FOUNDATIONS & PROOF LECTURE NOTES

by Dr Lynne Walling

Note: You are expected to spend 3-4 hours per week working on this course outside of the lectures and tutorials. In this time you are expected to review the lecture notes, the comments on your homework, and the model solutions; work on your current homework assignment; neatly rewrite your homework solutions for submission to your tutor.

In these notes, many proofs refer to previously proved results or previously stated assumptions by restating the results or assumptions; this is how I expect you to refer to these things when you take the exam in this course. However, you may find it useful in studying to annotate these notes with the propostion/theorem/corollary number or the page number containing the result or assumption being invoked.

References for the course:

- These notes: *Transcription of Lynne Walling's Lectures on Foundations & Proof.*
- Larry Gerstein, *Discrete Mathematics and Algebraic Structures*, W.H. Freeman and Company, 1987.
- D.J. Velleman, *How to Prove It: A Structured Approach*, Cambridge University Press, 2006.
- P.J. Eccles, *An Introduction to Mathematical Reasoning: Numbers, Sets and Functions*, Cambridge University Press, 1997.

The course is organised into the following sections.

§1. Introduction: Sets and Functions (including notation; discussion of sets; Cartesian products; definition of a function; injective, surjective, and bijective functions; composition of functions; invertible functions; proof by contradiction)

§2. Truth tables, equivalences, and contrapositive (including notation used in truth tables; equivalence of propositions; the contrapositive of a proposition)

§3. Negations and contrapositives of propositions with quantifiers (including notation for a proposition dependent on a variable; an algorithmic approach for negating complex propositions; an equivalent definition of injective)

§4. Set operations (including union, intersection, difference of two sets, complement of a set; De Morgan's Laws and similar propositions involving unions, intersections, differences, and complements of sets; indexed sets; inverse image of a set under a function; relations between inverse images)

§5. Partitioning sets, equivalence relations, and congruences (including relations on a set; definitions of reflexive, symmetric, and transitive relations; a correspondence between a partition of a set and an equivalence relation on that set; congruences)

§6. Algorithms, recursion, and mathematical induction (including the division algorithm in the integers; highest common factors; Euclid's algorithm; the Chinese Remainder Theorem; using mathematical induction to prove relations on sets constructed using set operations)

§7. Strong induction and the Fundamental Theorem of Arithmetic (including the definition of a prime number; a proof that there are infinitely many prime numbers; an application of the Fundamental Theorem of Arithmetic to find all prime numbers p so that $5p + 9$ is the square of an integer)

§8. Cardinality (including the definition of a countable set; statement of the Cantor-Schröder-Bernstein Theorem; basic results regarding the cardinality of subsets of the positive integers; proof that the Cartesian product $\mathbb{Z}_+ \times \mathbb{Z}_+$ is countable; proof that the union of a countable number of pairwise disjoint countable sets is a countable set)

§9. Uncountable sets and power sets (including Cantor's diagonalisation proof that the unit interval $(0, 1)$ is uncountable; proof of Cantor's Theorem that the cardinality of the power set of a set A is strictly larger than the cardinality of A)

§10. More proofs using contradiction, construction, and induction (including more practice problems; how to easily determine whether an integer is divisible by 9)

1. Introduction: Sets and Functions

Mathematics is pure language - the language of science. It is unique among languages in its ability to provide precise expression for every thought or concept that can be formulated in its terms... It is also an art - the most intellectual and classical of the arts. (Quote from A. Adler's article "Mathematics and Creativity" in *The World Treasury of Physics, Astronomy, and Mathematics.*)

This course is heavily based on (1) definitions that are used to capture mathematial concepts, and on (2) using these definitions to solve mathematical problems. So we begin by defining some terms and introducing some notation we will use frequently.

A set is a collection considered as a unit. We are familiar with many sets, such as the set of integers, the set of rational numbers, and so on. In mathematics we use certain sets so often that we have abbreviated notation for them:

$\mathbb{Z}$ is the set of integers; so $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$.

$\mathbb{Q}$ is the set of rational numbers; so $\mathbb{Q} = \left\{ \frac{a}{b} : \ a, b \in \mathbb{Z}, \ b \neq 0 \ \right\}$, meaning that $\mathbb{Q}$ is the set of all objects of the form $\frac{a}{b}$ that meet the conditions that $a, b \in \mathbb{Z}$ and $b \neq 0$ (recall that $a, b \in \mathbb{Z}$ means that a, b are elements of the set $\mathbb{Z}$).

$\mathbb{R}$ is the set of real numbers.

$\mathbb{C}$ is the set of complex numbers, so $\mathbb{C} = \{a + b\sqrt{-1} : \ a, b \in \mathbb{R} \ \}$.

$\{\}$ is the empty set (i.e. the set with no elements), which is also denoted by $\emptyset$.

Note: Suppose X is a set. We cannot say "choose $x \in X$" unless we know $X \neq \emptyset$; however, we can say "suppose $x \in X$", even when we don't know whether X is nonempty.

We write $\mathbb{Z}_+$ to denote the set of positive integers, $\mathbb{Q}_+$ the set of positive rational numbers, and $\mathbb{R}_+$ the set of positive real numbers. (**Note:** 0 is neither positive nor negative.) The notation

$$A = \{x \in \mathbb{R} : \ x > \sqrt{2} \ \}$$

means that A is the set of all real numbers x that meet the condition $x > \sqrt{2}$. We write $A \subseteq X$ when X is a set and A is a subset of X, meaning that every element of A is also an element of X. We write $A \subsetneq X$ when A is a a proper subset of the set X, meaning that A is a subset of X but A is not equal to X. (The use of the notation $A \subset X$ is not consistent throughout mathematical literature, so we will avoid using this notation.) We write $A \not\subseteq B$ when A is not a subset of B.

Note that $\emptyset$ is the only subset of $\emptyset$.

Example: $\mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$.

Note: Suppose A, X are sets. Showing $A = X$ is equivalent to showing $A \subseteq X$ and $X \subseteq A$.

For A and B subsets of some set X, the notation $A \cup B$ denotes the union of A and B, meaning

$$A \cup B = \{x \in X : \ x \in A \text{ or } x \in B \ \}.$$

Similarly, for A and B subsets of some set X, the notation $A \cap B$ denotes the intersection of A and B, meaning

$$A \cap B = \{x \in X : \ x \in A \text{ and } x \in B \ \}.$$

Example: Suppose that $A = \{1, 2, 3\}$ and $B = \{3, 4, 5\}$, which are subsets of $\mathbb{Z}$. Then

$$A \cup B = \{1, 2, 3, 4, 5\} \text{ and } A \cap B = \{3\}.$$

The sets $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are sets that have some useful properties. For instance, with $X = \mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$, and for any $a, b, c \in X$, we have $a+b, -a, ab \in X$, $a+b = b+a$, $ab = ba$, and $c(a+b) = ca+cb$. Further, for $X = \mathbb{Q}$ or $\mathbb{R}$ or $\mathbb{C}$ and $a \in X$ with $a \neq 0$, we have $\frac{1}{a} \in X$. We also know that for $a, b \in \mathbb{Z}_+$, we have $a \leq ab$, and $a = ab$ only when $b = 1$. In addition, we know that for any $a, b \in \mathbb{C}$, we have $ab = 0$ only when $a = 0$ or $b = 0$; this means that for $a, b, c \in \mathbb{C}$ with $ab = ac$ and $a \neq 0$, we have $a(b-c) = 0$ and hence $b - c = 0$ so $b = c$.

A noteable property of $\mathbb{R}$ is that it is linearly ordered, meaning that for every $x, y \in \mathbb{R}$, either $x \leq y$ or $y \leq x$, and if $x, y \in \mathbb{R}$ with $x \leq y$ and $y \leq x$ then $x = y$. Note that any subset of $\mathbb{R}$ is also linearly ordered. We say a (nonempty) subset A of $\mathbb{R}$ is bounded above if there is some $M \in \mathbb{R}$ so that $M \geq a$ for all $a \in A$, and we say A is bounded below if there is some $m \in \mathbb{R}$ so that $m \leq a$ for all $a \in A$.

Proposition 1.1. *Suppose A is a nonempty subset of $\mathbb{Z}$.*

(1) *If A is bounded above then A contains a maximal element, meaning A is bounded above by an element of A.*
(2) *If A is bounded below then A contains a minimal element, meaning A is bounded below by an element of A.*

Proof. (1) Suppose A is bounded above by $N \in \mathbb{R}$. Choose any $c \in A$. Then there are finitely many integers between c and N, so there are finitely many $a \in A$ so that $c \leq a \leq N$; A is bounded above by the largest of these elements a.

(2) Suppose A is bounded below by $n \in \mathbb{R}$. Choose any $c \in A$. Then there are finitely many elements of a so that $n \leq a \leq c$; A is bounded below by the smallest of these elements a. □

Corollary 1.2. *Any nonempty subset of $\mathbb{Z}_+$ has a minimal element.*

Proof. If $A \subset \mathbb{Z}_+$ and A is nonempty, then A is a subset of $\mathbb{Z}$ that is bounded below by 1, and hence by the above theorem A has a minimal element. □

A cautionary tale regarding sets. Consider the following situation: "The barber is a man in town who shaves all those, and only those, men in town who do not shave themselves." Who shaves the barber?

In 1901 Bertrand Russell presented a version of this paradox to the mathematical community; this resulted in widespread fear that the foundations of mathematics were "built on quicksand". This paradox shows that a condition that contains an inherent contradiction does not determine a set.

There are many sources that discuss Russell's Paradox (easily found by searching the internet); students are encouraged to peruse these.

In mathematics, we are very often concerned with functions (also called maps). Some functions model the behaviour of complex systems, while other functions allow us to compare two sets. We are accustomed to functions that are given by a formula, as when studying Calculus. For instance, you might have been given $f(x) = x^2$ and been instructed to graph this for $-2 \leq x \leq 2$. To do this, you would have plotted all points of the form $(x, f(x))$ (or equivalently, (x, x^2)) for all x-values between -2 and 2. So the set $\{(x, f(x)) : \ -2 \leq x \leq 2 \ \}$ is what you might have called the graph of f for $-2 \leq x \leq 2$. Here we develop a formal definition of a function.

Definition. Given sets X, Y, we define the Cartesian product of X and Y as

$$\{(x, y) : \ x \in X, \ y \in Y \ \},$$

and we denote this set by $X \times Y$. (So $X \times Y$ is the set of all ordered pairs (x, y) that meet the conditions that $x \in X$ and $y \in Y$.) Note that if X or Y is the empty set, then so is $X \times Y$.

Example: $\mathbb{R} \times \mathbb{R} = \{(x, y) : \ x, y \in \mathbb{R} \ \}$. So $\mathbb{R} \times \mathbb{R}$ is the Cartesian plane.

Example: Let $X = \{1, 2, 3\}$, $Y = \{4, 5, 6\}$. Then

$$X \times Y = \{(1,4), (1,5), (1,6), (2,4), (2,5), (2,6), (3,4), (3,5), (3,6)\}.$$

Definitions. With X, Y nonempty sets, a function f from X into Y is a set of ordered pairs $f \subseteq X \times Y$ with the property that for each element $x \in X$ there is exactly one $y \in Y$ so that $(x, y) \in f$. (Equivalently: With X, Y nonempty sets, a function f from X into Y is a set of ordered pairs $f \subseteq X \times Y$ with the property that for each element $x \in X$ there is exactly one pair in f with first coordinate x.) When f is a function with $(x, y) \in f$, we write $f(x)$ to denote y. (So a function f from X into Y pairs each element of X with exactly one element of Y, which we denote by $f(x)$.) Thus using this notation, when f is a function from X to Y,

$$f = \{(x, f(x)) : \ x \in X \ \}.$$

We write $f : X \to Y$ to denote that f is a function from X into Y (so implicit in the notation $f : X \to Y$ is that X, Y are nonempty sets). Suppose $f : X \to Y$. We say X is the domain of f and Y is the codomain of f. The range (or image) of f, denoted $f(X)$, is the set

$$f(X) = \{f(x) : \ x \in X \ \},$$

i.e. the set of all values $f(x)$ where x meets the condition that $x \in X$. Since $f(x) \in Y$ for any $x \in X$, we also have

$$f(X) = \{y \in Y : \ \text{for some } x \in X, \ f(x) = y \ \}.$$

More generally, for any $A \subseteq X$,

$$f(A) = \{f(x) : \ x \in A \ \}.$$

Note: What we have defined here as the function f is what you may have previously called the *graph* of f.

Example: Let $X = \{x \in \mathbb{R} : \ -2 \leq x \leq 2 \ \}$, $Y = \mathbb{R}$, and

$$f = \{(x, x^2) : \ x \in \mathbb{R} \text{ and } \ -2 \leq x \leq 2 \ \}.$$

So f is a function from X into $\mathbb{R}$, and $f(x) = x^2$.

Example: Let $X = \{1, 2, 3\}$, $Y = \{4, 5, 6\}$. Let $f = \{(1, 4), (2, 5), (3, 4)\}$, $g = \{(1, 4), (1, 5), (3, 6)\}$. Then f is a function from X into Y, since for each $x \in X$, there is exactly one $y \in Y$ so that $(x, y) \in f$. However, g is not a function from X into Y: We have $1 \in X$, but there are two values of $y \in Y$ (namely $y = 4$ and $y = 5$) so that $(1, y) \in g$; further, $2 \in X$, but there is no value of $y \in Y$ so that $(2, y) \in g$. We also have

$$f(X) = \{f(1),\ f(2),\ f(3)\} = \{4,\ 5\}.$$

Example: Define $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ by $f((m, n)) = n^2$. (So the range of f is $\{n^2 : \ n \in \mathbb{Z} \ \}$.) Let $A = \{(m, n) \in \mathbb{Z} \times \mathbb{Z} : \ n = 2m \ \}$. So

$$A = \{(m, 2m) : \ m \in \mathbb{Z} \ \}.$$

Then

$$\begin{aligned} f(A) &= \{f((m, n)) : \ (m, n) \in A \ \} \\ &= \{f((m, 2m)) : \ m \in \mathbb{Z} \ \} \\ &= \{(2m)^2 : \ m \in \mathbb{Z} \ \} \\ &= \{4m^2 : \ m \in \mathbb{Z} \ \}. \end{aligned}$$

We often need to quantify objects in mathematics, meaning we need to distinguish between a condition always being met, or the existence of a case where a condition is met. Sometimes we also need to distinguish whether there is a unique case where a condition is met. For instance, suppose we have a function $f : X \to Y$. This means that for every $x \in X$ there exists a unique $y \in Y$ so that $(x, y) \in f$. Notice that the order of the quantifying phrases is important:

"For every $x \in X$, there is a unique $y \in Y$ so that $(x, y) \in f$" means that the choice of x determines the value of y (in this particular situation, we have $y = f(x)$). Contrastingly, "There exists a unique $y \in Y$ so that for every $x \in X$, $(x, y) \in f$" means that there is a unique $y \in Y$ so that for every $x \in X$, we have $f(x) = y$, meaning f is a constant function (as in the case $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 5$).

Notation: We use the symbol $\forall$ to denote "for all", or equivalently, "for every". We use the symbol $\exists$ to denote "there exists", and we use $\exists!$ to denote "there exists a unique", or equivalently "there exists one and only one".

Notes: (1) To show that something is unique, a standard technique is to show first that one such thing exists, and to show then that if another exists, it is equal to the first. For example, suppose X, Y are sets and $f \subseteq X \times Y$. Then f is a function from X to Y if, $\forall x \in X$, $\exists y \in Y$ so that $(x, y) \in f$, and $\forall y' \in Y$, if $(x, y') \in f$ then $y' = y$.

(2) When we write "Suppose $c \in X$" or "Choose $c \in X$" or "Take $c \in X$" without stating further assumptions on c, we mean that we are choosing c

arbitrarily from X; thus anything we then conclude about c applies to every element of X.

(3) We will sometimes write "for $c \in X$" to mean "$\forall c \in X$", as the only condition being imposed on c is that it is in X. Somewhat similarly, we will sometimes write "for some $c \in X$" to mean "$\exists c \in X$".

Theorem 1.3. *Suppose $f : X \to Y$, $g : X \to Y$. Then $f = g$ if and only if $\forall x \in X$, $f(x) = g(x)$.*

Proof. First suppose that $\forall x \in X$, $f(x) = g(x)$. Thus

$$f = \{(x, f(x)) : \ x \in X \ \} = \{(x, g(x)) : \ x \in X \ \} = g.$$

Now suppose that $f = g$. Thus $(x, y) \in f$ if and only if $(x, y) \in g$. Take [arbitrary] $x \in X$, and then choose [the unique] $y \in Y$ so that $(x, y) \in f = g$; thus $y = f(x)$, and also $y = g(x)$. Hence for every $x \in X$, we have $f(x) = g(x)$. □

Definitions. We say a function $f : X \to Y$ is injective (or one-to-one, or an injection) if, $\forall x_1, x_2 \in X$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$. (Here $\forall x_1, x_2 \in X$ means $\forall x_1 \in X$, $\forall x_2 \in X$.) We say a function $f : X \to Y$ is surjective (or onto, or a surjection) if, $\forall\ y \in Y$, $\exists\ x \in X$ so that $f(x) = y$. (Thus $f : X \to Y$ is surjective if the range of f is Y.) A function is called bijective if it is both injective and surjective.

Note: We have defined (for example) a map $f : X \to Y$ to be injective if, $\forall x_1, x_2 \in X$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$. According to standard English usage, a definition is a precise statement of what a word or expression means. Thus saying that a map $f : X \to Y$ is injective is *equivalent* to saying that $\forall x_1, x_2 \in X$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$.

Example: Define $f : \mathbb{Z}_+ \to \mathbb{Z}_+$ by $f(x) = x^2$. This function is injective but not surjective.

Example: Let $\mathbb{R}_{\geq 0} = \{y \in \mathbb{R} : \ y \geq 0 \ \}$. Define $g : \mathbb{R} \to \mathbb{R}_{\geq 0}$ by $g(x) = x^2$; this function is surjective, but not injective.

Example: Define $h : \mathbb{R} \to \mathbb{R}$ by $h(x) = x^3$; this function is bijective.

Warning: Do not confuse the definition of injective with the definition of a function. For example, consider $f = \{(y^2, y) : \ y \in \mathbb{Z} \ \} \subseteq \mathbb{Z} \times \mathbb{Z}$. So for each $y \in \mathbb{Z}$, $\exists!\ x \in \mathbb{Z}$ so that $(x, y) \in f$ (namely $x = y^2$). But f is not a function, as, for example, $(4, 2), (4, -2) \in f$.

Example: Define $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R}$ by

$$f(m, n) = (m + n, m - n).$$

We claim that f is surjective. We begin by choosing $(u, v) \in \mathbb{R} \times \mathbb{R}$, the codomain of f.

[We want to find $(m, n) \in \mathbb{R} \times \mathbb{R}$, the domain of f, so that $f(m, n) = (u, v)$. So we need to find $m, n \in \mathbb{R}$ so that $m + n = u$ and $m - n = v$. To have these equalities, we need $m = u - n$ and $m = v + n$. To have these last two equalities, we need $u - n = v + n$, or equivalently, $u - v = 2n$, or equivalently, $\frac{u-v}{2} = n$. If we have $\frac{u-v}{2} = n$ and $m = u - n$, then we have

$$m = u - \frac{u - v}{2} = \frac{u + v}{2}.$$

Thus, having worked backwards to find m and n, we take these values for m and n and, with hope, can show that $f(m,n) = (u,v)$.]

Having chosen $(u,v) \in \mathbb{R} \times \mathbb{R}$, we set $m = \frac{u+v}{2}$, $n = \frac{u-v}{2}$. Thus $(m,n) \in \mathbb{R} \times \mathbb{R}$, the domain of f. Then

$$f(m,n) = (m+n, m-n) = \left(\frac{u+v}{2} + \frac{u-v}{2}, \frac{u+v}{2} - \frac{u-v}{2}\right) = (u,v).$$

This shows that f is surjective.

