it as an exercise to show that w is inductive.

1.33 Definition: The unique set w in the above theorem is called the set of **natural numbers**, and we denote it by $\mathbf{N}$. We write

$$\begin{aligned} \mathbf{N} &= \big\{x \,\big|\, x \in v \text{ for every inductive set } v\big\} \\ &= \{0, 1, 2, 3 \cdots\}. \end{aligned}$$

For $x, y \in \mathbf{N}$, we write $x < y$ when $x \in y$ and we write $x \le y$ when $x < y$ or $x = y$.

1.34 Notation: For a formula F, we write $\forall x \in u\ F$ as a shorthand notation for the formula $\forall x\,(x \in u \to F)$. Similarly, we write $\exists x \in u\ F$ as a shorthand notation for $\exists x\,(x \in u \wedge F)$.

1.35 Theorem: *(Principle of Induction) Let $F(x)$ be a formula with free variable x. Suppose that*

(1) $F(0)$, and
(2) $\forall x \in \mathbf{N}\,\big(F(x) \to F(x+1)\big)$.

Then $\forall x \in \mathbf{N}\ F(x)$.

Proof: Let $u = \big\{x \in \mathbf{N} \big| F(x)\big\}$. By (1) we have $0 \in u$. Let $x \in u$. Then $x \in \mathbf{N}$ and $F(x)$. Since $x \in \mathbf{N}$ we have $x + 1 \in \mathbf{N}$ (since $\mathbf{N}$ is inductive). Since $x \in \mathbf{N}$ and $F(x)$ we have $F(x+1)$ by (2). Since $x+1 \in \mathbf{N}$ and $F(x+1)$, we have $x+1 \in u$ (by the definition of u). We have shown that $0 \in u$ and that $\forall x\,(x \in u \to x+1 \in u)$, so u is inductive. Since u is inductive, we have $\mathbf{N} \subseteq u$ (by the definition of $\mathbf{N}$). Thus $x \in \mathbf{N} \implies x \in u \implies F(x)$.

1.36 Remark: In the above theorem, the expression $F(0)$ is short for $\forall x\,\big(x = 0 \to F(x)\big)$ which in turn is short for $\forall x\,\big(\forall y \neg\, y \in x \to F(x)\big)$. Similarly, $F(x+1)$ is short for the formula $\forall y\,\big(y = x{+}1 \to F(y)\big)$, where $F(y) = \big[F(x)\big]_{x \mapsto y}$.

1.37 Definition: Given a formula $F(x,y)$ with free variables x and y, the following formula is an axiom:

$$\forall u\Big(\forall x\,\exists!y\ F(x,y) \to \exists w\,\forall y\,\big(y\in w \leftrightarrow \exists x\in u\ F(x,y)\big)\Big)\,,$$

where $\exists!yF(x,y)$ is short for $\exists y\big(F(x,y)\wedge\forall z\,\big(F(x,z)\to z=y\big)\big)$ with $F(x,z)$ short for the formula $\forall y\,\big(y=z\to F(x,y)\big)$. More generally, given a formula $F(x,y,u_1,\cdots,u_n)$ with free variables $x,y,u_1,\cdots,u_n$, the following formula is an axiom:

$$\forall u\,\forall u_1\cdots\forall u_n\Big(\forall x\,\exists!y\ F(x,y,u_1,\cdots,u_n)\to\exists w\,\forall y\,\big(y\in w\leftrightarrow\exists x\in u\ F(x,y,u_1,\cdots,u_n)\big)\Big)\,.$$

An axiom of this form is called a **Replacement Axiom**.

1.38 Notation: Given sets $u,u_1,\cdots,u_n$ and given a formula $F(x,y,u_1,\cdots,u_n)$ with free variables $x,y,u_1,\cdots,u_n$ with the property that $\forall x\,\exists!y\ F(x,y,u_1,\cdots,u_n)$, for each set x we let $y=f(x)$ denote the unique set for which $F(x,y,u_1,\cdots,u_n)$ holds, and then we denote the unique set w, whose existence is stipulated by the above Replacement Axiom, by

$$\big\{f(x)\big|x\in u\big\}\,.$$

1.39 Example: If u is a set then the collection

$$\big\{\mathcal{P}(x)\big|x\in u\big\}$$

is also a set, by the Replacement Axiom taking $F(x,y)$ to be the formula $y=\mathcal{P}(x)$.

1.40 Definition: The **Axiom of Choice** is the formula given by

$$\forall u\Big(\big(\neg\phi\in u\wedge\ \forall x\in u\ \forall y\in u\,(\neg x=y\to x\cap y=\emptyset)\big)\to\exists w\,\forall v\in u\ \exists!\,x\in v\ x\in w\Big)$$

1.41 Remark: We have now stated each of the ZFC axioms formally. Up until now, we have used lower-case letters to denote all sets (and all elements of sets, which are also sets). From now on, we shall often use upper-case letters to denote sets, as is more customary.