Proposition 1.4. *Suppose $f : X \to Y$.*

(a) *f is injective if and only if $\forall y \in f(X)$, $\exists! x \in X$ so that $f(x) = y$.*
(b) *f is bijective if and only if $\forall\ y \in Y$, $\exists!\ x \in X$ so that $f(x) = y$.*

(So when f is bijective, f gives us a one-to-one correspondence between the elements of X and the elements of Y.)

Proof. (a) Suppose first that f is injective, and suppose $y \in f(X)$. Thus $\exists x \in X$ so that $f(x) = y$. Now suppose $x' \in X$ so that $x' \neq x$. Then since f is injective, $f(x') \neq f(x) = y$. Thus x is the only element of X so that $f(x) = y$; in other words, x is the unique element of X so that $f(x) = y$. In summary, for $y \in f(X)$, $\exists! x \in X$ so that $f(x) = y$.

Now suppose that $\forall y \in f(X)$, $\exists! x \in X$ so that $f(x) = y$. Suppose $x_1, x_2 \in X$ so that $x_1 \neq x_2$; let $y_1 = f(x_1)$. By assumption, x_1 is the only element of X that f maps to y_1. Hence $y_1 \neq f(x_2)$, so $f(x_1) \neq f(x_2)$. Thus we have shown that for $x_1, x_2 \in X$ with $x_1 \neq x_2$, we have $f(x_1) \neq f(x_2)$, meaning that f is injective.

(b) Say f is bijective; then $f(X) = Y$, so by (a), $\forall y \in Y$, $\exists! x \in X$ so that $f(x) = y$.

Now suppose that $\forall y \in Y$, $\exists! x \in X$ so that $f(x) = y$. Then f is surjective, since $\forall y \in Y$, $\exists x \in X$ so that $f(x) = y$. Therefore $f(X) = Y$, and by (a), f is injective. Hence f is bijective. □

Definitions. Suppose we have functions $f : X \to Y$, $g : Y \to Z$. We define the composition of g and f, denoted $g \circ f$, by

$$(g \circ f)(x) = g(f(x)) \text{ for any } x \in X.$$

Since f assigns to $x \in X$ exactly one value $f(x) \in Y$, and g assigns to $f(x) \in Y$ exactly one value in Z, we have that $g \circ f$ is a function from X to Z, i.e. $g \circ f : X \to Z$. We say a function $f : X \to Y$ is invertible if there exists a function $g : Y \to X$ so that $g \circ f$ is the identity function on X (meaning that for all $x \in X$, $(g \circ f)(x) = x$), and $f \circ g$ is the identity function on Y. Note that when g is an inverse for f, we also have that f is an inverse for g.

Example: Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = 2x + 3$ and define $g : \mathbb{R} \to \mathbb{R}$ by $g(x) = (x-3)/2$. Then for any $x \in \mathbb{R}$,

$$g \circ f(x) = g(f(x)) = \frac{f(x) - 3}{2} = \frac{(2x+3) - 3}{2} = x,$$

and

$$f \circ g(x) = f(g(x)) = 2g(x) + 3 = 2 \cdot \frac{x-3}{2} + 3 = x.$$

Hence $g \circ f$ is the identity function on the domain of f, and $f \circ g$ is the identity function on the domain of g. So f is invertible with g as an inverse.

Proposition 1.5. *Suppose $f : X \to Y$, $g : Y \to Z$, $h : Z \to W$. Then $h \circ (g \circ f) = (h \circ g) \circ f$.*

Proof. To show $h \circ (g \circ f) = (h \circ g) \circ f$, we need to show that for all $x \in X$, we have $h \circ (g \circ f)(x) = (h \circ g) \circ f(x)$. So take $x \in X$; then

$$h \circ (g \circ f)(x) = h(g \circ f(x)) = h(g(f(x)))$$

and

$$(h \circ g) \circ f(x) = h \circ g(f(x)) = h(g(f(x))).$$

Thus $h \circ (g \circ f) = (h \circ g) \circ f$. □

Theorem 1.6. *Suppose $f : X \to Y$, $g : Y \to Z$.*

(a) *If f and g are injective then so is $g \circ f$.*
(b) *If f and g are surjective then so is $g \circ f$.*

Proof. We will prove (a) and leave (b) as an exercise.

Suppose f, g are injective, and suppose $x_1, x_2 \in X$ so that $x_1 \neq x_2$. Since f is injective, this means that $f(x_1) \neq f(x_2)$. Set $y_1 = f(x_1)$, $y_2 = f(x_2)$. Thus $y_1, y_2 \in Y$ with $y_1 \neq y_2$. Since g is injective, this means $g(y_1) \neq g(y_2)$. Substituting for y_1, y_2, this means

$$g \circ f(x_1) = g(f(x_1)) = g(y_1) \neq g(y_2) = g(f(x_2)) = g \circ f(x_2).$$

Summarising, for any $x_1, x_2 \in X$, if $x_1 \neq x_2$ then $g \circ f(x_1) \neq g \circ f(x_2)$. Hence $g \circ f$ is injective. □

Note: This theorem shows that if $f : X \to Y$ and $g : Y \to Z$ are both bijective, then $g \circ f : X \to Z$ is also bijective.

As an exercise, one proves the following.

Theorem 1.7. *Suppose $f : X \to Y$, and $g : Y \to X$, $h : Y \to X$ are inverses of f. Then $g = h$; that is, if f has an inverse then its inverse is unique.*

Proof by contradiction: A proof by contradiction proceeds as follows. We want to prove a certain statement P is true. So instead, we assume that P is false, and we use this to deduce as true something we know to be false. Hence we conclude that it is impossible that P is falst, and thus P must be true. We use this technique in part of the proof of the next theorem.

Theorem 1.8. *Suppose $f : X \to Y$. Then f is invertible if and only if f is bijective.*

Proof. There are 2 statements we need to prove:

(1) f is invertible only if f is bijective, or equivalently, f is invertible implies f is bijective.
(2) f is invertible if f is bijective, or equivalently, f is bijective implies f is invertible;

To show (1): Suppose f is invertible, and let $g : Y \to X$ denote an inverse of f. Thus $g \circ f$ is the identity function on X, and $f \circ g$ is the identity function on Y.

To show that f is injective, suppose that $x_1, x_2 \in X$ so that $x_1 \neq x_2$. So

$$x_1 = g \circ f(x_1) = g(f(x_1)), \text{ and } x_2 = g \circ f(x_2) = g(f(x_2)).$$

[We now proceed to argue by contradiction: We want to deduce that $f(x_1) \neq f(x_2)$. So we show that if $f(x_1) = f(x_2)$ then we obtain a contradiction to something we know to be true, and thus it is impossible to have $f(x_1) = f(x_2)$.] For the sake of contradiction, suppose that $f(x_1) = f(x_2)$. Then we must have $g(f(x_1)) = g(f(x_2))$. But this contradicts our deduction above that $g(f(x_1)) \neq g(f(x_2))$. Hence it cannot be the case that $f(x_1) = f(x_2)$, so we must have that $f(x_1) \neq f(x_2)$. This shows that (having assumed that f is invertible) if $x_1, x_2 \in X$ with $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$; that is, this shows that f is injective.

We now show that f is surjective [note that we are still assuming that f is invertible with inverse g]. We begin by choosing [arbitrary] $y \in Y$. [We need to find $x \in X$ so that $f(x) = y$.] We know that $f \circ g$ is the identity function on Y, so

$$y = f \circ g(y) = f(g(y)).$$

Set $x = g(x)$. Thus $x \in X$, and

$$f(x) = f(g(y)) = y.$$

This shows that f is surjective.

Therefore we have shown that when f is invertible, then f is injective and surjective, i.e. f is bijective.

To show (2): Suppose f is bijective. [So we are assuming that f is injective and surjective.] Set

$$g = \{(y, x) \in Y \times X : \ (x, y) \in f \ \}.$$

Since f is bijective, $\forall y \in Y$, $\exists! x \in X$ so that $(x, y) \in f$; thus g is a function. So $g : Y \to X$, and for any $y \in Y$, $g(y) = x$ where $f(x) = y$. Now we need to show that g is an inverse of f. For this, first take any $x \in X$. Set $y = f(x)$. Thus by the definition of g, $g(y) = x$, so $(g \circ f)(x) = x$. As x was chosen arbitrarily from X, this shows $g \circ f$ is the identity function on X. Now choose any $y \in Y$. Since f is bijective, there is a unique $x \in X$ with $f(x) = y$. Thus $g(y) = x$, and hence $(f \circ g)(y) = f(x) = y$. Since y was chosen arbitrarily from Y, this shows $f \circ g$ is the identity function on Y. Hence when f is bijective, we have that f is invertible. Thus (1) $\implies$ (2). $\square$

Note: Suppose $f : X \to Y$ bijective. In proving the above theorem, we found a "recipe" for defining $f^{-1} : Y \to X$:

For any $y \in Y$, $f^{-1}(y) = x$ where $x \in X$ so that $f(x) = y$.

Example: Suppose $a, b, c, d \in \mathbb{R}$ so that $a < b$ and $c < d$. Let $[a, b]$ denote the closed interval from a to b; that is,

$$[a, b] = \{x \in \mathbb{R} : \ a \leq x \leq b \ \}.$$

We claim there is a bijection between $[a, b]$ and $[c, d]$. Intuitively, the idea is the we stretch or shrink the interval $[a, b]$ to be the same length as $[c, d]$, and shift this. The map $f_1(x) = x - a$ will take $[a, b]$ to $[0, b - a]$, then

$f_2(x) = x \cdot \frac{d-c}{b-a}$ will take $[0, b-a]$ to $[0, d-c]$, and then $f_3(x) = c+x$ will take $[0, d-c]$ to $[c, d]$. We set $f = f_3 \circ f_2 \circ f_1$; thus we set $f(x) = c + \frac{(x-a)(d-c)}{(b-a)}$. We want to show $f : [a, b] \to [c, d]$. To do this, take $x \in [a, b]$, we have $0 \leq x - a \leq b - a$. We know $a < b$ and $c < d$, so $b - a > 0$ and $d - c > 0$. Hence

$$0 \leq \frac{(x-a)(d-c)}{(b-a)} \leq (d-c),$$

and then

$$c \leq c + \frac{(x-a)(d-c)}{(b-a)} \leq d.$$

So we indeed have that $f : [a, b] \to [c, d]$. (**Warning:** One may be tempted to argue by first *assuming* that $c \leq f(x) \leq d$, and then *deducing* that $a \leq x \leq b$, but what we need to show is that *if* $x \in [a, b]$ *then* $f(x) \in [c, d]$. In one's scratch work one might first assume that $c \leq f(x) \leq d$ and then deduce that $a \leq x \leq b$, but then one must determine whether these steps can be reversed to obtain a proof of what is needed. More generally, to prove a statement of the form "If A then B", it is **incorrect** to begin by assuming what is to be deduced.)

Now we want to show that f is bijective. So we could argue that f is injective and surjective. Using the definition of injective that we have given, it is awkward to show that f is injective; in §3 we will use a result of §2 to produce an equivalent definition of injective, using the "contrapositive" of the definition we have given. (The contrapositive of a statement of the form "If A holds then B holds" is "If B does not hold then A does not hold"; in §2 we will see that the contrapositive of a statement is equivalent to the statement.) In arguing that this particular function is surjective, we would actually produce the inverse of f, so here we will argue that f is bijective by finding $g : [c, d] \to [a, b]$ so that $g \circ f$ is the identity map on $[a, b]$ and $f \circ g$ is the identity map on $[c, d]$.

Using the strategy we used to construct f, reversing the roles of a and c and the roles of b and d, we define $g(x) = a + \frac{(x-c)(b-a)}{(d-c)}$. [Alternatively, we could set $y = f(x)$ and solve for x, finding that $x = a + \frac{(y-c)(b-a)}{(d-c)}$, and then setting $g(y) = a + \frac{(y-c)(b-a)}{(d-c)}$.] Then for $x \in [c, d]$, we have $c \leq x \leq d$ and hence $a \leq a + \frac{(x-c)(b-a)}{(d-c)} \leq b$; so $g : [c, d] \to [a, b]$. Also, for $x \in [a, b]$,

$$\begin{aligned} g \circ f(x) &= g(f(x)) \\ &= a + (f(x) - c)\frac{(b-a)}{(d-c)} \\ &= a + \left(c + (x-a)\frac{(d-c)}{(b-a)} - c\right)\frac{(b-a)}{(d-c)} \\ &= a + (x-a) \\ &= x. \end{aligned}$$

Similarly,

$$\begin{aligned} f \circ g(x) &= f(g(x)) \\ &= c + (g(x) - a)\frac{(d-c)}{(b-a)} \\ &= c + \left(a + (x-c)\frac{(b-a)}{(d-c)} \right) \frac{(d-c)}{(b-a)} \\ &= x. \end{aligned}$$

Thus $g : [c, d] \to [a, b]$ is the inverse of f.

Note: Given the above definitions of f and g, it is necessary to ensure that $f([a, b]) \subseteq [c, d]$ and that $g([c, d]) \subseteq [a, b]$, else we cannot claim that $f : [a, b] \to [c, d]$ and $g : [c, d] \to [a, b]$, and knowing the domains and codomains of f and g is necessary to apply the preceding theorem. We could define $f : [a, b] \to \mathbb{R}$ by $f(x) = c + \frac{(x-a)(d-c)}{(b-a)}$ and $g : [0, 2] \to \mathbb{R}$ by $g(x) = a + \frac{(c-x)(b-a)}{(d-c)}$, and then proceed mechanically to argue that $g \circ f(x) = x$, $f \circ g(x) = x$; this will work because we could have extended the domains of f and g to $\mathbb{R}$, but unless $c = 0$ and $d = 2$, this does not *prove* that there is a bijection between $[a, b]$ and $[0, 1]$.

In the exercises, one proves the following. (Part (a) of this theorem is an exercise for §3, and part (b) is an exercise for this section.)

Proposition 1.9. *Suppose $f : X \to Y$, $g : Y \to X$ so that $g \circ f$ is the identity map on X, meaning that for all $x \in X$, we have $g \circ f(x) = x$.*

(a) *Suppose g is injective; then $f \circ g$ is the identity map on Y (and hence $g = f^{-1}$).*
(b) *Suppose f is surjective; then $f \circ g$ is the identity map on Y (and hence $g = f^{-1}$).*

As an exercise, one also proves the following.

Theorem 1.10. *Suppose $f : X \to Y$ and $g : Y \to Z$ are bijective (and hence we know $g \circ f$ is bijective). Then $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

One also proves this useful result.

Proposition 1.11. *Suppose $f : X \to Y$ is bijective, and $A \subseteq X$. Set $B = \{x \in X : x \notin A\}$. (Standard notation for B is $X \setminus A$.) Then $f(A) \cap f(B) = \emptyset$.*

2. Truth tables, equivalences, and contrapositive

We use the word "statement" interchangeably with the word "sentence", and we agree that a statement can be true or false or neither, but a statement cannot be simultaneously true and false. In a mathematical system, the true statements and false statements are the propositions of the system, and the label "true" or "false" associated with a given proposition is its truth value.

Notation: We use the symbol $\neg$ to mean "not". We use the symbol $\wedge$ to mean "and". We use the symbol $\vee$ to mean "or". (Note that we do not use

$\vee$ to mean "exclusive or"; that is, $P \vee Q$ is true if P is true or if Q is true or if both P and Q are true.) We use the symbol $\implies$ to mean "implies". (So $P \implies Q$ means that if P is true then Q is true.) We use the symbol $\iff$ to mean "if and only if"; so with P, Q propositions, $P \iff Q$ means that $P \implies Q$ and $Q \implies P$. (So $P \iff Q$ means that P is true exactly when Q is true, and P is false exactly when Q is false.) When $P \iff Q$, we say P and Q are equivalent.

Example: With $x \in \mathbb{Z}$, we could have P representing the proposition "$x \geq 5$" and Q representing the proposition "$x \leq 7$". Then $P \wedge Q$ would represent the proposition "$x \geq 5$ and $x \leq 7$".

When a proposition P is true, we sometimes express this by saying that P holds.

Example: Suppose P and Q represent propositions. $P \implies Q$ is the proposition that P implies Q, or in other words, the proposition that if P is true then Q is true. To state this more emphatically, $P \implies Q$ means that **if** P is true, then Q **must** be true. Note that $P \implies Q$ allows for P and Q to both be true, or for P to be false and Q to be true, or for P and Q to both be false. However, $P \implies Q$ does not allow for P to be true and Q to be false. (Initially, it can seem confusing that $P \implies Q$ is true when P and Q are false. However, having P and Q false does not contradict that Q must be true if P is true.) We can represent this scenario using what is called a "truth table", wherein we consider all possible combinations of the truth values of P and Q, and the consequent truth value of $P \implies Q$:

P	Q	$[P \implies Q]$
T	T	T
T	F	F
F	T	T
F	F	T

(The square brackets on the top line of the truth table are used simply to make it easier to distinguish the three propositions from each other.)

Note: We could prove the truth of the following propositions and theorems without using truth tables, but here we use truth table to establish some fundamental and useful results in a rather painless way.

Example: Suppose P and Q represent propositions. $P \wedge Q$ is true exactly when P and Q are both true. So the corresponding truth table is:

P	Q	$[P \wedge Q]$
T	T	T
T	F	F
F	T	F
F	F	F

Example: Suppose P and Q represent propositions. $P \vee Q$ is true exactly when P or Q is true. We do not use the word "or" to mean "exclusive or",

so $P \vee Q$ is true when P and Q are both true. So the corresponding truth table is:

P	Q	$[P \vee Q]$
T	T	T
T	F	T
F	T	T
F	F	F

Example: Suppose P and Q represent propositions. The corresponding truth table for $(\neg P) \vee Q$ is:

P	Q	$[\neg P \vee Q]$
T	T	T
T	F	F
F	T	T
F	F	T

Example: Suppose P and Q represent propositions. The corresponding truth table for $\neg Q \implies \neg P$ is:

P	Q	$[\neg Q \implies \neg P]$
T	T	T
T	F	F
F	T	T
F	F	T

This truth table can be easier to determine by expanding it as follows.

P	Q	$\neg P$	$\neg Q$	$[\neg Q \implies \neg P]$
T	T	F	F	T
T	F	F	T	F
F	T	T	F	T
F	F	T	T	T

Theorem 2.1. *Suppose P, Q are propositions.*

(a) *$P \implies Q$ is equivalent to $\neg Q \implies \neg P$.*
(b) *$P \implies Q$ is equivalent to $\neg P \vee Q$.*

Proof. We prove (a) and leave (b) as an exercise.

P	Q	$[P \implies Q]$	$[\neg Q \implies \neg P]$	$[(P \implies Q) \iff (\neg Q \implies \neg P)]$
T	T	T	T	T
T	F	F	F	T
F	T	T	T	T
F	F	T	T	T

So for all truth values of P and Q, $(P \implies Q)$ and $(\neg Q \implies \neg P)$ have the same truth values. Hence $[(P \implies Q) \iff (\neg Q \implies \neg P)]$. □

Definitions. We call the proposition $\neg Q \implies \neg P$ the contrapositive of the proposition $P \implies Q$. As seen above, the proposition $P \implies Q$ is equivalent to its contrapositive. The proposition $Q \implies P$ is called the converse of the proposition $P \implies Q$; as an exercise one shows that $Q \implies P$ is not equivalent to $P \implies Q$.

We also have this easily proved result.

Proposition 2.2. *Suppose P is a proposition. Then $P \iff \neg(\neg P)$.*

Proof.

P	$\neg P$	$\neg(\neg P)$
T	F	T
F	T	F

So the truth values of P and $\neg(\neg P)$ always agree, so the propositions P and $\neg(\neg P)$ are equivalent. □

The next proposition shows that $\wedge$ and $\vee$ are "associative", meaning that $P \wedge Q \wedge R$ and $P \vee Q \vee R$ are propositions that do not require parentheses.

Proposition 2.3. *Suppose P, Q, R are propositions.*

(a) $(P \wedge Q) \wedge R \iff P \wedge (Q \wedge R)$.
(b) $(P \vee Q) \vee R \iff P \vee (Q \vee R)$.

Proof. We prove (a), and leave the proof of (b) as an exercise.

P	Q	R	$(P \wedge Q)$	$[(P \wedge Q) \wedge R]$
T	T	T	T	T
T	T	F	T	F
T	F	T	F	F
T	F	F	F	F
F	T	T	F	F
F	T	F	F	F
F	F	T	F	F
F	F	F	F	F

Also:

P	Q	R	$(Q \wedge R)$	$[P \wedge (Q \wedge R)]$
T	T	T	T	T
T	T	F	F	F
T	F	T	F	F
T	F	F	F	F
F	T	T	T	F
F	T	F	F	F
F	F	T	F	F
F	F	F	F	F

So for any truth values of P, Q, R, these truth tables show that $[(P \wedge Q) \wedge R] \iff [P \wedge (Q \wedge R)]$.

Note that one could combine the above truth tables into one (large) table, or just combine some of the information from these two truth tables into one truth table as follows:

P	Q	R	$[(P \wedge Q) \wedge R]$	$[P \wedge (Q \wedge R)]$	$[(P \wedge Q) \wedge R] \iff [P \wedge (Q \wedge R)]$
T	T	T	T	T	T
T	T	F	F	F	T
T	F	T	F	F	T
T	F	F	F	F	T
F	T	T	F	F	T
F	T	F	F	F	T
F	F	T	F	F	T
F	F	F	F	F	T

□

The above proposition shows we can write $P \wedge Q \wedge R$ and $P \vee Q \vee R$, without there being confusion. As a trivial exercise, one can also show the following sometimes useful equivalences.

Proposition 2.4. *Suppose P, Q, R are propostions. Then*

$$P \wedge Q \wedge R \iff (P \wedge Q) \wedge (P \wedge R), \text{ and } P \vee Q \vee R \iff (P \vee Q) \vee (P \vee R).$$

Proposition 2.5. *Suppose P, Q, R are propositions.*

(a) $P \wedge (Q \vee R) \iff (P \wedge Q) \vee (P \wedge R)$.
(b) $P \vee (Q \wedge R) \iff (P \vee Q) \wedge (P \vee R)$.

Proof. We prove (a) and leave the proof of (b) as an exercise.

P	Q	R	$(Q \vee R)$	$[P \wedge (Q \vee R)]$
T	T	T	T	T
T	T	F	T	T
T	F	T	T	T
T	F	F	F	F
F	T	T	T	F
F	T	F	T	F
F	F	T	T	F
F	F	F	F	F

Also:

P	Q	R	$(P \wedge Q)$	$(P \wedge R)$	$[(P \wedge Q) \vee (P \wedge R)]$
T	T	T	T	T	T
T	T	F	T	F	T
T	F	T	F	T	T
T	F	F	F	F	F
F	T	T	F	F	F
F	T	F	F	F	F
F	F	T	F	F	F
F	F	F	F	F	F

So for any truth values of P, Q, R, these truth tables show that

$$[P \wedge (Q \vee R)] \iff [(P \wedge Q) \vee (P \wedge R)].$$

□

Theorem 2.6. *Suppose P, Q are propositions.*

(a) $\neg(P \wedge Q) \iff \neg P \vee \neg Q$.
(b) $\neg(P \vee Q) \iff \neg P \wedge \neg Q$.
(c) $\neg(P \implies Q) \iff (P \wedge \neg Q)$.

Proof. Using a truth table, we prove (a) and leave (b) and (c) as exercises.

P	Q	$\neg P$	$\neg Q$	$P \wedge Q$	$\neg(P \wedge Q)$	$\neg P \vee \neg Q$
T	T	F	F	T	F	F
T	F	F	T	F	T	T
F	T	T	F	F	T	T
F	F	T	T	F	T	T

Thus for any truth values of P, Q, R, the truth values of $\neg(P \wedge Q)$ and $\neg P \vee \neg Q$ are the same. This proves (1). □

As an exercise, one proves the following.

Proposition 2.7. *Suppose P, Q are propositions. Then $[P \vee Q] \iff [\neg P \implies Q]$.*

Note: With P, Q, R propositions, $P \implies Q \iff R$ and $P \iff Q \implies R$ do not have clear meanings. As exercises, one shows that the statements $P \implies (Q \iff R)$ and $(P \implies Q) \iff R$ are not equivalent, and $P \iff (Q \implies R)$ and $(P \iff Q) \implies R$ are not equivalent. Note that this also means an assertion such as

$$P \implies Q \iff R \implies S$$

has no clear meaning.

We give two proofs of the next theorem; one is a proof by contradiction, and the other is a proof by contrapositive.

Theorem 2.8. *(Pigeonhole Principle) Let A be a set with n elements and B a set with m elements where $m, n \in \mathbb{Z}_+$ with $m < n$. Then there is no injection from A into B. (So if n pigeons fly into m pigeonholes, then at least one pigeonhole contains more than one pigeon.)*

Proof. **Proof 1:** For the sake of contradiction, suppose $g : A \to B$ is injective. Enumerate the elements of A as $a_1, a_2, \ldots, a_n$ and the elements of B as $b_1, b_2, \ldots, b_m$. Let $C = \{g(a_i) : \ i \in \mathbb{Z}_+, \ i \leq n \ \}$. Thus C is a subset of B, and since g is injective, C is a set with n elements. But this means there is a subset of B containing more elements than are in B, which is impossible. Thus it cannot be possible to have an injective function $g : A \to B$.

Proof 2: The statement of the theorem is equivalent to "Let A be a set with n elements and B a set with m elements where $m, n \in \mathbb{Z}_+$. If $m < n$ then there is no injection from A into B." The contrapositive of this statement is "Let A be a set with n elements and B a set with m elements where $m, n \in \mathbb{Z}_+$. If there is an injection from A into B then $m \geq n$." We will prove this latter statement. Suppose $g : A \to B$ is injective. Enumerate the elements of A as $a_1, a_2, \ldots, a_n$ and the elements of B as $b_1, b_2, \ldots, b_m$. Let $C = \{g(a_i) : \ i \in \mathbb{Z}_+, \ i \leq n \ \}$. Thus C is a subset of B, and since g is injective, C is a set with n elements. Thus B must have at least n elements, meaning $m \geq n$.

□

Sometimes one can prove a result by contrapositive using an argument that is almost identical to proving the result by contradiction (as above). However, there are occassions where this is not the case; we will see an example of this later in the course when we prove by contradiction that the interval $(0, 1) \subseteq \mathbb{R}$ is what we call "uncountable".

3. Negations and contrapositives of propositions with quantifiers

Suppose $P(x)$ is a proposition involving x (where $x \in X$, X some set). Suppose the proposition

$$\forall x \in X, P(x)$$

is **not** true. Then there must be an exceptional $x \in X$ so that $P(x)$ does not hold. That is,

$$\neg(\forall x \in X, \ P(x)) \implies (\exists x \in X \text{ so that } \neg P(x)).$$

Conversely, suppose the proposition

$$\exists x \in X \text{ so that } \neg P(x)$$

is true. Then it is not the case that $P(x)$ holds for all $x \in X$, meaning

$$(\exists x \in X \text{ so that } \neg P(x)) \implies \neg(\forall x \in X, \ P(x)).$$

Thus

$$\neg(\forall x \in X, \ P(x)) \iff (\exists x \in X \text{ so that } \neg P(x)).$$

This means we also have

$$\begin{aligned} \neg(\exists x \in X \text{ so that } \neg P(x)) &\iff \neg(\neg(\forall x \in X, \ P(x)) \\ &\iff (\forall x \in X, \ P(x)). \end{aligned}$$

(Recall that for a proposition R, $\neg(\neg R)$ is equivalent to R.) Letting $Q(x) = \neg P(x)$, this gives us

$$\neg(\exists x \in X \text{ so that } Q(x)) \iff (\forall x \in X, \ \neg Q(x)).$$

Note: We are inserting phrases like "such that" to make our sentences more readable without changing their meanings.

Example: Recall that by definition, $f : X \to Y$ is injective if and only if

$$\forall x_1 \in X,\ \forall x_2 \in X,\ x_1 \neq x_2 \implies f(x_1) \neq f(x_2).$$

Let $P(x_1)$ be the proposition that $\forall x_2 \in X,\ x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$. (So f is injective if and only if $\forall x_1 \in X,\ P(x_1)$.) We know that

$$\neg(\forall x_1 \in X,\ P(x_1)) \text{ is equivalent to } (\exists x_1 \in X \text{ so that } \neg P(x_1)).$$

Now let $Q(x_1, x_2)$ be the proposition $x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$. Then

$$\begin{aligned} \neg P(x_1) &\iff \neg(\forall x_2 \in X,\ Q(x_1, x_2)) \\ &\iff \exists x_2 \in X \text{ so that } \neg Q(x_1, x_2). \end{aligned}$$

Also, using results from §2, we have

$$\begin{aligned} \neg Q(x_1, x_2) &\iff [x_1 \neq x_2 \ \wedge\ \neg(f(x_1) \neq f(x_2))] \\ &\iff [x_1 \neq x_2 \ \wedge\ f(x_1) = f(x_2)]. \end{aligned}$$

Summarising, $f : X \to Y$ is **not** injective if and only if

$$\exists x_1 \in X,\ \exists x_2 \in X \text{ so that } x_1 \neq x_2 \ \wedge\ f(x_1) = f(x_2).$$

Example: Suppose $f : X \to Y$. By definition, we know f is surjective if and only if

$$\forall y \in Y,\ \exists x \in x \text{ so that } f(x) = y.$$

Let $P(y)$ be the proposition $\exists x \in x$ so that $f(x) = y$. Thus

$$\begin{aligned} f \text{ is not surjective} &\iff \neg(\forall y \in Y,\ P(y)) \\ &\iff \exists y \in Y \text{ such that } \neg P(y) \\ &\iff \exists y \in Y \text{ so that } \neg(\exists x \in X \text{ so that } f(x) = y) \\ &\iff \exists y \in Y \text{ so that } [\forall x \in X,\ \neg(f(x) = y)] \\ &\iff \exists y \in Y \text{ so that } [\forall x \in X,\ f(x) \neq y]. \end{aligned}$$

Example: For every $n \in \mathbb{Z}_+$, suppose $a_n \in \mathbb{R}$. Consider the following proposition:

$$\exists c \in \mathbb{R} \text{ so that } \forall \varepsilon > 0,\ \exists N \in \mathbb{Z}_+ \text{ so that } \forall n \geq N,\ |a_n - c| < \varepsilon.$$

We negate this proposition in a series of steps so that each consecutive pair of propositions are clearly equivalent:

$$\begin{aligned} &\neg[\exists c \in \mathbb{R} \text{ so that } \forall \varepsilon > 0,\ \exists N \in \mathbb{Z}_+ \text{ so that } \forall n \geq N,\ |a_n - c| < \varepsilon] \\ &\iff \forall c \in \mathbb{R},\ \neg[\forall \varepsilon > 0,\ \exists N \in \mathbb{Z}_+ \text{ so that } \forall n \geq N,\ |a_n - c| < \varepsilon] \\ &\iff \forall c \in \mathbb{R},\ \exists \varepsilon > 0 \text{ so that } \neg[\exists N \in \mathbb{Z}_+ \text{ so that } \forall n \geq N,\ |a_n - c| < \varepsilon] \\ &\iff \forall c \in \mathbb{R},\ \exists \varepsilon > 0 \text{ so that } \forall N \in \mathbb{Z}_+,\ \neg[\forall n \geq N,\ |a_n - c| < \varepsilon] \\ &\iff \forall c \in \mathbb{R},\ \exists \varepsilon > 0 \text{ so that } \forall N \in \mathbb{Z}_+,\ \exists n \geq N \text{ so that } \neg[|a_n - c| < \varepsilon] \\ &\iff \forall c \in \mathbb{R},\ \exists \varepsilon > 0 \text{ so that } \forall N \in \mathbb{Z}_+,\ \exists n \geq N \text{ so that } |a_n - c| \geq \varepsilon. \end{aligned}$$

Note: With P, Q propositions, the proposition "$P \implies Q$" is equivalent to the proposition "if P then Q". When we have complex proposition involving quantifiers and an implication, it can be important to know where the word "if" belongs. Here we consider an example of this.

Example: Suppose $A \subseteq \mathbb{R}$ with $A \neq \emptyset$. For $L \in \mathbb{R}$, we say L is an upper bound for A if, $\forall a \in A$, $a \leq L$. We say $L \in \mathbb{R}$ is a least upper bound for A if (1) L is an upper bound for A, and (2) if $M \in \mathbb{R}$ is an upper bound for A, then $L \leq M$. Let $P(M)$ be the proposition that M is an upper bound for A (so $P(M)$ means that $\forall a \in A,\ a \leq M$). Thus L is a least upper bound for A if and only if $[P(L) \wedge (\forall M \in \mathbb{R},\ P(M) \implies (L \leq M))]$. Notice that the quantifier on $a \in A$ is part of the proposition $P(M)$.

How can L fail to be a least upper bound for A? This can happen if L is not an upper bound for A, or if there is an upper bound M for A with $M < L$. More formally, we have

L is not a least upper bound for A

$$\begin{aligned}
&\iff \neg[P(L) \wedge (\forall M \in \mathbb{R},\ P(M) \implies L \leq M)] \\
&\iff \neg P(L) \vee \neg(\forall M \in \mathbb{R},\ P(M) \implies L \leq M) \\
&\iff \neg P(L) \vee (\exists M \in \mathbb{R} \text{ so that } \neg(P(M) \implies L \leq M)) \\
&\iff \neg P(L) \vee (\exists M \in \mathbb{R} \text{ so that } P(M) \wedge \neg(L \leq M) \\
&\iff \neg P(L) \vee (\exists M \in \mathbb{R} \text{ so that } P(M) \wedge (L > M) \\
&\iff [\exists a \in A \text{ so that } a > L] \vee [\exists M \in \mathbb{R} \text{ so that } (\forall a \in A,\ a \leq M) \wedge (L > M)],
\end{aligned}$$

consistent with discussion above. However, if we were to proceed mechanically without thought, we might assert

L is not a least upper bound for A

$$\begin{aligned}
&\iff \neg[(\forall a \in A,\ a \leq L) \wedge (\forall M \in \mathbb{R},\ \forall a \in A,\ a \leq M \implies L \leq M)] \\
&\iff (\exists a \in A \text{ so that } a > L) \vee (\exists M \in \mathbb{R},\ \exists a \in A \text{ so that } (a \leq M) \wedge (L > M)),
\end{aligned}$$

but this last proposition is **not** equivalent to "L is not a least upper bound for A". The problem is that we could interpret "$\forall a \in A,\ a \leq M \implies L \leq M$" as "$\forall a \in A,$ if $a \leq M$ then $L \leq M$," or as "if, $\forall a \in A,\ a \leq M,$ then $L \leq M$." Some texts try to avoid this confusion by writing "$a \leq M\ \forall a \in A \implies L \leq M$," which can only be interpreted as "if, $a \leq M\ \forall a \in A,$ then $L \leq M$."

Example: Let $[0,1) = \{x \in \mathbb{R} :\ 0 \leq x < 1\ \}$ and $(0,1) = \{x \in \mathbb{R} :\ 0 < x < 1\ \}$. Define $f : [0,1) \to (0,1)$ by

$$f(x) = \begin{cases} 1 - \frac{1}{n+1} & \text{if } x = 1 - \frac{1}{n} \text{ for some } n \in \mathbb{Z}_+, \\ x & \text{otherwise.} \end{cases}$$

To understand this definition, we need to understand the condition "otherwise":

$$\begin{aligned}
\neg[x = 1 - \frac{1}{n} \text{ for some } n \in \mathbb{Z}_+] &\iff \neg[\exists n \in \mathbb{Z}_+ \text{ so that } x = 1 - \frac{1}{n}] \\
&\iff [\forall n \in \mathbb{Z}_+,\ \neg(x = 1 - \frac{1}{n})] \\
&\iff [\forall n \in \mathbb{Z}_+,\ x \neq 1 - \frac{1}{n}].
\end{aligned}$$

Contrapositives of propositions with quantifiers. Suppose $P(x), Q(x)$ are propositions involving $x \in X$ where X is some set. We have seen that

$P(x) \implies Q(x)$ is equivalent to its contrapositive: $\neg Q(x) \implies \neg P(x)$. Hence

$[\forall x \in X,\ (P(x) \implies Q(x))]$ is equivalent to $[\forall x \in X,\ (\neg Q(x) \implies \neg P(x))]$.

Similarly,

$[\exists x \in X,\ (P(x) \implies Q(x))]$ is equivalent to $[\exists x \in X,\ (\neg Q(x) \implies \neg P(x))]$.

This analysis extends to implication with multiple quantifiers; in the next proposition we discuss such a situation.

Theorem 3.1. *Suppose $f : X \to Y$. The map f is injective if and only if*

$$[\forall x_1, x_2 \in X,\ f(x_1) = f(x_2) \implies x_1 = x_2].$$

Proof. By definition,

$$[f \text{ is injective }] \iff [\forall x_1, x_2 \in X,\ (x_1 \neq x_2 \implies f(x_1) \neq f(x_2))].$$

The contrapositive of the statement

$$x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$$

is

$$\neg(f(x_1) \neq f(x_2)) \implies \neg(x_1 \neq x_2),$$

or equivalently,

$$f(x_1) = f(x_2) \implies x_2 = x_2.$$

Thus

$$\forall x_1, x_2 \in X,\ x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$$

is equivalent to

$$\forall x_1, x_2 \in X,\ f(x_1) = f(x_2) \implies x_1 = x_2,$$

which proves the proposition. □

Note: With $f : X \to Y$, some texts define f to be injective if:

$$\forall x_1, x_2 \in X, f(x_1) = f(x_2) \implies x_1 = x_2.$$

Since the above statement is equivalent to the definition given in §1, either can be used as the definition of injective. The definition in §1 is meant to capture more obviously that a map f is injective when it maps distinct elements of the domain to distinct elements of the codomain, but the above equivalent statement is often easier to use when proving a map is injective.

4. Set operations

Throughout this section, we rely on basic results from §2.

Suppose that A, B are subsets of some set X.

Recall: $A \cup B$ denotes the union of A and B, meaning

$$A \cup B = \{x \in X :\ x \in A \text{ or } x \in B\ \}.$$

So for $x \in X$, $x \in A \cup B$ if and only if $x \in A \vee x \in B$.
$A \cap B$ denotes the intersection of A and B, meaning

$$A \cap B = \{x \in X :\ x \in A \text{ and } x \in B\ \}.$$

So for $x \in X$, $x \in A \cap B$ if and only if $x \in A \wedge x \in B$. When $A \cap B = \emptyset$ we say A and B are disjoint.
$A \smallsetminus B$ denotes the difference of A and B, meaning

$$A \smallsetminus B = \{x \in X : \ x \in A \text{ and } x \notin B \ \}.$$

So for $x \in X$, $x \in A \smallsetminus B$ if and only if $x \in A \wedge x \notin B$.
A^c denotes the complement of A, meaning

$$A^c = \{x \in X : \ x \notin A \ \}.$$

We have the following simple proposition.

Theorem 4.1. *Let X be a set, and for $x \in X$, let $P(x)$ be the proposition that x satisfies condition P, and let $Q(x)$ be the proposition that x satisfies condition Q. Set*

$$A = \{x \in X : \ P(x) \ \}, \ B = \{x \in X : \ Q(x) \ \}.$$

Then

$$A \cap B = \{x \in X : \ P(x) \wedge Q(x) \ \} \text{ and } A \cup B = \{x \in X : \ P(x) \vee Q(x) \ \}.$$

Proof. For $x \in X$, we have $x \in A$ if and only if $P(x)$; similarly, $x \in B$ if and only if $Q(x)$. Thus

$$\begin{aligned} A \cap B &= \{x \in X : \ x \in A \wedge x \in B \ \} \\ &= \{x \in X : \ P(x) \wedge Q(x) \ \} \end{aligned}$$

and

$$\begin{aligned} A \cup B &= \{x \in X : \ x \in A \vee x \in B \ \} \\ &= \{x \in X : \ P(x) \vee Q(x) \ \}. \end{aligned}$$

□

Proposition 4.2. *Suppose A, B, C are subsets of a set X.*

(a) $A \cap (B \cap C) = (A \cap B) \cap C$.
(a) $A \cup (B \cup C) = (A \cup B) \cup C$.