1.42 Definition: A **binary relation** R on a set X is a subset $R \subseteq X \times X$. More generally, a **binary relation** is any set R whose elements are ordered pais. For a binary relation R, we usually write xRy instead of $(x, y) \in R$.

1.43 Definition: Let R and S be binary relations. The **domain** of R is

$$\text{Domain}(R) = \left\{x \,\middle|\, \exists y \; xRy\right\}$$

and the **range** of R is

$$\text{Range}(R) = \left\{x \,\middle|\, \exists y \; xRy\right\}.$$

For any set A, the **image** of A under R is

$$R(A) = \left\{y \,|\, \exists x \in A \; xRy\right\}$$

and the **inverse image** of A under R is

$$R^{-1}(A) = \left\{x \,\middle|\, \exists y \in A \; xRy\right\}.$$

The **inverse** of R is

$$R^{-1} = \left\{(y, x) \middle| (x, y) \in R\right\}$$

and the composite S **composed with** R is

$$S \circ R = \left\{(x, z) \middle| \exists y \; xRy \wedge ySz\right\}.$$

1.44 Theorem: *Let A be a set and let R and S be binary relations. Then*

(1) $\text{Domain}(R)$, $\text{Range}(R)$, $R(A)$ *and* $R^{-1}(A)$ *are sets, and*
(2) R^{-1} *and* $S \circ R$ *are binary relations.*

Proof: The proof is left as an exercise.

1.45 Definition: An **equivalence relation** on a set X is a binary relation R on X such that

(1) R is **reflexive**, that is $\forall x \in X \; xRx$,
(2) R is **symmetric**, that is $\forall x, y \in X \; (xRy \to yRx)$, and
(3) R is **transitive**, that is $\forall x, y, z \in X \; \big((xRy \wedge yRz) \to xRz\big)$.

1.46 Definition: Let R be an equivalence relation on the set X. For $a \in X$, the **equivalence class** of a modulo R is the set

$$[a]_R = \left\{x \in X \middle| xRa\right\}.$$

1.47 Definition: A **partition** of a set X is a set S of non-empty pairwise disjoint sets whose union is X, that is a set S such that

(1) for all $A \in S$ we have $A \neq \emptyset$,
(2) for all $A, B \in S$, if $A \neq B$ then $A \cap B = \emptyset$, and
(2) $\bigcup S = X$.

1.48 Theorem: *Given a set X, we have the following correspondence between equivalence relations on X and partitions of X.*

(1) Given an equivalence relation R on X, the set of all equivalence classes

$$S_R = \left\{ [a]_R \,\middle|\, a \in X \right\}$$

is a partition of X.
(2) Given a partition S of X, the relation R_S on X defined by

$$R_S = \left\{ (x,y) \in X \times X \middle| \exists A \in S \, (x \in A \wedge y \in A) \right\}$$

is an equivalence relation on X.
(3) Given an equivalence relation R on X we have $R_{S_R} = R$, and given a partition S of X we have $S_{R_S} = S$.

Proof: The proof is left as an exercise.

1.49 Notation: Given an equivalence relation R on X, the set of all equivalence classes, which we denoted by S_R in the above theorem, is usually denoted by X/R, so

$$X/R = \left\{ [a]_R \,\middle|\, a \in X \right\}.$$

1.50 Definition: Let R be an equivalence relation. A **set of representatives** for R is a subset of X which contains exactly one element from each equivalence class in X/R.

1.51 Remark: The Axiom of Choice is equivalent to the statement that every equivalence relation has a set of representatives.

1.52 Definition: Given sets X and Y, a **function** from X to Y is a binary relation $f \subseteq X \times Y$ with the property that

$$\forall x \in X \; \exists ! \, y \in Y \; (x,y) \in f \, .$$

More generally, a **function** is a binary relation with the property that

$$\forall x \in \text{Domain}(f) \; \exists ! \, y \; (x,y) \in f \, .$$

For a function f, we usually write $y = f(x)$ instead of xfy. It is customary to use the notation $f : X \to Y$ when $X = \text{Domain}(f)$ and Y is any set with Range(f) $\subseteq Y$.

1.53 Definition: Let $f : X \to Y$. The function f is called **one-to-one** (or **injective**) when

$$\forall y \in Y \; \exists \text{ at most one } x \in X \; y = f(x)$$

and f is called **onto** (or **surjective**) when

$$\forall y \in Y \; \exists \text{ at least one } x \in X \; y = f(x) \, .$$

1.54 Definition: Let $f : X \to Y$. Let I_X and I_Y denote the identity functions on X and Y respectively (that is $I_X(x) = x$ for all $x \in X$ and $I_Y(y) = y$ for all $y \in Y$). A **left inverse** of f is a function $g : Y \to X$ such that $g \circ f = I_X$. A **right inverse** of f is a function $h : Y \to X$ such that $f \circ h = I_Y$. Note that if f has a left inverse g and a right inverse h, then we have $g = g \circ I_Y = g \circ f \circ h = I_X \circ h = h$. In this case we say that g is the (unique two-sided) **inverse** of f.