(Thus the set operations $\cup$ and $\cap$ are associative.)

Proof. We prove (a) and leave (b) as an exercise.

Suppose $x \in X$. Let P be the propostion $x \in A$, Q the proposition $x \in B$, and R the propostion $x \in C$. Recall that $P \wedge (Q \wedge R) \iff (P \wedge Q) \wedge R$. Thus:

$$\begin{aligned} x \in A \cap (B \cap C) &\iff (x \in A) \wedge (x \in B \cap C) \\ &\iff (x \in A) \wedge (x \in B \wedge x \in C) \\ &\iff P \wedge (Q \wedge R) \\ &\iff (P \wedge Q) \wedge R \\ &\iff (x \in A \wedge x \in B) \wedge x \in C \\ &\iff (x \in A \cap B) \wedge (x \in C) \\ &\iff x \in (A \cap B) \cap C. \end{aligned}$$

Thus the elements of X that are in $A \cap (B \cap C)$ are exactly the elements of X that are in $(A \cap B) \cap C$, so $A \cap (B \cap C) = (A \cap B) \cap C$. □

Theorem 4.3. *Let A, B, C be subsets of a set X.*

(a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
(b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Proof. We prove (a) and leave (b) as an exercise.

Suppose $x \in X$. Let P be the propostion $x \in A$, Q the proposition $x \in B$, and R the proposition $x \in C$. Recall that $P \wedge (Q \vee R) \iff (P \wedge Q) \vee (P \wedge R)$. Then:

$$\begin{aligned} x \in A \cap (B \cup C) &\iff x \in A \wedge x \in B \cup C \\ &\iff x \in A \wedge (x \in B \vee x \in C) \\ &\iff P \wedge (Q \vee R) \\ &\iff (P \wedge Q) \vee (P \wedge R) \\ &\iff (x \in A \wedge x \in B) \vee (x \in A \wedge x \in C) \\ &\iff (x \in A \cap B) \vee (x \in A \cap C) \\ &\iff x \in (A \cap B) \cup (A \cap C). \end{aligned}$$

Thus the elements of $A \cap (B \cup C)$ are exactly the elements of $(A \cap B) \cup (A \cap C)$, and hence $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. □

Proposition 4.4. *Suppose A, B are subsets of a set X.*

(a) $A \smallsetminus B = A \cap B^c$.
(b) $(A \smallsetminus B)^c = A^c \cup B$.

Proof. We prove (a) and leave (b) as an exercise.

Suppose $x \in X$; then we have

$$\begin{aligned} x \in A \smallsetminus B &\iff x \in A \wedge x \notin B \\ &\iff x \in A \wedge x \in B^c \\ &\iff x \in A \cap B^c, \end{aligned}$$

Thus the elements of X that are in $A \smallsetminus B$ are exactly the elements of X that are in $A \cap B^c$, so $A \smallsetminus B = A \cap B^c$. □

xs

Theorem 4.5. *(De Morgan's Laws) Suppose A, B, C are subsets of a set X.*

(a) $A \smallsetminus (B \cup C) = (A \smallsetminus B) \cap (A \smallsetminus C)$.
(b) $A \smallsetminus (B \cap C) = (A \smallsetminus B) \cup (A \smallsetminus C)$.
(c) $(A \cap B)^c = A^c \cup B^c$. *(Thus for $x \in X$, $x \notin A \cap B \iff x \notin A \vee x \notin B$.)*
(d) $(A \cup B)^c = A^c \cap B^c$. *(Thus for $x \in X$, $x \notin A \cup B \iff x \notin A \wedge x \notin B$.)*

Proof. We prove (a), (d) and leave (b), (c) as exercises.

(a) Suppose $x \in X$. As an easy exercise using truth tables, one shows that with P, Q, R propositions, $P \wedge (Q \wedge R) \iff (P \wedge Q) \wedge (P \wedge R)$. Thus:

$$\begin{aligned} x \in A \smallsetminus (B \cup C) &\iff (x \in A) \wedge (x \notin B \cup C) \\ &\iff (x \in A) \wedge \neg(x \in B \cup C) \\ &\iff (x \in A) \wedge \neg(x \in B \vee x \in C) \\ &\iff (x \in A) \wedge (x \notin B \wedge x \notin C) \\ &\iff (x \in A \wedge x \notin B) \wedge (x \in A \wedge x \notin C) \\ &\iff (x \in A \smallsetminus B) \wedge (x \in A \smallsetminus C) \\ &\iff x \in (A \smallsetminus B) \cap (A \smallsetminus C). \end{aligned}$$

Thus the elements of $A \smallsetminus (B \cup C)$ and $(A \smallsetminus B) \cap (A \smallsetminus C)$ are the same, meaning $A \smallsetminus (B \cup C) = (A \smallsetminus B) \cap (A \smallsetminus C)$.

(d) Suppose $x \in X$. Thus:

$$\begin{aligned} x \in (A \cup B)^c &\iff \neg(x \in A \cup B) \\ &\iff \neg(x \in A \vee x \in B) \\ &\iff \neg(x \in A) \wedge \neg(x \in B) \\ &\iff x \in A^c \wedge x \in B^c \\ &\iff x \in A^c \cap B^c. \end{aligned}$$

Since $x \in (A \cup B)^c$ if and only if $x \in A^c \cap B^c$, we have $(A \cup B)^c = A^c \cap B^c$. (Note that we have also shown that $x \in (A \cup B)^c \iff x \in A^c \wedge x \in B^c$, so $x \notin A \cup B \iff x \notin A \wedge x \notin B$.)

ALTERNATIVELY: Suppose $x \in X$. Then, using (a) we have

$$\begin{aligned} x \in (A \cup B)^c &\iff x \in X \smallsetminus (A \cup B) \\ &\iff xin[(X \smallsetminus A) \cap (X \smallsetminus B)] \\ &\iff x \in A^c \cap B^c. \end{aligned}$$

Since $x \in (A \cup B)^c$ if and only if $x \in A^c \cap B^c$, we have $(A \cup B)^c = A^c \cap B^c$. $\square$

Notation: It is often convenient to denote the elements of a set using indices, or subscripts. For example, suppose A is a set with 5 elements; we can denote these elements as a_1, a_2, a_3, a_4, a_5. Then we can write

$$A = \{a_i : \ i \in I \ \} \text{ where } I = \{1, 2, 3, 4, 5\};$$

here I is called an indexing set. This notation is particularly useful when dealing with infinite sets. For instance, we will see that there are infinitely many primes within the set of integers; ordering the primes in increasing order, let p_i denote the ith prime where $i \in \mathbb{Z}_+$. Then

$$\{p_i : \ i \in \mathbb{Z}_+ \ \}$$

denotes the set of all primes. Alternatively, we sometimes denote this set with the notation $\{p_i\}_{i \in \mathbb{Z}_+}$.

Let $\{A_i\}_{i \in I}$ be a collection of subsets of a set X where I is an indexing set. Then we write $\cup_{i \in I} A_i$ to denote the union of all the sets A_i, $i \in I$. That is,

$$\cup_{i \in I} A_i = \{x \in X : \ \exists i \in I \text{ so that } x \in A_i \ \}.$$

Somewhat similarly, we write $\cap_{i\in I}A_i$ to denote the intersection of all the sets A_i, $i \in I$. That is,

$$\cap_{i\in I}A_i = \{x \in X : \ \forall i \in I, \ x \in A_i \ \}.$$

Proposition 4.6. *Let X be a set with subset A, and an indexed collection of subsets $\{B_i\}_{i\in I}$, where I is an indexing set. Then we have:*

(a) $A \setminus \cap_{i\in I}B_i = \cup_{i\in I}(A \setminus B_i)$.
(b) $A \setminus \cup_{i\in I}B_i = \cap_{i\in I}(A \setminus B_i)$.

Proof. We prove (a) and leave (b) as an exercise.

We know $x \in \cap_{i\in I}B_i$ if and only if $\forall i \in I$, $x \in B_i$. So $\neg(x \in \cap_{i\in I}B_i)$ if and only if $\exists i \in I$ so that $x \notin B_i$.

Suppose $x \in A \setminus \cap_{i\in I}B_i$. Then $x \in A$, and for some $i \in I$, we have $x \notin B_i$. So for some $i \in I$, $x \in A \setminus B_i$. Thus $x \in \cup_{i\in I}(A \setminus B_i)$. This shows that $A \setminus \cap_{i\in I}B_i \subseteq \cup_{i\in I}(A \setminus B_i)$.

Now suppose that $x \in \cup_{i\in I}(A \setminus B_i)$. Thus for some $i \in I$, we have $x \in A \setminus B_i$. So for some $i \in I$, $x \in A$ and $x \notin B_i$. Since $\exists i \in I$ so that $x \notin B_i$, we have $x \notin \cap_{i\in I}B_i$. Thus $x \in A \setminus \cap_{i\in I}B_i$. This shows that $\cup_{i\in I}(A \setminus B_i) \subseteq A \setminus \cap_{i\in I}B_i$. Together with the result of the preceding paragraph, we get $A \setminus \cap_{i\in I}B_i = \cup_{i\in I}(A \setminus B_i)$. □

Theorem 4.7. *Suppose $f : X \to Y$ and $X = U \cup V$. Then $f(X) = f(U) \cup f(V)$. Further, if f is injective and $U \cap V = \emptyset$, then $f(U) \cap f(V) = \emptyset$.*

Proof. Since $U, V \subseteq X$, clearly $f(U), f(V) \subseteq f(X)$, so $f(U) \cup f(V) \subseteq f(X)$. On the other hand, take $x \in X$. Then $x \in U$ or $x \in V$, so $f(x) \in f(U)$ or $f(x) \in f(V)$. Therefore $f(x) \in f(U) \cup f(V)$; as this holds for all $x \in X$, we have $f(X) \subseteq f(U) \cup f(V)$. Hence $f(X) = f(U) \cup f(V)$.

Now suppose f is injective and $U \cap V = \emptyset$. For the sake of contradiction, suppose there is some $y \in f(U) \cap f(V)$. Thus there is some $u \in U$ so that $y = f(u)$, and there is some $v \in V$ so that $y = f(v)$. Hence $f(u) = y = f(v)$. Since f is injective, we have $u = v$. Hence $u \in U \cap V$ [as $u = v$ and $v \in V$], contradicting the assumption that $U \cap V = \emptyset$. Thus there cannot be any $y \in f(U) \cap f(V)$, meaning $f(U) \cap f(V) = \emptyset$. □

Definition. Suppose $f : X \to Y$, $V \subseteq Y$. We define the inverse image of V under f as

$$f^{-1}(V) = \{x \in X : \ f(x) \in V \ \}.$$

Note that $f^{-1}(\emptyset) = \emptyset$.

Warning: This notation does **not** mean f^{-1} is necessarily a function!

Example: Say $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is defined by $f((x, y)) = 2x - 5y$. Then, from linear algebra, the "kernel" of f is

$$\begin{aligned} f^{-1}(\{0\}) &= \{(x, y) \in \mathbb{R} \times \mathbb{R} : \ f((x, y)) \in \{0\} \ \} \\ &= \{(x, y) \in \mathbb{R} \times \mathbb{R} : \ 2x - 5y = 0 \ \} \\ &= \{(x, y) \in \mathbb{R} \times \mathbb{R} : \ y = 2x/5 \ \} \\ &= \{(x, 2x/5) : \ x \in \mathbb{R} \ \}. \end{aligned}$$

Example: Suppose still that $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is defined by $f(x, y) = 2x - 5y$. Let $V = (0, 1)$, an open interval in $\mathbb{R}$. Then

$$\begin{aligned} f^{-1}(V) &= \{(x, y) \in \mathbb{R} \times \mathbb{R} : \ f((x, y)) \in V \ \} \\ &= \{(x, y) \in \mathbb{R} \times \mathbb{R} : \ 2x - 5y \in (0, 1) \ \} \\ &= \{(x, y) \in \mathbb{R} \times \mathbb{R} : \ 0 < 2x - 5y < 1 \ \} \\ &= \left\{ (x, y) \in \mathbb{R} \times \mathbb{R} : \ \frac{5}{2}y < x < \frac{5}{2}y + \frac{1}{2} \ \right\}. \end{aligned}$$

So we can also describe $f^{-1}(V)$ as

$$f^{-1}(V) = \left\{ \left(\frac{5}{2}y + \varepsilon, y \right) : \ \epsilon, y \in \mathbb{R}, \ 0 < \varepsilon < \frac{1}{2} \ \right\}.$$

Example: Define $g : \mathbb{R} \to \mathbb{R}$ by $g(x) = |x^3|$. Take $V = [4, \infty)$. Then

$$\begin{aligned} g^{-1}(V) &= \{x \in \mathbb{R} : \ g(x) \in V \ \} \\ &= \{x \in \mathbb{R} : \ |x^3| \in [4, \infty) \ \} \\ &= \{x \in \mathbb{R} : \ x^3 \geq 4 \ \vee \ -x^3 \geq 4 \ \} \\ &= \{x \in \mathbb{R} : \ x \geq \sqrt[3]{4} \ \vee \ x \leq \sqrt[3]{-4} \ \} \\ &= (-\infty, \sqrt[3]{-4}] \cup [\sqrt[3]{4}, \infty). \end{aligned}$$

Theorem 4.8. *Let $f : X \to Y$, and let $U \subseteq X$, $V \subseteq Y$. Then we have:*

(a) *$f(f^{-1}(V)) \subseteq V$, and when f is surjective, $f(f^{-1}(V)) = V$.*
(b) *$U \subseteq f^{-1}(f(U))$, and when f is injective, $U = f^{-1}(f(U))$.*

Proof. We prove (a) and leave (b) as an exercise.

If $V = \emptyset$, then $f^{-1}(V) = \emptyset$ and $f(f^{-1}(V)) = \emptyset = V$. So suppose $V \neq \emptyset$.

Choose $y \in f(f^{-1}(V))$. Thus $y = f(w)$ for some $w \in f^{-1}(V)$. By the definition of $f^{-1}(V)$, we have $f(w) \in V$. Hence $y = f(w) \in V$. Since y was chosen arbitrarily from $f(f^{-1}(V))$, this shows that every element of $f(f^{-1}(V))$ lies in V, i.e. $f(f^{-1}(V)) \subseteq V$.

Now suppose f is surjective. We have already established that $f(f^{-1}(V)) \subseteq V$, so to show $f(f^{-1}(V)) = V$, we need to show $V \subseteq f(f^{-1}(V))$. Suppose $v \in V$. Since f is surjective, $\exists x \in X$ so that $f(x) = v$. Thus $f(x) \in V$, so $x \in f^{-1}(V)$. Hence $v = f(x) \in f(f^{-1}(V))$. Since v was chosen arbitrarily from V, this shows $V \subseteq f(f^{-1}(V))$. Since we chose v arbitrarily from V, this shows that $f(f^{-1}(V)) = V$ [under the assumption that f is surjective]. $\square$

Theorem 4.9. *Suppose $f : X \to Y$ and $V_1, V_2 \subseteq Y$. Then*

$$f^{-1}(V_1 \cap V_2) = f^{-1}(V_1) \cap f^{-1}(V_2)$$

and

$$f^{-1}(V_1 \cup V_2) = f^{-1}(V_1) \cup f^{-1}(V_2).$$

Proof. We prove the first statement and leave the second as an exercise.

We have

$$\begin{aligned} f^{-1}(V_1 \cap V_2) &= \{x \in X : \ f(x) \in V_1 \cap V_2 \ \} \\ &= \{x \in X : \ f(x) \in V_1 \wedge f(x) \in V_2 \ \}. \end{aligned}$$

On the other hand,

$$\begin{aligned} f^{-1}(V_1) \cap f^{-1}(V_2) &= \{x \in X : \ f(x) \in V_1 \ \} \cap \{x \in X : \ f(x) \in V_2 \ \} \\ &= \{x \in X : \ f(x) \in V_1 \wedge f(x) \in V_2 \ \}. \end{aligned}$$

Therefore $f^{-1}(V_1 \cap V_2) = f^{-1}(V_1) \cap f^{-1}(V_2)$.

Alternatively, one could present this argument as follows:

$$\begin{aligned} f^{-1}(V_1 \cap V_2) &= \{x \in X : \ f(x) \in V_1 \cap V_2 \ \} \\ &= \{x \in X : \ f(x) \in V_1 \wedge f(x) \in V_2 \ \} \\ &= \{x \in X : \ f(x) \in V_1 \ \} \cap \{x \in X : \ f(x) \in V_2 \ \} \\ &= f^{-1}(V_1) \cap f^{-1}(V_2). \end{aligned}$$

□

5. Partitioning sets, equivalence relations, and congruences

According to standard usage of English, partitioning a set means we break it into non-overlapping pieces. More precisely, we have the following.

Definition. A partition of a nonempty set X is a collection $\{A_i : \ i \in I \ \}$ of nonempty subsets of X so that

(1) $\forall x \in X$, $\exists i \in I$ so that $x \in A_i$;
(2) $\forall x \in X$, $\forall i, j \in I$, if $x \in A_i \wedge x \in A_j$ then $A_i = A_j$.

(In some texts the subsets A_i are called blocks of the partition.)

Example: Let $X = \{1, 2, 3, 4, 5, 6\}$. Then

$$\{\{1\}, \ \{2, 3\}, \ \{4, 5, 6\}\}$$

is a partition of X. Another partition of X is

$$\{\{1, 2, 3\}, \ \{4, 6\}, \ \{5\}\}.$$

Partitions of sets are inextricably linked to "equivalence relations"; to define these, we first need some other definitions.

Definitions. A relation $\sim$ on a nonempty set X corresponds to a subset $R_\sim$ of $X \times X$; we write $x \sim y$ when $(x, y) \in R_\sim$, and we say x is related to y. Given a relation $\sim$ on X, we say:

(1) $\sim$ is reflexive if: $\forall x \in X$, we have $x \sim x$;
(2) $\sim$ is symmetric if: $\forall x, y \in X$, $x \sim y \implies y \sim x$;
(3) $\sim$ is transitive if: $\forall x, y, z \in X$, $(x \sim y \wedge y \sim z) \implies x \sim z$.

A relation is an equivalence relation if it is reflexive, symmetric, and transitive.

Example: Let T be the set of all triangles in $\mathbb{R} \times \mathbb{R}$. For $t_1, t_2 \in T$, consider the following relation: $t_1 \sim t_2$ if t_1 is similar to t_2 (meaning there is a correspondence between the interior angles of t_1 and the interior angles of t_2 so that corresponding angles are equal). Then $\sim$ is an equivalence relation (check!).

Example: Let $X = \mathbb{Z}$, and let $R_\sim = \{(x, x) : \ x \in \mathbb{Z} \ \}$. So $\forall x \in \mathbb{Z}$, $x \sim x$ (so $\sim$ is reflexive). We claim that $\sim$ is an equivalence relation on

$\mathbb{Z}$: We already noted $\sim$ is reflexive. Suppose $x, y \in \mathbb{Z}$ so that $x \sim y$. Thus $(x, y) \in R_\sim$, so $x = y$. Hence $(y, x) = (x, x) \in R_\sim$, so $y \sim x$. Thus $\sim$ is symmetric. Suppose $x, y, z \in \mathbb{Z}$ so that $x \sim y$ and $y \sim z$. Thus $x = y$, and $y = z$, so $x = y = z$. Hence $(x, z) = (x, x) \in R_\sim$, so $x \sim z$. Thus $\sim$ is transitive. So $\sim$ is an equivalence relation.

Note: If $\sim$ is an equivalence relation on some nonempty set X, then we necessarily have

$$\{(x, x) : \ x \in X \ \} \subseteq R_\sim$$

since $\sim$ is reflexive.

Example: Define a relation $\sim$ on $\mathbb{Z}$ by $x \sim y$ if $x < y$. So $\sim$ is not reflexive, as there are $x \in \mathbb{Z}$ so that $\neg(x \sim x)$; in particular, $1 \in \mathbb{Z}$ and $\neg(1 < 1)$ so $\neg(1 \sim 1)$. Also, $\sim$ is not symmetric, as there are $x, y \in \mathbb{Z}$ so that $x \sim y$ but $\neg(y \sim x)$; in particular, $2, 3 \in \mathbb{Z}$ and $2 < 3$ so $2 \sim 3$, but $\neg(3 < 2)$ so $\neg(3 \sim 2)$. However, $\sim$ is transitive: Suppose $x, y, z \in \mathbb{Z}$ so that $x \sim y$ and $y \sim z$. Thus $x < y$ and $y < z$, so $x < y < z$. Hence $x < z$, so $x \sim z$.

Definition. Suppose $\sim$ is an equivalence relation on a (nonempty) set X. For $x \in X$, we define

$$[x]_\sim = \{y \in X : \ y \sim x \ \},$$

and we call $[x]_\sim$ the equivalence class of x (relative to the relation $\sim$).

Proposition 5.1. *Suppose $\sim$ is an equivalence relation on a (nonempty) set X. For any $x, y \in X$, $[x]_\sim \neq [y]_\sim$ if and only if $[x]_\sim \cap [y]_\sim = \emptyset$.*

Proof. To ease notation, for $x \in X$ let us temporarily write $[x]$ for $[x]_\sim$.

Take $x, y \in X$. We need to prove

(1) $[x] \neq [y] \implies [x] \cap [y] = \emptyset$, and

(2) $[x] \cap [y] = \emptyset \implies [x] \neq [y]$.

To do this, we will prove the contrapositive of each statement:

(1) $[x] \cap [y] \neq \emptyset \implies [x] = [y]$, and

(2) $[x] = [y] \implies [x] \cap [y] \neq \emptyset$.

To prove (1): Suppose $[x] \cap [y] \neq \emptyset$. Thus there is some $z \in [x] \cap [y]$. Hence $z \in [x]$, so $z \sim x$; similarly, $z \in [y]$, so $z \sim y$. Since $\sim$ is symmetric, we have $x \sim z$; since $\sim$ is transitive, we have $x \sim y$. Now choose $w \in [x]$; thus $w \sim x$, and since $x \sim y$ and $\sim$ is transitive, $w \sim y$. Hence $w \in [y]$; as this holds for all $w \in [x]$, we have $[x] \subseteq [y]$. A virtually identical argument shows that for any $w \in [y]$ we have $w \in [x]$, so $[y] \subseteq [x]$. Hence $[x] = [y]$.

To prove (2): Suppose $[x] = [y]$. We know $x \in [x]$ as $\sim$ is reflexive and so $x \sim x$. Hence $x \in [x] = [x] \cap [y]$, so $[x] \cap [y] \neq \emptyset$. □

Theorem 5.2. *Suppose $\sim$ is an equivalence relation on a (nonempty) set X. Then*

$$\Pi = \{[x]_\sim : \ x \in X \ \}$$

is a partition of X.

Proof. To ease notation, for $x \in X$ let us temporarily write $[x]$ for $[x]_\sim$.

Take $a \in X$. Then $[a] \in \Pi$; hence every element of X is in one of the sets in Π.

Now suppose that for $a \in X$, we have $a \in [x]$ and $a \in [y]$ where $x, y \in X$. Then $[x] \cap [y] \neq \emptyset$, so by the preceding proposition we have $[x] = [y]$. Thus Π is a partition of X. □

On the other hand, we have the following.

Theorem 5.3. *Suppose $\Pi = \{A_i : i \in I\}$ is a partition of a (nonempty) set X (so I is an indexing set). For $x, y \in X$, define $x \sim y$ if $\exists i \in I$ so that $x, y \in A_i$. Then $\sim$ is an equivalence relation on X.*

Proof. We first show $\sim$ is reflexive: Take $x \in X$. Since Π is a partition of X, there is some $i \in I$ so that $x \in A_i$. Thus $x \sim x$.

Next we show $\sim$ is symmetric: Suppose $x, y \in X$ so that $x \sim y$. Thus there is some $i \in I$ so that $x, y \in A_i$. Hence $y, x \in A_i$, so $y \sim x$.

Finally, we show $\sim$ is transitive: Suppose $x, y, z \in X$ so that $x \sim y$ and $y \sim z$. Thus there is some $i \in I$ so that $x, y \in A_i$ and some $j \in I$ so that $y, z \in A_j$. Hence $y \in A_i$ and $y \in A_j$; since Π is a partition, we must have $A_i = A_j$. Thus $x, z \in A_i$ so $x \sim z$.

This shows $\sim$ is an equivalence relation on X. □

Congruences. Here we present an explicit and fundamental example of an equivalence relation on $\mathbb{Z}$. We begin with a familiar definition.

Defintion. For $x, y \in \mathbb{Z}$, we say x divides y if $\exists z \in \mathbb{Z}$ so that $y = xz$. We write $x|y$ to denote "x divides y". Similarly, we write $x \nmid y$ to denote "x does not divide y", meaning that $\forall z \in \mathbb{Z},\ y \neq xz$.

Fix $n \in \mathbb{Z}_+$. We define a relation on $\mathbb{Z}$ as follows: For $a, b \in \mathbb{Z}$, we write $a \equiv b \pmod{n}$ if $n|a - b$. When $a \equiv b \pmod{n}$, we say a is congruent to b modulo n. We leave it as an exercise to show that this relation is in fact an equivalence relation on $\mathbb{Z}$. So when $a, b \in \mathbb{Z}$ and $n \in \mathbb{Z}_+$ with $a \equiv b \pmod{n}$, then a and b are in the same congruence class modulo n.

This is a particularly interesting equivalence relation because of the following.

Theorem 5.4. *Fix $n \in \mathbb{Z}_+$. Suppose $a, b, c, d \in \mathbb{Z}$ so that $a \equiv c \ (mod\ n)$, $b \equiv d \ (mod\ n)$. Then*

$$a + b \equiv c + d \ (mod\ n),\ ab \equiv cd \ (mod\ n).$$

Proof. By assumption, we have $n|a - c$ and $n|b - d$. Thus for some $x, y \in \mathbb{Z}$, we have $a - c = nx$ and $b - d = ny$. Hence

$$(a + b) - (c + d) = (a - c) + (b - d) = nx + ny = n(x + y).$$

Since $x + y \in \mathbb{Z}$, this means $n|(a + b) - (c + d)$, so $a + b \equiv c + d \pmod{n}$. Also, since $a = c + nx$ and $b = d + ny$, we have

$$ab = (c + nx)(d + ny) = cd + n(cy + dx + nxy)$$

and hence $ab - cd = n(cy + dx + nxy)$. Since $cy + dx + nxy \in \mathbb{Z}$, we have $n|ab - cd$, so $ab \equiv cd \pmod{n}$. □

This result helps simplify many computations modulo a positive integer n.

Example: We compute $3^5 + 2^8 \pmod 7$ without working unnecessarily hard.

We have $3^2 \equiv 9 \equiv 2 \pmod 7$. So $3^4 = 3^2 \cdot 3^2 \equiv 2 \cdot 2 \equiv 4 \pmod 7$. Hence

$$3^5 \equiv 3^4 \cdot 3 \equiv 12 \equiv 5 \pmod 7.$$

Somewhat similarly, $2^3 \equiv 8 \equiv 1 \pmod 7$, so

$$2^6 \equiv 2^3 \cdot 2^3 \equiv 1 \cdot 1 \equiv 1 \pmod 7.$$

So $2^8 \equiv 2^6 \cdot 2^2 \equiv 1 \cdot 4 \equiv 4 \pmod 7$. Hence

$$3^5 + 2^8 \equiv 5 + 4 \equiv 2 \pmod 7.$$

6. Algorithms, recursion, and mathematical induction

An algorithm is a logical step-by-step procedure for solving a problem in a finite number of steps. Many algorithms are recursive, meaning that after one or more initial steps, a general method is given for determining each subsequent step on the basis of steps already taken.

As an example of a recursive algorithm, we discuss Euclid's algorithm for finding the highest common factor of two nonzero integers.

First, recall that we have a "division algorithm" for $a, b \in \mathbb{Z}_+$:

Theorem 6.1. *Suppose $a, b \in \mathbb{Z}_+$. Then $\exists! q, r \in \mathbb{Z}$ so that $b = aq + r$ where $0 \leq r < a$.*

Proof. Consider the set $A = \{u \in \mathbb{Z} : \ au \leq b \ \}$. Since $0 \in A$, we know A is nonempty, and since $b \leq ab$, A is bounded above; hence we can choose q to be the maximal element in A. [Thus q is the largest integer so that $aq \leq b$.] Set $r = b - aq$. So $b = aq + r$ with $0 \leq r < a$. [If $r \geq a$, then we would have $a(q+1) \leq b$, contrary to our choice of q.] Note also that q, r are the unique integers so that $b = aq + r$ with $0 \leq r < a$. To see this, suppose $q', r' \in \mathbb{Z}$ so that

$$b = aq' + r' \text{ with } 0 \leq r' < a.$$

Thus $aq + r = aq' + r'$, so $a(q - q') = r' - r$. Since $0 \leq r < a$ and $0 \leq r' < a$, we have $-a < r' - r < a$. Note that 0 is the only integer strictly between $-a$ and a that is divisible by a. Since $q - q'$ is an integer with $a(q - q') = r' - r$, we must have $r' - r = 0$ and $q - q' = 0$, meaning $r' = r$ and $q' = q$. Hence there are unique $q, r \in \mathbb{Z}$ so that $b = aq + r$ with $0 \leq r < a$. □

Note: An immediate consequence is that with $n \in \mathbb{Z}_+$, $\forall b \in \mathbb{Z}$, $\exists! r \in \mathbb{Z}$ so that $b \equiv r \pmod n$ with $0 \leq r < n$. Hence $\mathbb{Z}$ is partitioned into n congruence classes modulo n. For $n \geq 3$, these congruence classes are

$$\{a \in \mathbb{Z} : \ a \equiv 0 \pmod n \ \},$$
$$\{a \in \mathbb{Z} : \ a \equiv 1 \pmod n \ \},$$
$$\{a \in \mathbb{Z} : \ a \equiv 2 \pmod n \ \},$$
$$\{a \in \mathbb{Z} : \ a \equiv 3 \pmod n \ \},$$
$$\vdots$$
$$\{a \in \mathbb{Z} : \ a \equiv n - 1 \pmod n \ \}.$$

Equivalently, for $n \geq 3$ these congruence classes are

$$n\mathbb{Z},\ 1+n\mathbb{Z},\ 2+n\mathbb{Z},\ 3+n\mathbb{Z}, \ldots,\ (n-1)+n\mathbb{Z}.$$

As an exercise, one proves the following.

Proposition 6.2. *Suppose $a, b, c \in \mathbb{Z}_+$. Then $\exists! q, r \in \mathbb{Z}$ so that $b = aq + r$ with $c \leq r < a + c$.*

Remark: We can extend the division algorithm to show that for $a, b \in \mathbb{Z}$ with $a \neq 0$, there exist unique $q, r \in \mathbb{Z}$ so that $b = aq + r$ with $0 \leq r < |a|$.

Definitions. With $a, b, c \in \mathbb{Z}$, c is a common divisor of a and b if $c|a$ and $c|b$. Note that 1 is always a common divisor of a and b, and if $a \neq 0$, no integer larger than $|a|$ can be a common divisor of a and b. Also note that every $x \in \mathbb{Z}$ is a divisor of 0, as $0 = 0 \cdot x$. With $a, b \in \mathbb{Z}$, a, b not both 0, we write $\text{hcf}(a, b)$ (or equivalently, $gcd(a, b)$) to denote the highest common factor (or equivalently, greatest common divisor) of a and b, meaning $\text{hcf}(a, b)$ is the largest common divisor of a and b.

(For $a, b \in \mathbb{Z}$, not both 0, let C be the set of common divisors of a and b that are positive. So

$$C = \{d \in \mathbb{Z}_+ :\ d|a \text{ and } d|b\ \}.$$

$C \neq \emptyset$ since $1 \in C$. Let M be the maximum of $|a|$ and $|b|$. Then no integer larger than M is a common divisor of a and b, so C is bounded above by M. Thus C has a maximal element, and this is $\text{hcf}(a, b)$, which is positive.)

When $\text{hcf}(a, b) = 1$, we say a, b are relatively prime.

Note that $\text{hcf}(0, 0)$ does not exist, since every integer is a divisor of 0 (for $x \in \mathbb{Z}$, $0 = x \cdot 0$ so $x|0$). For $a \in \mathbb{Z}$ with $a \neq 0$, $\text{hcf}(a, 0) = |a|$.

As an exercise, one proves the following.

Proposition 6.3. *Suppose $a, b \in \mathbb{Z}_+$ and $c = hcf(a, b)$. Take $x, y \in \mathbb{Z}$ so that $a = cx$, $b = cy$. Then $hcf(x, y) = 1$.*

Theorem 6.4. *Take $a, b \in \mathbb{Z}$ so that a, b are not both 0, and let $c = hcf(a, b)$. Then there exist $s, t \in \mathbb{Z}$ so that $c = as + bt$.*

Proof. Let d be the minimum value in the set

$$A = \{au + bv :\ (u, v \in \mathbb{Z}) \wedge (au + bv > 0)\ \}.$$

(One checks that this subset of $\mathbb{Z}$ is nonempty, and it is bounded below by 0, so it has a minimum value.) Take $s, t \in \mathbb{Z}$ so that $d = as + bt$. Note that $c|d$ since $c|a$ and $c|b$ [so $a = cx$, $b = cy$ for some $x, y \in \mathbb{Z}$, and thus $d = as + bt = c(xs + yt)$]. Hence $c \leq d$.

Take $q, r \in \mathbb{Z}$ so that $a = dq + r$ with $0 \leq r < d$. So

$$r = a - dq = a - (as + bt)q = a(1 - sq) + b(-tq).$$

If $r > 0$ then $r \in A$ with $r < d$, contrary to how we chose d. Hence we must have $r = 0$, which means $d|a$. A virtually identical argument shows that $d|b$, so d is a common divisor of a and b. As $c = \text{hcf}(a, b)$, we have $d \leq c$.

Hence $c \leq d$ and $d \leq c$, which means $c = d$. So $\text{hcf}(a, b) = c = d = as + bt$. □

Remark: Suppose $c = \text{hcf}(a, b) = as + bt$ where $a, b, s, t \in \mathbb{Z}$ with a, b not both 0. Thus $\exists a', b' \in \mathbb{Z}$ so that $a = ca'$, $b = cb'$. Then one can show that for any $k \in \mathbb{Z}$, we have $c = a(s + b'k) + b(t - a'k)$, and if $s, t \in \mathbb{Z}$ so that $as' + bt' = c$ then $s' = s + b'k$, $t' = t - a'k$ for some $k \in \mathbb{Z}$.

As an exercise, one proves the following.

Proposition 6.5. *Suppose $a, b, c \in \mathbb{Z}$ so that $c \neq 0$, $c|ab$, and $hcf(b, c) = 1$. Then $c|a$.*

Note that for $a, b \in \mathbb{Z}$, a, b not both 0, this proof shows the existence of some $s, t \in \mathbb{Z}$ so that $\text{hcf}(a, b) = as + bt$, but it does not tell us the actual values of s and t. Euclid's algorithm will produce values such values s, t; to help us prove this, we need the following, whose proof is left as an exercise.

Proposition 6.6. *Suppose $a, b, x \in \mathbb{Z}$, with a, b not both 0. Then $hcf(|a|, |b|) = hcf(a, b) = hcf(b, a + bx)$.*

Euclid's algorithm. Take $a, b \in \mathbb{Z}$ with $b \neq 0$; we will first compute $\text{hcf}(a, b)$ and then we will construct $s, t \in \mathbb{Z}$ so that $\text{hcf}(a, b) = as + bt$.

Step 1: Choose $q_1, r_1 \in \mathbb{Z}$ so that $a = bq_1 + r_1$ with $0 \leq r_1 \leq |b|$. If $r_1 = 0$ then we stop; otherwise we continue.

Step 2: Choose $q_2, r_2 \in \mathbb{Z}$ so that $b = r_1q_2 + r_2$ with $0 \leq r_2 < r_1$. If $r_2 = 0$ then we stop; otherwise we continue.

Step k ($k \geq 3$): Choose $q_k, r_k \in \mathbb{Z}$ so that $r_{k-2} = r_{k-1}q_k + r_k$ with $0 \leq r_k < r_{k-1}$. If $r_k = 0$ then we stop; otherwise we continue.

Notice that after k steps, we have $|b| > r_1 > r_2 > \cdots > r_k \geq 0$. Thus after at most $|b|$ steps, the algorithm must terminate.

If the algorithm terminates after 1 step, then $\text{hcf}(a, b) = |b|$, and we know

$$|b| = \begin{cases} a \cdot 0 + b \cdot 1 & \text{if } b > 0, \\ a \cdot 0 + b \cdot (-1) & \text{if } b < 0. \end{cases}$$

So suppose the algorithm terminates after n steps where $n > 1$; we claim that $r_{n-1} = \text{hcf}(a, b)$. To see this, first note that $r_1 = a - bq_1$, $r_2 = b - r_1q_2$, and for $3 \leq k < n$, we have $r_k = r_{k-2} - r_{k-1}q_k$. Then the preceding proposition tells us

$$\text{hcf}(a, b) = \text{hcf}(b, r_1) = \text{hcf}(r_1, r_2) = \cdots = \text{hcf}(r_{n-1}, r_n).$$

Since the algorithm terminates after n steps, this means $r_{n-1} > 0$ but $r_n = 0$; hence $\text{hcf}(r_{n-1}, r_n) = \text{hcf}(r_{n-1}, 0) = r_{n-1}$.

To realise r_{n-1} as $as + bt$, we substitute, using the equalities that $r_k = r_{k-2} - r_{k-1}q_k$ for $3 \leq k < n$, $r_2 = b - r_1q_2$, and $r_1 = a - bq_1$.

Example: We compute $\text{hcf}(1451, 323)$ and find $s, t \in \mathbb{Z}$ so that $\text{hcf}(1451, 323) = 1451s + 323t$.

Step 1: $1451 = 323 \cdot 4 + 159$ (so $q_1 = 4$, $r_1 = 159$).

Step 2: $323 = 159 \cdot 2 + 5$ (so $q_2 = 2$, $r_2 = 5$).

Step 3: $159 = 5 \cdot 31 + 4$ (so $q_3 = 31$, $r_3 = 4$).

Step 4: $5 = 4 \cdot 1 + 1$ (so $q_4 = 1$, $r_4 = 1$).

Step 5: $4 = 1 \cdot 4 + 0$ (so $q_5 = 4$, $r_5 = 0$).

Hence $\text{hcf}(1451, 323) = r_4 = 1$.

Solving the above equations for r_4, r_3, r_2, r_1 gives us:

$$\begin{aligned} 1 &= 5 - 4 \cdot 1, \\ 4 &= 159 - 5 \cdot 31, \\ 5 &= 323 - 159 \cdot 2, \\ 159 &= 1451 - 323 \cdot 4. \end{aligned}$$

Thus

$$\begin{aligned} 1 &= 5 - (159 - 5 \cdot 31) \cdot 1 \\ &= 5 \cdot 32 - 159 \cdot 1 \\ &= (323 - 159 \cdot 2) \cdot 32 - 159 \cdot 1 \\ &= 323 \cdot 32 - 159 \cdot 65 \\ &= 323 \cdot 32 - (1451 - 323 \cdot 4) \cdot 65 \\ &= 323 \cdot 292 - 1451 \cdot 65. \end{aligned}$$

(So $1 = \text{hcf}(1451, 323) = 1451s + 323t$ where $s = -65$, $t = 292$.)

Remark: As a later exercise, one shows that for $x, y \in \mathbb{Z}$ with $x, y \neq 0$ and $\text{hcf}(x, y) = 1$, there are infinitely many ways to choose $u, v \in \mathbb{Z}$ so that $xu + yv = 1$. Recall that with $c = \text{hcf}(a, b)$ we have $a = cx$, $b = cy$ where $x, y \in \mathbb{Z}$ with $\text{hcf}(x, y) = 1$. Consequently for any $a, b \in \mathbb{Z}$ with $a, b \neq 0$, there are infinitely many ways to choose $s, t \in \mathbb{Z}$ so that $as + bt = \text{hcf}(a, b)$.

As an application of Euclid's algorithm, we prove the following.