1.55 Theorem: *Let $f : X \to Y$. Then*

(1) f is one-to-one if and only if f has a left inverse.
(2) f is onto if and only if f has a right inverse.
(3) f is one-to-one and onto if and only if f has a (two-sided) inverse.

Proof: The proof is left as an exercise. We remark that the Axiom of Choice is needed.

1.56 Definition: A function $f : X \to Y$ is called **invertible** (or **bijective**) when it is one-to-one and onto, or equivalently, when it has a (unique two-sided) inverse.

1.57 Remark: The Axiom of Choice is equivalent to the statement that for every set S, there exists a function $f : S \to \bigcup S$ with the property that $\forall X \in S \big(X \neq \emptyset \to f(X) \in X\big)$. Such a function f is called a **choice function** for the set S.

1.58 Theorem: *(The Recursion Theorem)*

(1) Let A be a set, let $a \in A$, and let $g : A \times \mathbf{N} \to A$. Then there exists a unique function $f : \mathbf{N} \to A$ such that

$$f(0) = a \text{ and } f(n+1) = g\big(f(n), n\big) \text{ for all } n \in \mathbf{N}\,.$$

(2) Let A and B be sets, let $g : A \to B$, and let $h : A \times B \times \mathbf{N} \to B$. Then there exists a unique function $f : A \times \mathbf{N} \to B$ such that for all $a \in A$ we have

$$f(a, 0) = g(a) \text{ and } f(a, n+1) = h\big(a, f(a, n), n\big) \text{ for all } n \in \mathbf{N}\,.$$

Proof: To prove part (1), note first that for each $n \in \mathbf{N}$ we can construct a (unique) function $f_n : \{0, 1, \cdots, n\} \to A$ such that $f(0) = a$ and $f_n(k+1) = g\big(f_n(k), k\big)$ for all k with $0 \leq k < n$ (that the functions f_n exist and are unique can be proven by induction). Notice that since $\{0, 1, \cdots, n\} = n{+}1$, we have $f_n : (n{+}1) \to A$, so $f_n \subseteq (n{+}1){\times}A \subseteq \mathbf{N}{\times}A$, and so all of the functions f_n are subsets of $\mathbf{N} \times A$. We can combine all these functions into a single function $f : \mathbf{N} \to A$ as follows. First we let

$$F = \Big\{ f \subseteq \mathbf{N}{\times}A \;\Big|\; \exists n{\in}\mathbf{N}\Big(f : (n{+}1) \to A\,,\ f(0) = a\,,\ \forall k{\in}(n{+}1)\ f(k{+}1) = g\big(f(k), k\big)\Big)\Big\}\,,$$

and then we let

$$f = \bigcup F\,.$$

We leave it as an exercise to prove that indeed f is a function which satisfies the conditions of the theorem.

We can prove part (2) in a similar manner. First we let

$$F = \Big\{ f \subseteq A \times \mathbf{N} \times B \;\Big|\; \exists n{\in}\mathbf{N}\Big(f : A \times (n+1) \to B \text{ and } \forall a{\in}A\Big(f(a, 0) = g(a) \,\wedge\, \forall k{\in}(n+1)\ f(a, k+1) = h\big(a, f(a, k), k\big)\Big)\Big)\Big\}\,,$$

then we let $f = \bigcup F$.

1.59 Definition: By part (2) of the Recursion Theorem, there is a unique function $s : \mathbf{N} \times \mathbf{N} \to \mathbf{N}$ such that for all $a, b \in \mathbf{N}$ we have

$$s(a, 0) = a \ , \quad s(a, b + 1) = s(a, b) + 1 \, .$$

We call $s(a, b)$ the **sum** of a and b in $\mathbf{N}$, and we write it as

$$a + b = s(a, b) \, .$$

Also, there is a unique function $p : \mathbf{N} \times \mathbf{N} \to \mathbf{N}$ such that for all $a, b \in \mathbf{N}$ we have

$$p(a, 0) = 0 \ , \quad p(a, b + 1) = p(a, b) + a \, .$$

We call $p(a, b)$ the **product** of a and b in $\mathbf{N}$, and we write it as

$$a \cdot b = p(a, b) \, .$$

1.60 Remark: It can be shown (using induction) that the sum and product satisfy all the usual properties in $\mathbf{N}$.