Theorem 6.7. *(Chinese Remainder Theorem) Suppose $m, n \in \mathbb{Z}+$ with hcf$(m, n) = 1$. For any $a, b \in \mathbb{Z}$, there is some $x \in \mathbb{Z}$ so that*

$$x \equiv a \ (mod\ m) \text{ and } x \equiv b \ (mod\ n).$$

Further, for $x' \in \mathbb{Z}$, we have $x' \equiv a \ (mod\ m)$ and $x' \equiv b \ (mod\ n)$ if and only if $x' \equiv x \ (mod\ mn)$.

Proof. Since $\text{hcf}(m, n) = 1$, there exist $s, t \in \mathbb{Z}$ so that $ms + nt = 1$. Thus

$$1 \equiv ms + nt \equiv nt \ (\text{mod } m)$$

and

$$1 \equiv ms + nt \equiv ms \ (\text{mod } n).$$

Take $x = msb + nta$. Then

$$x \equiv nta \equiv 1 \cdot a \equiv a \ (\text{mod } m)$$

and

$$x \equiv msb \equiv 1 \cdot b \equiv b \ (\text{mod } m).$$

We leave it as an exercise to show that for $x' \in \mathbb{Z}$, we have

$$x' \equiv a \ (\text{mod } m) \text{ and } x' \equiv b \ (\text{mod } n) \text{ if and only if } x' \equiv x \ (\text{mod } mn).$$

□

Mathematical induction. Mathematical induction is a method of proof wherein we show the smallest instance of a given proposition is true, and from that deduce that each successive instance of the given proposition is

true. Thus we establish a base case, or some base cases, then set up a recursive process to establish the succeeding cases.

More formally, suppose $P(n)$ is the proposition that the integer n has property P. To prove that $P(n)$ holds for all $n \in \mathbb{Z}_+$ using induction, we first prove $P(1)$ holds (this is called the base case). Then we show that for any $k \in \mathbb{Z}_+$, $P(k) \implies P(k+1)$ (this is called the induction step); to do this, one supposes that $P(k)$ holds (called the induction hypothesis), and then argues that this implies $P(k+1)$ must hold. Hence for any $n \in \mathbb{Z}$ with $n > 1$, this second step shows that $P(1) \implies P(2)$, $P(2) \implies P(3)$, ..., $P(n-1) \implies P(n)$. Having established that $P(1)$ holds, $P(1) \implies P(2)$ shows that $P(2)$ holds; then $P(2) \implies P(3)$ shows that $P(3)$ holds; and so on. A proof using induction to show that $P(n)$ holds for all $n \in \mathbb{Z}_+$ is called a proof by induction on n.

Remarks:

(1) Proving $P(k) \implies P(k+1)$ for all $k \in \mathbb{Z}_+$ does not allow us to conclude $P(n)$ holds for some $n \in \mathbb{Z}_+$ unless we have established that $P(m)$ holds for some $m \in \mathbb{Z}_+$ with $m < n$.

(2) An induction argument gives us an algorithm, which we can only apply finitely many times. Hence if $P(n)$ is a proposition that states that the integer n has property P where we begin by showing $P(1)$ holds, the induction step $P(k) \implies P(k+1)$ does not allow us to conclude that property $P(\infty)$ holds. For example, consider the proposition $P(n)$ that says "for any subset A of $\mathbb{Z}$ with n elements, A has a maximal element"; while this proposition is true for any subset A of $\mathbb{Z}$ where A has finitely many elements, this proposition clearly does not hold for $A = \mathbb{Z}_+$. (Curious students may want to read other sources about a proof technique called "transfinite induction", which is beyond the scope of this course.)

(3) A proof by induction on n does not need to begin by establishing $P(1)$. More generally, if we establish that $P(n_0)$ holds for some (fixed) $n_0 \in \mathbb{Z}$, and that $P(k) \implies P(k+1)$ for any $k \in \mathbb{Z}$ with $k \geq n_0$, then the principle of mathematical induction allows us to conclude that $P(n)$ holds for all $n \in \mathbb{Z}$ with $n \geq n_0$.

We now present some examples of proofs by induction.

Proposition 6.8. *For every $n \in \mathbb{Z}_+$,*

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}.$$

Proof. For $n \in \mathbb{Z}_+$, let $P(n)$ be the proposition that

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}.$$

(Base case:) We have $1 = \frac{1(1+1)}{2}$, so $P(1)$ holds.

(Induction step:) Suppose $k \geq 1$ and $P(k)$ holds; recall that $P(k)$ is the proposition that

$$1 + 2 + 3 + \cdots + k = \frac{k(k+1)}{2}.$$

[We need to deduce that $P(k+1)$ holds.] Then

$$\begin{aligned} 1+2+3+\cdots+k+(k+1) &= \frac{k(k+1)}{2}+(k+1) \\ &= \frac{k(k+1)}{2}+\frac{2(k+1)}{2} \\ &= \frac{k^2+3k+2}{2} \\ &= \frac{(k+1)(k+2)}{2}. \end{aligned}$$

Hence $P(k+1)$ holds if $P(k)$ holds, or equivalently, $P(k) \implies P(k+1)$.

By the principle of mathematical induction, this shows that for every $n \in \mathbb{Z}_+$, $P(n)$ holds. □

Proposition 6.9. *Let X be a set, and let $A, B_1, B_2, \ldots, B_n \subseteq X$ where $n \in \mathbb{Z}_+$. Then we have the following results.*

(a) $A \cup (B_1 \cap B_2 \cap \cdots \cap B_n) = (A \cup B_1) \cap (A \cup B_2) \cap \cdots \cap (A \cup B_n)$.
(b) $A \cap (B_1 \cup B_2 \cup \cdots \cup B_n) = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_n)$.
(c) *For $n \geq 2$,* $(B_1 \cap B_2 \cap \cdots \cap B_n)^c = B_1^c \cup B_2^c \cup \cdots \cup B_n^c$.
(d) *For $n \geq 2$,* $(B_1 \cup B_2 \cup \cdots \cup B_n)^c = B_1^c \cap B_2^c \cap \cdots \cap B_n^c$.

Proof. We prove (a) and leave (b), (c), (d) as exercises.

(Base case:) First note that $A \cup B_1 = A \cup B_1$.

(Induction step:) Now suppose that $k \geq 1$ and that $A \cup (B_1 \cap B_2 \cap \cdots \cap B_k) = (A \cup B_1) \cap (A \cup B_2) \cap \cdots \cap (A \cup B_k)$. Let $C = B_1 \cap B_2 \cap \cdots \cap B_k$. Then

$$A \cup (B_1 \cap B_2 \cap \cdots \cap B_{k+1}) = A \cup (C \cap B_{k+1}).$$

By Proposition 4.3, $A \cup (C \cap B_{k+1}) = (A \cup C) \cap (A \cup B_{k+1})$. By our induction hypothesis,

$$\begin{aligned} A \cup C &= A \cup (B_1 \cap B_2 \cap \cdots \cap B_k) \\ &= (A \cup B_1) \cap (A \cup B_2) \cap \cdots \cap (A \cup B_k). \end{aligned}$$

Hence

$$\begin{aligned} &A \cup (B_1 \cap B_2 \cap \cdots \cap B_{k+1}) \\ &= A \cup (C \cap B_{k+1}) \\ &= (A \cup B_1) \cap (A \cup B_2) \cap \cdots \cap (A \cup B_k) \cap (A \cup B_{k+1}). \end{aligned}$$

Thus by the priniciple of mathematical induction, (a) holds for all $n \in \mathbb{Z}_+$. □

7. Strong induction and the Fundamental Theorem of Arithmetic

An argument by strong induction proceeds as follows. Suppose $P(n)$ is the proposition that the integer n has property P. Fix $n_0 \in \mathbb{Z}$. To prove that $P(n)$ holds for all $n \in \mathbb{Z}$ with $n \geq n_0$ using strong induction, we first establish that $P(n_0)$ holds, and then we show that for $k \in \mathbb{Z}$ with $k \geq n_0$,

$$[P(n_0) \wedge P(n_0+1) \wedge \cdots \wedge P(k)] \implies P(k+1).$$

As an example, we will prove the Fundamental Theorem of Arithmetic, which states that for $n \in \mathbb{Z}$ with $n > 1$, n can be written uniquely as a product of primes.

Before we can prove the Fundamental Theorem of Arithmetic, we need to establish some other basic results.

Definition. We say an integer p is prime if $p > 1$, and the only positive divisors of p are 1 and p.

Remark: Later we will see that there are infinitely many primes.

Proposition 7.1. *Suppose $q_1, \ldots, q_r \in \mathbb{Z}$ where $r \in \mathbb{Z}$ with $r \geq 2$, and suppose p is a prime so that $p|q_1 \cdots q_r$. Then for some $i \in \mathbb{Z}$ with $1 \leq i \leq r$, we have $p|q_i$.*

Proof. We proceed by induction on r.

[Base case] Suppose that $p|q_1q_2$. If $p|q_1$ then we are done. So suppose $p \nmid q_1$. Thus $\text{hcf}(p, q_1) = 1$ [since the only positive divisors of p are 1 and p, and p is not a common factor of p and q_1]. Hence by Proposition 6.5, $p|q_2$.

[Induction step] Now suppose $k \in \mathbb{Z}$ with $k \geq 2$, and suppose that if $p|q_1 \cdots q_k$ where $q_1, \ldots, q_k \in \mathbb{Z}$, then $p|q_i$ for some $i \in \mathbb{Z}$ with $1 \leq i \leq k$ [this is the induction hypothesis]. Suppose $q_1, \ldots, q_k, q_{k+1} \in \mathbb{Z}$ with $p|q_1 \cdots q_kq_{k+1}$. Set $t = q_1 \cdots q_k$. Thus $p|tq_{k+1}$, so by the base case in this argument, $p|t$ or $p|q_{k+1}$. If $p|t$ then our induction hypothesis tells us that $p|q_i$ for some $i \in \mathbb{Z}$ with $1 \leq i \leq k$. If $p \nmid t$ then $p|q_{k+1}$. Hence $p|q_i$ for some $i \in \mathbb{Z}$ with $1 \leq i \leq k + 1$.

Thus by the principle of mathematical induction, the proposition is proved. □

Theorem 7.2. *(Fundamental Theorem of Arithmetic) For every $n \in \mathbb{Z}$ so that $n > 1$, we have $n = p_1p_2 \cdots p_r$ for some primes $p_1, p_2, \ldots, p_r$ with $p_1 \leq p_2 \leq \cdots \leq p_r$. Further, if we also have $n = q_1q_2 \cdots q_s$ for primes $q_1, q_2, \ldots, q_s$ with $q_1 \leq q_2 \leq \cdots \leq q_s$, we have $r = s$ and $p_i = q_i$ for all $i \in \mathbb{Z}$ with $1 \leq i \leq r$.*

Proof. We first argue by strong induction to show that each integer $n > 1$ is a product of primes.

First, we note that 2 is a prime, so 2 is a product of primes (where there is only one prime in this product).

Now suppose that $k \in \mathbb{Z}$ with $k \geq 2$, and suppose that for all integers $m \in \mathbb{Z}$ with $2 \leq m \leq k$, m is a product of primes. Consider the integer $k+1$. If $k+1$ is prime, then we are done. So suppose $k+1$ is not prime; thus 1 and $k + 1$ are not the only positive integers dividing $k + 1$. Hence there is some $a \in \mathbb{Z}_+$ so that $1 < a < k + 1$ with $a|k + 1$; this means there is some $b \in \mathbb{Z}_+$ so that $ab = k + 1$. Since $1 < a$, we have $b < ab$, so $b < k + 1$. Also, $1 \leq b$; since $a < k + 1$ and $ab = k + 1$, we have $1 < b$. Thus a and b are products of primes, so $k+1$ is as well. Hence by the principle of mathematical induction, every integer $n > 1$ is a product of primes.

Now we want to show that for any integer $n > 1$, there is a unique way to realise n as a product of primes. More precisely, we want to show that if $p_1p_2 \cdots p_r = q_1q_2 \cdots q_s$ with $p_1, p_2, \ldots, p_r, q_1, q_2, \ldots, q_s$ primes so that $p_1 \leq p_2 \leq \cdots \leq p_r$ and $q_1 \leq q_2 \leq \cdots \leq q_s$, then $r = s$ and $p_i = q_i$ for all $i \in \mathbb{Z}$ with $1 \leq i \leq r$. To prove this, we argue by induction on $r \in \mathbb{Z}_+$.

More formally, for $r \in \mathbb{Z}_+$, we let $P(r)$ be the proposition that if $p_1 \cdots p_r = q_1 \cdots q_s$ with $p_1 \leq \cdots \leq p_r$ primes and $q_1 \leq \cdots \leq q_s$ primes, we have $r = s$ and $p_i = q_i$ for all $i \in \mathbb{Z}$ with $1 \leq i \leq r$.

Suppose first that $p_1 = q_1 \cdots q_s$ where p_1 is prime and $q_1 \leq \cdots \leq q_s$ are prime (here $s \in \mathbb{Z}_+$). Since p_1 is prime and thus cannot be a product of two or more primes, it must be the case that $s = 1$ and $p_1 = q_1$. [This proves the base case for the induction argument, i.e. this shows $P(1)$ holds.]

Now suppose that $k \in \mathbb{Z}_+$, and that whenever $n = p_1 \cdots p_k = q_1 \cdots q_s$ with $p_1, \cdots, p_k, q_1, \ldots, q_s$ primes and $p_1 \leq \cdots \leq p_k$, $q_1 \leq \cdots \leq q_s$, we have $k = s$ and $p_i = q_i$ for all $i \in \mathbb{Z}$ with $1 \leq i \leq k$. [This is the induction hypothesis.] Suppose now that $a = p_1 \cdots p_k p_{k+1} = q_1 \cdots q_t$ with $p_1, \ldots, p_k, p_{k+1}, q_1, \ldots, q_t$ primes and $p_1 \leq \cdots \leq p_k \leq p_{k+1}$, $q_1 \leq \cdots \leq q_t$. Note that since $k \geq 1$, a is not prime, and hence $t \geq 2$. Let p be the largest prime so that $p|a$. [Note that there are only finitely many primes dividing a since each such prime q satisfies $2 \leq q \leq a$, and there are only finitely many integers between 2 and a.] Thus we have $p \geq p_{k+1}$. Also, since $p|a$, we have $p|p_1 \cdots p_k p_{k+1}$, hence $p|p_i$ for some $i \in \mathbb{Z}$, $1 \leq i \leq k+1$. Since p_i is prime, we must have $p = p_i$. So $p = p_i \leq p_{k+1}$. Also, by our choice of p, we have $p \geq p_{k+1}$; hence we must have $p = p_{k+1}$.

A virtually identical argument shows that $p = q_t$, and hence $p_{k+1} = q_t$. Thus $p_1 \cdots p_k = q_1 \cdots q_{t-1}$. By the induction hypothesis, we have $k = t - 1$ and $p_i = q_i$ for $i \in \mathbb{Z}$ with $1 \leq i \leq k$. Therefore we have $k + 1 = t$ and $p_i = q_i$ for all $i \in \mathbb{Z}$ with $1 \leq i \leq k + 1$.

Consequently, by the principle of mathematical induction, the factorisation of an integer $n > 1$ as a product of (nondecreasing) primes is unique. □

Corollary 7.3. *For $x \in \mathbb{Q}_+$, $\exists!$ $a, b \in \mathbb{Z}_+$ so that $hcf(a, b) = 1$ and $x = \frac{a}{b}$.*

Proof. Take $x \in \mathbb{Q}_+$. Thus $\exists a, b \in \mathbb{Z}$, $a, b \neq 0$, so that $x = \frac{a}{b}$. Since $x > 0$, we have $x = |x| = \frac{|a|}{|b|}$, so we have that x is a quotient of two elements of $\mathbb{Z}_+$. So suppose $a, b > 0$. Let $c = \text{hcf}(a, b)$, and take $a', b' \in \mathbb{Z}_+$ so that $a = ca'$ and $b = cb'$. Thus $\text{hcf}(a', b') = 1$ and $x = \frac{a'}{b'}$. [This shows that there is at least one way to write any $x \in \mathbb{Q}_+$ as $\frac{a}{b}$ where $a, b \in \mathbb{Z}_+$ with $\text{hcf}(a, b) = 1$.]

Now suppose $x \in \mathbb{Q}_+$ and $a, b, c, d \in \mathbb{Z}_+$ so that $x = \frac{a}{b} = \frac{c}{d}$ with $\text{hcf}(a, b) = 1 = \text{hcf}(c, d)$. Thus $ad = bc$. Suppose $a = 1$; then $d = bc$, so $c|d$. Since $\text{hcf}(c, d) = 1$ and $c > 0$, this means $c = 1$ and hence $a = c$ and $b = d$. So suppose $a > 1$; then $a = p_1 \cdots p_r$ for some $r \in \mathbb{Z}_+$ and primes $p_1, \ldots, p_r$. We now argue by induction on r to show that $a = c$ and $b = d$. First, suppose $a = p_1$ (p_1 prime). We have $p_1|bc$ and $\text{hcf}(a, b) = 1$, so $\text{hcf}(p_1, b) = 1$ and hence $p_1|c$. Thus $c = p_1 c'$ for some $c' \in \mathbb{Z}_+$. So $d = bc'$ and hence $c'|d$. Since $\text{hcf}(c, d) = 1$ and c' is a positive factor of d, we must have $c' = 1$ and hence $a = p_1 = c$ and $b = d$. Now suppose $k \geq 1$ and whenever $p_1, \ldots, p_k$ are prime and $b', c', d' \in \mathbb{Z}_+$ so that $\text{hcf}(p_1 \cdots p_k, b') = 1 = \text{hcf}(c', d')$ with $p_1 \cdots p_k d' = b'c'$, we have $p_1 \cdots p_k = c'$. Suppose $a = p_1 \cdots p_k p_{k+1}$ where $p_1, \ldots, p_k, p_{k+1}$ are prime. Thus $p_{k+1}|c$, so $c = p_{k+1}c'$ for some $c' \in \mathbb{Z}_+$. Therefore $p_1 \cdots p_k d = bc'$, so by the induction hypothesis, $p_1 \cdots p_k = c'$. Hence $a = c$ and $b = d$. □

Corollary 7.4. *There are infinitely many prime numbers in $\mathbb{Z}_+$.*

Proof. (Euclid's proof) For the sake of contradiction, suppose there are only finitely many primes, and enumerate these as $p_1, p_2, \ldots, p_m$ where $m \in \mathbb{Z}_+$ is the number of primes. Now set $n = p_1 p_2 \cdots p_m + 1$. Clearly $m \geq 2$, as 2 and 3 are prime. Hence $n > 1$. By the Fundamental Theorem of Arithmetic, n can be factored as a product of primes; let q be a prime dividing n. So $n = qk$ for some $k \in \mathbb{Z}_+$ [$k > 0$ since $n > 0$ and $q > 0$]. Since there are only finitely many primes, we must have $q = p_j$ for some $j \in \mathbb{Z}_+$, $1 \leq j \leq m$. Hence

$$n = p_j k = p_1 p_2 \cdots p_m + 1,$$

so

$$1 = n - p_1 p_2 \cdots p_m = p_j \left(k - \frac{p_1 p_2 \cdots p_m}{p_j} \right).$$

Hence the prime p_j divides 1. Contradiction! Thus there cannot be a finite number of primes in $\mathbb{Z}_+$. □

Remark: One can use induction and the Fundamental Theorem of Arthmetic to prove the following generalisation of the Chinese Remainder Theorem: Suppose $r \in \mathbb{Z}$ with $r \geq 2$, and $m_1, \ldots, m_r \in \mathbb{Z}_+$ are pairwise relatively prime, meaning that $\text{hcf}(m_i, m_j) = 1$for $i, j \in \mathbb{Z}$ with $1 \leq i \leq r$, $1 \leq j \leq r$ and $i \neq j$. For any $a_1, \ldots, a_r \in \mathbb{Z}$, there is some $x \in \mathbb{Z}$ so that $x \equiv a_i \pmod{m_i}$ for all $i \in \mathbb{Z}$ with $1 \leq i \leq r$. Further, with x as above and $x' \in \mathbb{Z}$, we have $x' \equiv a_i \pmod{m_i}$ for $i = 1, 2, \ldots, r$ if and only if $x' \equiv x \pmod{m_1 m_2 \cdots m_r}$.

Application: We make use of the Fundamental Theorem of Arithmetic to find all primes p so that $5p + 9 = n^2$ for some $n \in \mathbb{Z}_+$.

[*Strategy:* First, we suppose we have a prime p so that $5p + 9 = n^2$ for some $n \in \mathbb{Z}_+$, and we deduce constraints on p. Then we consider all primes p subject to these constraints and determine for which of these p we have that $5p + 9 = n^2$ for some $n \in \mathbb{Z}_+$.]

Suppose p is prime and $n \in \mathbb{Z}_+$ so that $5p + 9 = n^2$. Since so $5p = (n+3)(n-3)$, and $n + 3 > 0$. By the Fundamental Theorem of Arithmetic, the only positive factors of $5p$ are $1, 5, p, 5p$. Since $n \in \mathbb{Z}_+$, we know that $n + 3$ is positive.

Suppose $n + 3 = 1$. Then $n - 3 = -5$, meaning $5p = (n+3)(n-3) = -5$. But this implies $p = -1$, which is not prime. So we cannot have $n + 3 = 1$.

Suppose $n + 3 = 5$. Then $n - 3 = -1$, so $5p = (n + 3)(n - 3) = -5$ and hence $p = -1$. But this is impossible [since -1 is not prime].

Suppose $n+3 = p$. Thus $n-3 = p-6$, so $5p = p(p-6)$. Hence $5 = p-6$, so $p = 11$, which is prime. [So $n + 3 = p$ does not lead to a contradiction.]

Suppose $n+3 = 5p$. Thus $5p = (n+3)(n-3) = 5p(n-3)$. Hence $n-3 = 1$, and so $n = 4$. Then $5p = (n + 3)(n - 3) = 7$; but this is impossible, since 5 does not divide [the prime] 7.

This shows that if p is a prime so that $5p + 9 = n^2$ for some $n \in \mathbb{Z}_+$ then $p = 11$. On the other hand, with $p = 11$, we have $5p + 9 = 55 + 9 = 64 = 8^2$.

Hence p is a prime with $5p+9 = n^2$ for some $n \in \mathbb{Z}_+$ if and only if $p = 11$.

8. Cardinality

Definitions. We say that two nonempty sets A and B have the same cardinality if there is a bijective map $f : A \to B$, and we write $|A| = |B|$. Note that (1) $h : A \to A$ defined by $h(x) = x$ is bijective; (2) when $f : A \to B$ is bijective, $f^{-1} : B \to A$ exists and is also bijective; and (3) if $f : A \to B$ and $g : B \to C$ are bijective, then so is $g \circ f : A \to C$. (We are tempted to say that having the same cardinality is an equivalence relation on "the set of all sets"; however, it is known that there is no injection from $\mathcal{P}(X) = \{A : \ A \subseteq X \ \}$ into X. So if X were "the set of all sets", we would have $\mathcal{P}(X) \subseteq X$, and the inclusion map $\iota : \mathcal{P}(X) \to X$ defined by $\iota(A) = A$ would contradict that there is no injection from $\mathcal{P}(X)$ into X. Thus "the set of all sets" does not exist.)

We say a set Y has at least as many elements as a set X if there is an injective map $f : X \to Y$, and in this case we write $|X| \leq |Y|$. If there is an injection from X into Y but no bijection between X and Y, we write $|X| < |Y|$. (Note that when $X \subseteq Y$ and $X \neq \emptyset$, the map $\iota : X \to Y$ defined by $\iota(x) = x$ is injective.)

Let A be a set. When $A = \emptyset$, we set $|A| = 0$. Now suppose $n \in \mathbb{Z}_+$ and $f : \{1, 2, \ldots, n\} \to A$ is bijective, we write $|A| = n$ and we say A has n elements. Further, we can enumerate the elements of A as $a_1, a_2, \ldots, a_n$ where $a_i = f(i)$, and since f is injective, $a_i = a_j$ if and only if $i = j$.

When $|A| \in \mathbb{Z}_{\geq 0}$, we say A is a finite set. When A is not a finite set we say A is an infinite set.

Suppose A is a finite set with $|A| = n$ ($n \in \mathbb{Z}_{\geq 0}$) and B is a subset of A; we accept without proof that $|B| = m$ where $m \leq n$ ($m \in \mathbb{Z}_{\geq 0}$), and that $A = B$ if and only if $m = n$. We also accept without proof that $\mathbb{Z}_+$ is infinite. Note that we have assumed the following: Suppose $B \subseteq A$; if A is finite then B is finite. The contrapositive of this statement is: Suppose $B \subseteq A$; if B is infinite then A is infinite.

We will eventually see that $|\mathbb{Z}_+| = |\mathbb{Z}| = |\mathbb{Q}_+| = |\mathbb{Q}|$, but $|\mathbb{Z}_+| < |\mathbb{R}|$.

Proposition 8.1. *Suppose $A, B \subseteq X$ where X is some set, where A, B are nonempty finite sets with $A \cap B = \emptyset$. Then $|A \cup B| = |A| + |B|$.*

Proof. Let $s, t \in \mathbb{Z}_+$ so that $|A| = s$ and $|B| = t$. Thus we can enumerate the elements of A as $a_1, a_2, \ldots, a_s$ where $a_i = a_j$ only if $i = j$ (here i, j are integers between 1 and s). Similarly, we can enumerate the elements of B as $b_1, b_2, \ldots, b_t$ where $b_i = b_j$ only if $i = j$ (here i, j are integers between 1 and t). We also know that for any integers i, j with $1 \leq i \leq s$ and $1 \leq j \leq t$, we have $a_i \neq b_j$ since $A \cap B = \emptyset$.

Define $f : \{1, 2, \ldots, s + t\} \to A \cup B$ by

$$f(n) = \begin{cases} a_n & \text{if } 1 \leq n \leq s, \\ b_{n-s} & \text{if } s < n \leq s + t. \end{cases}$$

As an exercise, one shows that f is bijective. □

Definition. We say a set X is countable if there is a bijective function $f : \mathbb{Z}_+ \to X$, or equivalently, if there is a bijective map $g : X \to \mathbb{Z}_+$. (Note: Some texts say a set if countable if it is finite or if there is a bijective function

$f : \mathbb{Z}_+ \to X$, and when there is a bijective function $f : \mathbb{Z}_+ \to X$, these texts say X is countably infinite.)

Note: Suppose X is a countable set; so by definition there is a bijective map $f : \mathbb{Z}_+ \to X$. Thus we can enumerate the elements of X as $x_1, x_2, x_3, \ldots$ where $x_i = f(i)$ for $i \in \mathbb{Z}_+$.

Example: The set of positive even integers is countable: Let $A = \{2x : x \in \mathbb{Z}_+ \}$. Define $f : \mathbb{Z}_+ \to A$ by $f(x) = 2x$. To see f is injective, suppose $x, y \in \mathbb{Z}_+$ so that $f(x) = f(y)$. Thus $2x = 2y$, so $x = y$, showing that f is injective. To see f is surjective, take $a \in A$. Thus $a = 2x$ for some $x \in \mathbb{Z}_+$, and hence $a = 2x = f(x)$; so f is surjective. This shows that f is bijective, and hence A is countable. Similarly, the set of odd positive integers, $\{2x - 1 : \ x \in \mathbb{Z}_+ \}$, can be shown to be countable.

Theorem 8.2. *Suppose $f : X \to Y$ is injective and $A \subseteq X$. Then $|A| = |f(A)|$.*

Proof. Let $B = f(A)$. Define $g : A \to B$ by $g(a) = f(a)$. By the definition of B, $B = f(A) = g(A)$, so g is surjective. Suppose $a, a' \in A$ so that $g(a) = g(a')$. Then $f(a) = f(a')$, and since f is injective, this means $a = a'$. Hence g is injective. Thus g is bijective, so $|A| = |g(A)|$. We also know that $g(A) = B = f(A)$, so $|A| = |g(A)| = |f(A)|$. □

The next theorem may seem intuitively obvious, but a proper proof is beyond the scope of this course.

Theorem 8.3. (a) *Every infinite set contains a countable subset.*

(b) *(Cantor-Schröder-Bernstein Theorem) If X, Y are sets with $|X| \leq |Y|$ and $|Y| \leq |X|$ then $|X| = |Y|$. That is, if X, Y are sets so that there exist injective functions $g : X \to Y$ and $h : Y \to X$ then there is a bijective function $f : X \to Y$.*

(In some texts, this theorem is called the Cantor-Bernstein Theorem or the Schöder-Bernstein Theorem; an interesting proof of this theorem due to Halmos can be found in the book by Pierre Grillet, which is available as an electronic book from the University of Bristol library.)

Corollary 8.4. *Suppose $X \subseteq \mathbb{Z}_+$. Then X is finite or countable.*

Proof. If X is finite then we are done. So suppose X is infinite. We have an injective map $g : X \to \mathbb{Z}_+$ given by $g(x) = x$, so $|X| \leq |\mathbb{Z}_+|$. On the other hand, we know X contains a countable subset A. Hence there is a bijective map $h : \mathbb{Z}_+ \to A$. Define $f : \mathbb{Z}_+ \to X$ by $f(n) = h(n)$. So f gives us an injective map from $\mathbb{Z}_+$ into X. Thus $|\mathbb{Z}_+| \leq |X|$, and so by the Cantor-Bernstein Theorem, there is a bijective map $f : \mathbb{Z}_+ \to X$. Hence X is countable. □

The next result is very useful when proving a set is countable.

Corollary 8.5. *Suppose X is an infinite set. Then X is countable if and only if there is an injective map $f : X \to \mathbb{Z}_+$.*

Proof. First suppose there is an injective map $f : X \to \mathbb{Z}_+$. Thus $|X| = |f(X)|$, and $f(X)$ is a subset of $\mathbb{Z}_+$. Since X is not finite and $|X| = |f(X)|$, $f(X)$ is not finite. Hence (by the previous corollary), $f(X)$ is countable, and hence X is countable.

Now suppose that X is countable. Thus there is a bijective $g : \mathbb{Z}_+ \to X$. Since g is bijective, g^{-1} exists. With $f = g^{-1}$, we have that $f : X \to \mathbb{Z}_+$ is bijective and hence injective. □

As exercises, one proves the following.

Proposition 8.6. *Suppose X is a countable set.*

(a) *Suppose A is a subset of X; then A is finite or countable.*
(b) *Suppose A is a subset of X. If A is finite then $X \setminus A$ is countable.*
(c) *X contains a subset B so that B and $X \setminus B$ are countable.*
(d) *Suppose $f : C \to X$ is injective; then C is finite or countable.*

Theorem 8.7. *Suppose $A, B \subseteq X$ where X is some set; suppose A is a countable set and B is a nonempty, finite set with $A \cap B = \emptyset$. Then $A \cup B$ is countable.*

Proof. [The idea of this proof is that of the "Hilbert hotel", where there is always room for another guest: The Hilbert hotel has countably many rooms, labeled $1, 2, 3, \ldots$ (so for each number in $\mathbb{Z}_+$, there is a room with that number). One night, all the rooms are occupied, and another potential guest arrives at the hotel looking for a room. The manager says, no problem! Then the manager announces to the guests that every guest is to move to the next room (so the guests in room n move to room $n + 1$). Thus all the guests still have rooms, and room 1 has been made available to the new arrival.]

Since A is countable, there is an injective map $f : A \to \mathbb{Z}_+$. B is finite, so we can list the distinct elements of B as $b_1, \ldots, b_m$ where $m = |B| \in \mathbb{Z}_+$. Define $g : A \cup B \to \mathbb{Z}_+$ by

$$g(x) = \begin{cases} i & \text{if } x = b_i, \\ f(x) + m & \text{if } x \in A. \end{cases}$$

We claim that g is injective. To see this, take $x, y \in A \cup B$ so that $x \neq y$. If $x, y \in B$ then $x = b_i$ and $y = b_j$ for some $i, j \in \mathbb{Z}_+$ so that $i \leq m$, $j \leq m$ with $i \neq j$, and hence $g(x) = i \neq j = g(y)$. If $x \in B$ and $y \in A$ then $x = b_i$ for some $i \in \mathbb{Z}_+$ with $i \leq m$, and hence $g(x) = i < m + 1 \leq g(y) + m + 1$. If $x, y \in A$, then since $x \neq y$ and f is injective, we have $f(x) \neq f(y)$ and so $g(x) = f(x) + m + 1 \neq f(y) + m + 1 = g(y)$. Therefore g is injective. Since $A \subseteq A \cup B$ and A is infinite, $A \cup B$ is infinite. Since $g : A \cup B \to \mathbb{Z}_+$ is injective and $A \cup B$ is infinite, $A \cup B$ is countable. □

Theorem 8.8. *$\mathbb{Z}_+ \times \mathbb{Z}_+$ is countable.*

Proof. We have that $\mathbb{Z}_+ \times \mathbb{Z}_+$ is infinite, as $\{(x, 1) : \ x \in \mathbb{Z}_+ \}$ is an infinite subset of $\mathbb{Z}_+ \times \mathbb{Z}_+$.

We arrange the elements of $\mathbb{Z}_+ \times \mathbb{Z}_+$ in a grid:

$$\begin{array}{cccccc} (1,1) & (1,2) & (1,3) & (1,4) & \cdots \\ (2,1) & (2,2) & (2,3) & (2,4) & \cdots \\ (3,1) & (3,2) & (3,3) & (3,4) & \cdots \\ (4,1) & (4,2) & (4,3) & (4,4) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{array}$$

We order the elements of this grid along the cross-diagonals:

$$(1,1);\ (1,2),\ (2,1);\ (1,3),\ (2,2),\ (3,1);\ldots$$

We will define $f : \mathbb{Z}_+ \times \mathbb{Z}_+ \to \mathbb{Z}_+$ so that $f((1,1)) = 1$, $f((1,2)) = 2$, $f((2,1)) = 3$, $f((1,3)) = 4$, $f(2,2) = 5$, $f(3,1) = 6$, etc. We now find a formula to define f.

The kth cross-diagonal contains the pairs $(1,k), (2,k-1), (3,k-2), \ldots, (k,1)$. So this cross-diagonal has k pairs. Thus the number of pairs in the first $k-1$ cross-diagonals is

$$1 + 2 + 3 + \cdots + (k-1) = \frac{(k-1)k}{2}.$$

Thus we define $f : \mathbb{Z}_+ \times \mathbb{Z}_+$ by

$$f((i, k+1-i)) = \frac{(k-1)k}{2} + i.$$

To show f is injective, suppose $x, y \in \mathbb{Z}_+ \times \mathbb{Z}_+$ so that $f(x) = f(y)$. Thus $\exists i, k \in \mathbb{Z}_+$ so that $i \leq k$ and $x = (i, k+1-i)$ (so x is on the kth cross-diagonal). Suppose first that y is also on the kth cross-diagonal; thus $\exists j \in \mathbb{Z}_+$ so that $j \leq k$ and $y = (j, k+1-j)$. Then

$$\frac{(k-1)k}{2} + i = f(x) = f(y) = \frac{(k-1)k}{2} + j.$$

Hence $i = j$ and so $x = y$. Now suppose y is not on the kth cross-diagonal. So there exist $j, m \in \mathbb{Z}_+$ so that $j \leq m$ and $y = (j, m+1-j)$. So y is on the mth cross-diagonal where $m \neq k$; hence $m > k$ or $k > m$. Without loss of generality, assume $m > k$. [If it is the case that $k > m$ then we rename x as y and y as x.] Thus $m = k + r$ for some $r \in \mathbb{Z}_+$. So

$$f(x) = \frac{k-1)k}{2} + i \leq \frac{(k-1)k}{2} + k,$$

and

$$\begin{aligned} f(y) &= \frac{(k+r-1)(k+r)}{2} + j \\ &= \frac{(k-1)k}{2} + kr + \frac{(r-1)r}{2} + j \\ &\geq \frac{(k-1)k}{2} + k + 1 \end{aligned}$$

(since $kr \geq k$, $r(r-1) \geq 0$, and $j \geq 1$). Therefore $f(x) \neq f(y)$, contradicting the assumption that $f(x) = f(y)$.

Hence, if $f(x) = f(y)$ then x and y are on the same cross-diagonal and $x = y$. This shows f is injective. $\square$

Suppose X, Y are countable. Then certainly $X \times Y$ is infinite: Choose $y_0 \in Y$. Define $f : X \times \{y_0\} \to X$ by $f(x, y_0) = x$. One easily shows f is bijective, so $|X \times \{y_0\}| = |X|$, and hence $X \times \{y_0\}$ is countable. As $X \times \{y_0\} \subseteq X \times Y$, $X \times Y$ is infinite (as it contains an infinite subset).

As an exercise, one proves the following somewhat anti-intuitive result.

Corollary 8.9. *$\mathbb{Q}_+$ and $\mathbb{Q}$ are countable.*

Also as an exercise, one proves the following.

Proposition 8.10. *Suppose X, Y are countable. Then:*

(a) *$X \times Y$ is countable.*
(b) *Suppose also $X \cap Y = \emptyset$. Then $X \cup Y$ is countable.*

Note: Since $\mathbb{Z}$ is an infinite subset of $\mathbb{Q}$ and $\mathbb{Q}$ is countable, $\mathbb{Z}$ must be countable.

Corollary 8.11. *Let $\{A_n : n \in \mathbb{Z}_+\}$ be a (countable) collection of countable sets that are pairwise disjoint. Then $\cup_{n\in\mathbb{Z}_+} A_n$ is countable.*

Proof. First note that since A_1 is infinite and $A_1 \subseteq \cup_{n\in\mathbb{Z}_+} A_n$, we know that $\cup_{n\in\mathbb{Z}_+} A_n$ is also infinite. We have seen that if X is a countable set, Y is an infinite set, and $\exists$ an injective map $g : Y \to X$, then Y is countable. Thus to prove that $\cup_{n\in\mathbb{Z}_+} A_n$ is countable, we will prove there is an injective function $g : \cup_{n\in\mathbb{Z}_+} A_n \to \mathbb{Z}_+ \times \mathbb{Z}_+$.

For each $n \in \mathbb{Z}_+$, enumerate the elements of A_n as $a_{n1}, a_{n2}, a_{n3}, \ldots$. [Recall that since A_n is countable, there is a bijective function $f_n : \mathbb{Z}_+ \to A_n$; for $k \in \mathbb{Z}_+$, set $a_{nk} = f_n(k)$.] Now define $g : \cup_{n=1}^{\infty} A_n \to \mathbb{Z}_+ \times \mathbb{Z}_+$ by $g(a_{mk}) = (m, k)$. [Since the A_n are pairwise disjoint, g is actually a function.] To see g is injective, suppose $g(a_{mk}) = g(a_{st})$ for some $m, k, s, t \in \mathbb{Z}_+$. Thus $(m, k) = (s, t)$, so $m = s$, $k = t$, and hence $a_{mk} = a_{st}$. Thus g is injective. So we have an injective function from $\cup_{n=1}^{\infty}$ into a countable set; since $\cup_{n=1}^{\infty} A_n$ contains the infinite set A_1 and hence is infinite, $\cup_{n=1}^{\infty} A_n$ is countable.

□

9. Uncountable sets and power sets

Definition. A set X is called uncountable if it is infinite but not countable.

We want to show that $\mathbb{R}$ is uncountable. To do this we will show that the interval $(0, 1) = \{x \in \mathbb{R} : 0 < x < 1\}$ is uncountable; as an exercise one shows that there is a bijection between the interval $(0, 1)$ and $\mathbb{R}$.