1.61 Definition: We define the set of **integers** to be the set

$$\mathbf{Z} = \big(\mathbf{N} \times \mathbf{N}\big)/R$$

where R is the equivalence relation given by

$$(a, b)R(c, d) \iff a + d = b + c \, .$$

For $a, b, c, d \in \mathbf{N}$, we define

$$[(a, b)] \leq [(c, d)] \iff b + c \leq a + d$$
$$[(a, b)] + [(c, d)] = [(a + c, b + d)]$$
$$[(a, b)] \cdot [(c, d)] = [(ac + bd, ad + bc)] \, .$$

For $n \in \mathbf{N}$, we write $n = [(n, 0)]$ and $-n = [(0, n)]$, so that every element of $\mathbf{Z}$ can be written as $\pm n$ for some $n \in \mathbf{N}$, and we can identify $\mathbf{N}$ with a subset of $\mathbf{Z}$.

1.62 Remark: It can be shown that the ordering and the sum and product defined above are well-defined and satisfy the usual properties in $\mathbf{Z}$.

1.63 Definition: We define the set of **rational numbers** to be the set

$$\mathbf{Q} = \big(\mathbf{Z} \times \mathbf{Z}^+\big)/R$$

where $\mathbf{Z}^+ = \{x \in \mathbf{Z} \big| x > 0\}$ and R is the equivalence relation given by

$$(a, b)R(c, d) \iff ad = bc \, .$$

For $a, b, c, d \in \mathbf{Z}$ with $b, d > 0$, we define

$$[(a, b)] \leq [(c, d)] \iff a \cdot d \leq b \cdot c$$
$$[(a, b)] + [(c, d)] = [(a \cdot d + b \cdot c, b \cdot d)]$$
$$[(a, b)] \cdot [(c, d)] = [(a \cdot c, b \cdot d)] \, .$$

For $a \in \mathbf{Z}$ and $b \in \mathbf{Z}^+$, it is customary to write $\frac{a}{b} = [(a, b)]$. Also for $a \in \mathbf{Z}$ we write $a = [(a, 1)]$, and we identify $\mathbf{Z}$ with a subset of $\mathbf{Q}$.

1.64 Remark: It can be shown that the above ordering, sum and product are well-defined and satisfy the usual rules in $\mathbf{Q}$.

1.65 Definition: We define the set of **real numbers** to be the set

$$\mathbf{R} = \left\{ u \subseteq \mathbf{Q} \,\middle|\, u \neq \emptyset \ , \ u \neq \mathbf{Q} \ , \ \forall a \in u \ \forall x \in \mathbf{Q} \ (x \leq a \to x \in u) \ , \ \forall a \in u \ \exists b \in u \ a < b \right\}.$$

For $u, v \in \mathbf{R}$, we define

$$u \leq v \iff u \subseteq v$$
$$u + v = \left\{ a + b \,\middle|\, a \in u \ , \ b \in v \right\}.$$

We define $0 \in \mathbf{R}$ to be the set $0 = \{x \in \mathbf{Q} \,|\, x < 0 \text{ in } \mathbf{Q}\}$. Given $u \in \mathbf{R}$ we define $-u$ to be the interior of the complement of $\{-a \,|\, a \in u\}$, that is

$$-u = \left\{ b \in \mathbf{Q} \,\middle|\, \exists \ r \in \mathbf{Q} \text{ with } r > 0 \text{ such that } \ -(b + r) \notin u \right\}$$

and we define

$$|u| = \begin{cases} u & \text{if } 0 \leq u, \\ -u & \text{if } u \leq 0. \end{cases}$$

For $0 \leq u, v \in \mathbf{R}$ we define

$$u \cdot v = \left\{ a \cdot b \in \mathbf{Q} \,\middle|\, 0 \leq a \in u \ , \ 0 \leq b \in v \right\} \cup \left\{ c \in \mathbf{Q} \middle| c < 0 \right\},$$

and for any $u, v \in \mathbf{R}$ we define $u \cdot v = |u| \cdot |v|$. For $a \in \mathbf{C}$ we write $a = \left\{ x \in \mathbf{Q} \middle| x < a \right\} \in \mathbf{R}$ and we identify $\mathbf{Q}$ with a subset of $\mathbf{R}$.

1.66 Remark: It can be shown that the above ordering, sum and product are well-defined and satisfy the usual rules in $\mathbf{R}$.

1.67 Definition: We define the set of **complex** numbers to be the set

$$\mathbf{C} = \mathbf{R} \times \mathbf{R}.$$

We define addition and multiplication in $\mathbf{C}$ by

$$(a, b) + (c, d) = (a + c, b + d)$$
$$(a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

We write $i = (0, 1)$. For $x \in \mathbf{R}$ we write $x = (x, 0)$, and we identify $\mathbf{R}$ with a subset of $\mathbf{C}$.

1.68 Remark: It can be shown that the above sum and product are well-defined and satisfy the usual rules in $\mathbf{C}$.