We assume that every real number between 0 and 1 has a decimal expansion of the form

$$0.a_1a_2a_3\cdots = \sum_{k\in\mathbb{Z}_+} a_k 10^{-k}$$

where $a_k \in \{0,1,2,3,4,5,6,7,8,9\}$ for each $k \in \mathbb{Z}_+$. Note that

$$\begin{aligned} 0.999\cdots &= \sum_{k\in\mathbb{Z}_+} 9\cdot 10^{-k} \\ &= 9\cdot\frac{1/10}{1-1/10} \\ &= 1 \end{aligned}$$

(recall that $\sum_{k\in\mathbb{Z}_+} 10^{-k}$ is a convergent geometric series). Consequently if there is some $N \in \mathbb{Z}_+$ so that $a_N \neq 9$ and $a_n = 9$ for all $n \in \mathbb{Z}_+$ with $n > N$, then $0.a_1a_2a_3\cdots = 0.a_1a_2\cdots a_{N-1}b_N$ where $b_N = a_N + 1$. We will assume the result that for every $\alpha \in \mathbb{R}$ with $0 < \alpha < 1$, there is a unique way to write α as $0.a_1a_2a_3\cdots$ so that $a_k \in \{0,1,2,3,4,5,6,7,8,9\}$ for all $k \in \mathbb{Z}_+$ and $\neg(\exists N \in \mathbb{Z}_+$ so that $\forall n \in \mathbb{Z}_+,\ n > N \implies a_n = 9)$.

Theorem 9.1. *The interval* $(0,1) = \{x \in \mathbb{R} :\ 0 < x < 1\ \}$ *is uncountable.*

Proof. (Cantor's diagonalisation argument) We know the interval $(0,1)$ is infinite, since $f : \mathbb{Z}_+ \to (0,1)$ defined by $f(k) = 10^{-k}$ is easily shown to be injective. For the sake of contradiction, suppose $(0,1)$ is countable. Thus we can enumerate the elements of $(0,1)$ as $\alpha_1, \alpha_2, \alpha_3, \ldots$. Write each α_k as a decimal expansion as described above:

$$\alpha_k = 0.a_{k1}a_{k2}a_{k3}\cdots$$

where $a_{ki} \in \{0,1,2,3,4,5,6,7,8,9\}$ and $\neg(\exists N \in \mathbb{Z}_+$ so that $\forall n \in \mathbb{Z}_+,\ n > N \implies a_{kn} = 9)$. For each $k \in \mathbb{Z}_+$, set

$$b_k = \begin{cases} 1 & \text{if } a_{kk} \neq 1, \\ 2 & \text{if } a_{kk} = 1. \end{cases}$$

Set $\beta = 0.b_1b_2b_3\cdots$. Thus $\beta \in \mathbb{R}$ with $0 < \beta < 1$ and $\neg(\exists N \in \mathbb{Z}_+$ so that $\forall n \in \mathbb{Z}_+,\ n > N \implies b_n = 9)$. Hence by assumption, $\beta = \alpha_m$ for some $m \in \mathbb{Z}_+$. But $b_m \neq a_{mm}$, contradicting the uniqueness of the representation of β as a decimal expansion not ending in an infinite sequence of 9s. Thus the assumption that the interval $(0,1)$ is countable leads to a contradiction, so $(0,1)$ must be uncountable. □

Remark: Suppose we have $m \in \mathbb{Z}_+$ and $a_1, a_2, \ldots, a_m \in \{0,1,2,3,4,5,6,7,8,9\}$ (not all 0), and

$$\alpha = 0.a_1a_2\cdots a_ma_1a_2\cdots a_ma_1a_2\cdots a_m\cdots = 0.\overline{a_1a_2\cdots a_m}.$$

Then α is a rational number: Let $b = \sum_{k=1}^{m} a_k\cdot 10^{m-k}$. Then $b \in \mathbb{Z}_+$ and

$$\begin{aligned} \alpha &= \sum_{n\in\mathbb{Z}_+} b\cdot 10^{-mn} \\ &= b\cdot\frac{10^{-m}}{1-10^{-m}} \\ &= \frac{b}{10^m-1}. \end{aligned}$$

Note that the map $g : (0,1) \to \mathbb{R}$ given by $g(x) = x$ is injective, so $|(0,1)| \leq |\mathbb{R}|$. Since $|\mathbb{Z}_+| < |(0,1)|$, we get $|\mathbb{Z}_+| < |\mathbb{R}|$, meaning $\mathbb{R}$ is uncountable. In the following corollary, we show $|(0,1)| = |\mathbb{R}|$, which is another way to argue that $\mathbb{R}$ is uncountable.

Corollary 9.2. *There is a bijection between the interval* $(0,1)$ *and* $\mathbb{R}$ *(and hence* $\mathbb{R}$ *is uncountable).*

Definition. For A a set, we let

$$\mathcal{P}(A) = \{C : \ C \subseteq A \ \}.$$

We call $\mathcal{P}(A)$ the power set of A.

Examples:

(a) $\mathcal{P}(\emptyset) = \{\emptyset\}$, so $|\mathcal{P}(\emptyset)| = 1$.

(b) For any nonempty set X, we know $\emptyset, X$ are distinct subsets of X, and hence $|\mathcal{P}(X)| \geq 2$.

(c) $\mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$, so $|\mathcal{P}(\{1,2\})| = 4 = 2^2$.

(d) $\mathcal{P}(\{1,2,3\}) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$. So $|\mathcal{P}(\{1,2,3\})| = 8 = 2^3$.

As an exercise, one proves the following.

Theorem 9.3. *Suppose* A *is a finite set with* $|A| = n$ *for some* $n \in \mathbb{Z}$ *with* $n \geq 0$. *Then* $|\mathcal{P}(A)| = 2^n$.

Remark: Suppose A is a finite set with n elements; enumerate these elements as $a_1, a_2, \ldots, a_n$. Let $Y = \{(c_1, c_2, \ldots, c_n) : \ c_i = 0 \text{ or } 1 \ \forall i \in \mathbb{Z} \text{ with } 1 \leq i \leq n \ \}$. Define $f : Y \to \mathcal{P}(A)$ by

$$f((c_1, c_2, \ldots, c_n)) = \{a_i \in A : \ c_i = 1 \text{ for some } i \in \mathbb{Z} \text{ with } 1 \leq i \leq n \ \}.$$

Thus $f((c_1, c_2, \ldots, c_n))$ is a subset of A; one can show that f is bijective.

As exercises, one proves the following.

Proposition 9.4. *Let* A, B *be sets.*

(a) $(A \subseteq B) \iff (\mathcal{P}(A) \subseteq \mathcal{P}(B))$.

(b) $\mathcal{P}(A) \cup \mathcal{P}(B) \subseteq \mathcal{P}(A \cup B)$.

(c) $\mathcal{P}(A) \cap \mathcal{P}(B) = \mathcal{P}(A \cap B)$.

Theorem 9.5. *(Cantor's Theorem) Let* X *be a set. Then* $|X| < |\mathcal{P}(X)|$.

Proof. When $X = \emptyset$, then we know $|X| = 0 < 1 = |\mathcal{P}(X)|$. So suppose $X \neq \emptyset$, and define $f : X \to \mathcal{P}(X)$ by $f(x) = \{x\}$. We show f is injective: Suppose $x_1, x_2 \in X$ so that $f(x_1) = f(x_2)$. Thus $\{x_1\} = \{x_2\}$, and hence $x_1 = x_2$. Therefore f is injective, so $|X| \leq |\mathcal{P}(X)|$.

Now we want to show there is no bijection between X and $\mathcal{P}(X)$. For the sake of contradiction, suppose there is a bijection $g : X \to \mathcal{P}(X)$. (So for each $x \in X$, $g(x)$ is a subset of X.) Define $A = \{x \in X : \ x \notin g(x) \ \}$. Then A is a subset of X, so $A \in \mathcal{P}(X)$. Also, since we have assumed g is bijective, there is some $z \in X$ so that $g(z) = A$. By the definition of g, $z \in A$ if and only if $z \notin g(z) = A$. Thus we have a contradiction (namely that $z \in A \iff z \notin A$). Hence our assumption that there is a bijective function $g : X \to \mathcal{P}(X)$ must be false. So $|X| < |\mathcal{P}(X)|$. $\square$

10. More proofs using contradiction, construction, and induction

Proposition 10.1. *For prime $p \in \mathbb{Z}$, $\sqrt{p}$ is irrational.*

Proof. For the sake of contradiction, suppose $\sqrt{p} \in \mathbb{Q}$. Thus $\sqrt{p} = \frac{a}{b}$ where $a, b \in \mathbb{Z}$ with $b \neq 0$; we can assume $\text{hcf}(a, b) = 1$. Then $p = \frac{a^2}{b^2}$, so $pb^2 = a^2$. Thus $p|a$; hence $a = pc$ for some $c \in \mathbb{Z}$. This means we have $pb^2 = (pc)^2 = p^2c^2$, so $b^2 = pc^2$ and hence $p|b$. But then $p|\text{hcf}(a, b)$, contradicting that $\text{hcf}(a, b) = 1$. Hence $\sqrt{p}$ cannot be rational. □

Proposition 10.2. *There are infinitely many primes in $\mathbb{Z}_+$; in fact, there are countably many primes.*

Proof. Let X be the set of primes. Note that since $X \subseteq \mathbb{Z}_+$, we know X is either finite or countable.

For the sake of contradiction, suppose there are only finitely many primes in $\mathbb{Z}_+$; let t be the number of primes. We know there is at least one prime, namely 2, so $t \geq 1$. Let $p_1, \ldots, p_t$ be all the primes in $\mathbb{Z}_+$. Consider $m = p_1 \cdots p_t + 1$. Since $m \in \mathbb{Z}$ with $m > 1$, by the Fundamental Theorem of Arithmetic we know there is some prime $q \in \mathbb{Z}_+$ so that $q|m$. So there is some $m' \in \mathbb{Z}$ so that $m = qm'$, and hence $1 = qm' - p_1 \cdots p_t$. Since we have assumed there are finitely many primes, we must have $q = p_j$ for some $j \in \mathbb{Z}$ with $1 \leq j \leq t$. Hence $1 = p_j m' - p_1 \cdots p_t$, so $p_j|1$. But since $q = p_j$ is prime and thus $p_j > 1$, this is impossible. Thus there cannot be finitely many primes. □

Given $a, b, c \in \mathbb{Z}$, we can use Euclid's algorithm to find all $x, y \in \mathbb{Z}$ so that $ax + by = c$. Before we prove the general theorem, let us consider a specific example.

Example: We want to construct all $x, y \in \mathbb{Z}$ so that $6x + 8y = 2$.

First note that for $x, y \in \mathbb{Z}$, we have $6x + 8y = 2$ if and only if we have $3x + 4y = 1$. Since $\text{hcf}(3, 4) = 1$, we know (by Euclid's algorithm) that $\exists s, t \in \mathbb{Z}$ so that $3s + 4t = 1$. (By inspection, we see that $3 \cdot (-1) + 4 \cdot 1 = 1$, so in this case we don't need to use Euclid's algorithm to find $s, t \in \mathbb{Z}$ so that $3s + 4t = 1$.) Now suppose we also have $x, y \in \mathbb{Z}$ so that $3x + 4y = 1$. Hence $3s + 4t = 3x + 4y$, so $3(s - x) = 4(y - t)$. Thus $3|4(y - t)$, and since $\text{hcf}(3, 4) = 1$, $3|y - t$. Hence $\exists k \in \mathbb{Z}$ so that $y - t = 3k$, or equivalently, $y = t + 3k$. A virtually identical argument shows that $4|s - x$, so $\exists k' \in \mathbb{Z}$ so that $x = s - 4k'$. Therefore

$$3s + 4t = 3x + 4y = 3(s - 4k') + 4(t + 3k),$$

hence $0 = -12k' + 12k$, or equivalently, $k' = k$. In summary, we have shown that if $s, t, x, y \in \mathbb{Z}$ so that $3s + 4t = 1 = 3x + 4y$, then $\exists k \in \mathbb{Z}$ so that $x = s - 4k$ and $y = t + 3k$.

On the other hand, suppose $3s + 4t = 1$ (which is the case when $s = -1$, $t = 1$). Take any $k \in \mathbb{Z}$ and set $x = s - 4k$, $y = t + 3k$. Then

$$3x + 4y = 3s + 4t = 1.$$

So $3x + 4y = 1$ if and only if $x = -1 - 4k$, $y = 1 + 3k$ for some $k \in \mathbb{Z}$, and hence $6x + 8y = 2$ if and only if $x = -1 - 4k$, $y = 1 + 3k$ for some $k \in \mathbb{Z}$.

More generally, we have the following proposition and corollary, which one proves as exercises.

Proposition 10.3. *Fix $a, b, c \in \mathbb{Z}$ so that $a, b \neq 0$. Let $d = hcf(a, b)$. Take $a', b' \in \mathbb{Z}$ so that $a = da'$ and $b = db'$.*

(a) *If $d \nmid c$ then there do not exist $x, y \in \mathbb{Z}$ so that $ax + by = c$.*
(b) *Suppose $d|c$. Then $\exists s, t \in \mathbb{Z}$ so that $as + bt = c$. Also, for $x, y \in \mathbb{Z}$, we have $ax + by = c$ if and only if $\exists k \in \mathbb{Z}$ so that $x = s - b'k$ and $y = t + a'k$.*

Corollary 10.4. *Fix $a, b, n \in \mathbb{Z}$ so that $n \geq 1$. There $\exists x \in \mathbb{Z}$ so that $ax \equiv b \pmod{n}$ if and only if $hcf(a, n)|b$.*

Proposition 10.5. *Suppose $m \in \mathbb{Z}_+$ with $m \geq 2$ and $A_1, \ldots, A_m$ are nonempty, finite sets.*

(a) *Suppose $A_1, \ldots, A_m$ are pairwise disjoint, meaning that for $i, j \in \mathbb{Z}_+$ with $i, j \leq m$ and $i \neq j$, we have $A_i \cap A_j = \emptyset$. Then*

$$|A_1 \cup \cdots \cup A_m| = |A_1| + \cdots + |A_m|.$$

(b) $|A_1 \times \cdots \times A_m| = |A_1| \cdots |A_m|$.

Proof. We prove (b) and leave (a) as an exercise.

(b) We argue by induction on m.

[Base case.] Let $|A_1| = s$, $|A_2| = t$. Thus we know there exist bijections $f : \{1, 2, \ldots, s\} \to A_1$ and $g : \{1, 2, \ldots, t\} \to A_2$. For $i, j \in \mathbb{Z}_+$ with $i \leq s$, $j \leq t$, set $a_i = f(i)$ and set $b_j = g(j)$. Notice that since f and g are bijections, $a_1, a_2, \ldots, a_s$ are distinct and $b_1, b_2, \ldots, b_t$ are distinct.

We define $h : \{1, 2, \ldots, st\} \to A_1 \times A_2$ as follows. Take $n \in \{1, 2, \ldots, st\}$. Recall that as a consequence of the division algorithm for $\mathbb{Z}$, $\exists! q, r \in \mathbb{Z}$ so that $n = tq + r$ where $1 \leq r \leq t$. Note that since $n \geq 1$, we must have $q \geq 0$, for if $q < 0$ then $q \leq -1$ and $n \leq -t + r \leq 0$. Also note that $q < s$, else $st + 1 \leq n = tq + r \leq st$. Hence $a_{q+1} \in A_1$, and $b_r \in A_2$. We define

$$h(n) = (a_{q+1}, b_r) \text{ where } q, r \in \mathbb{Z} \text{ so that } n = tq + r \text{ where } 1 \leq r \leq t.$$

Since the conditions on q and r determine them uniquely, there is no ambiguity in the meaning of $h(n)$, or in other words, h is well-defined.

We need to show that h is bijective. Suppose first that $m, n \in \{1, 2, \ldots, st\}$ so that $h(m) = h(n)$. Take the unique $q, r, q', r' \in \mathbb{Z}$ so that $n = tq + r$, $m = tq' + r'$ where $1 \leq r \leq t$, $1 \leq r' \leq t$. Then $(a_{q'+1}, b_{r'}) = f(m) = f(n) = (a_{q+1}, b_r)$. Thus we have $a_{q'+1} = a_{q+1}$ and $b_{r'} = b_r$; hence $q' + 1 = q + 1$ (since $a_1, a_2, \ldots, a_s$ are distinct) and $r' = r$ (since $b_1, b_2, \ldots, b_t$ are distinct). Thus $m = tq' + r' = tq + r = n$, showing that h is injective. Now take an arbitrary element $(a_i, b_j) \in A_1 \times A_2$; thus $1 \leq i \leq s$ and $1 \leq j \leq t$, so $1 \leq t(i - 1) + j \leq st$. Hence with $n = t(i - 1) + j$, we have $h(n) = (a_i, b_j)$, showing that h is surjective. Thus h is bijective. So $|A_1 \times A_2| = st = |A_1||A_2|$.

[Induction step.] Suppose that $k \in \mathbb{Z}$ with $k \geq 2$, and suppose that $|A_1 \times \cdots \times A_k| = |A_1| \cdots |A_k|$. Set $A = A_1 \times \cdots \times A_k$. Thus

$$|A_1 \times \cdots \times A_k \times A_{k+1}| = |A \times A_{k+1}|,$$

and by the base case, we know $|A \times A_{k+1}| = |A| \cdot |A_{k+1}|$. So using the induction hypothesis, we get

$$|A_1 \times \cdots \times A_k \times A_{k+1}| = |A| \cdot |A_{k+1}| = |A_1| \cdots |A_k| \cdot |A_{k+1}|.$$

Hence by the principle of mathematical induction, (b) holds for all $m \in \mathbb{Z}$ with $m \geq 2$. □

As an exercise, one proves the following.

Proposition 10.6. *The union of countably many nonempty, pairwise disjoint finite sets is countable.*

Proposition 10.7. *Suppose A and B are nonempty finite sets with $|A| = |B|$.*

(a) *Suppose $f : A \to B$ is injective. Then f is bijective.*
(b) *Suppose $f : A \to B$ is surjective. Then f is bijective.*

Proof. Let $n \in \mathbb{Z}_+$ so that $n = |A|$. So $|B| = n$, and there are bijections $g : \{1, 2, \ldots, n\} \to A$ and $h : \{1, 2, \ldots, n\} \to B$; for $i \in \{1, 2, \ldots, n\}$, set $a_i = g(i)$, $b_i = h(i)$. Since g is injective, this means $a_1, \ldots, a_n$ are distinct; similarly, $b_1, \ldots, b_n$ are distinct.

(a) Suppose $f : A \to B$ is injective. Then $f(a_1), \ldots, f(a_n)$ are distinct, so $|f(A)| = |A| = n$. Since $f(A) \subseteq B$ and $|f(A)| = |B| = n \leq \infty$, we must have $f(A) = B$. Thus f is surjective and hence bijective.

(b) Suppose $f : A \to B$ is surjective. For the sake of contradiction, suppose f is not injective. Thus $\exists i, j \in \{1, 2, \ldots, n\}$ so that $a_i \neq a_j$ but $f(a_i) = f(a_j)$. Since $a_i \neq a_j$, we have $i \neq j$. Thus

$$f(A) = \{a_k : \ k \in \mathbb{Z}, \ 1 \leq k \leq n, \ k \neq j \ \}.$$

So $|f(A)| < n$. But since f is surjective, we know $f(A) = B$ and hence $|f(A)| = |B| = n$. This gives us a contradiction, so we must have that f is injective and hence bijective. □

Finally, we offer a "party trick" based on the theory presented in these notes.

Take $x \in \mathbb{Z}_+$. Written as a decimal expansion, we write x as

$$a_m a_{m-1} \cdots a_1 a_0$$

where $a_i \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ for $0 \leq i \leq m$. Thus

$$x = \sum_{i=0}^{m} a_i 10^i.$$

We know $10 \equiv 1 \pmod 9$. One uses induction to show that for all $i \in \mathbb{Z}_+$, we have $10^i \equiv 1 \pmod 9$. Hence, again using induction, one shows that

$$\sum_{i=0}^{m} a_i 10^i \equiv \sum_{i=0}^{m} a_i \pmod 9.$$

Recall that x is divisible by 9 if and only if $x \equiv 0 \pmod 9$, so x is divisible by 9 if and only if the digits of x sum to a number divisible by 9.

One can devise a similar party trick to test for divisibility by 11. In this case one uses that for $i \in \mathbb{Z}$, $i \geq 0$, $10^i \equiv 1 \pmod{11}$ when i is even, and $10^i \equiv -1 \pmod{11}$ when i is odd. [So what is the party trick?]