MA132/138: FOUNDATIONS/SETS AND NUMBERS, SHARED LECTURE NOTES

SAUL SCHLEIMER

Here we discuss some of the foundational material that will be assumed in many other mathematics modules at Warwick. All exercises in these notes have solutions (in Appendix A). There is also a short glossary (Appendix B), an index of notation, and a list of references.

While the material is old, its presentation here is new. In particular these lecture notes have not been previously used. So, please inform me via the anonymous form (`https://forms.gle/wAsUQUMo79BKubHk6`) or via an email (s.schleimer@warwick.ac.uk) of any errors (or of possible improvements).

Acknowledgements. I thank the many teachers, colleagues, and students without whom this work would not exist. In particular: John W. Addison Jr. taught me logic and set theory before I knew what those were. His course introduced me to Halmos' book [8] which continues to be a valuable resource. Simon Thomas mentored me (and explained all of the tricky bits) the first time I taught logic and set theory. Oleg Kozlovski, Christopher Lazda, David Wood, and Robert Kropholler were/are my fellow-lecturers in MA132/138 in years previous and present; I thank them for their good humour, gentle corrections, and patience.

1. Sets and subsets

1.1. **Sets.** A *set* contains its *elements*. We do not give a more formal definition; instead we understand sets by their properties – that is, by the axioms they satisfy.

Notation 1.2. Suppose that X is a set. Then we may write X as a sequence of elements, preceded and succeeded by curly braces. $\diamond$

Notation 1.3. We take $\mathbb{N}$ to be the set of natural numbers (including zero). Then we may write

$$\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\}$$

This is pronounced as "the natural numbers equals the set of zero, one, two, three, four, and so on." The three dots (an *ellipsis*) indicate that the sequence is to be completed in the "usual way".[1] ◇

Here we include zero as a natural number. Throughout these notes we assume some knowledge of the order on $\mathbb{N}$ as well as the arithmetic operations (addition, multiplication, and subtraction and division where defined) and their basic properties (commutativity, associativity, distributivity, and cancellativity).

As a simpler example of a set, suppose that X is the set containing exactly the natural numbers 0, 1, and 2. Then we may write

$$X = \{0, 1, 2\}$$

Notation 1.4. Suppose that X is a set. Suppose that x is an element of X. Then we may denote this by writing $x \in X$. Suppose that y is not an element of X. Then we may denote this by writing $y \notin X$. ◇

For example, we have $0 \in \mathbb{N}$ while $1/2 \notin \mathbb{N}$. The former is pronounced as "zero is an element of the natural numbers" and the latter is pronounced "one-half is not an element of the natural numbers". We now state our first axiom.

Axiom 1.5 (Extension)**.** *Suppose that X and Y are sets. Suppose that for any $x \in X$ we have $x \in Y$. Suppose also that for any $y \in Y$ we have $y \in X$. Then $X = Y$.* □

From this axiom we deduce the following equalities of sets.

$$\{0, 1, 2\} = \{0, 2, 1\} = \{1, 2, 0\} = \{0, 0, 0, 1, 2\}$$

That is: neither the order of the elements nor the number of times they appear within the braces effects the identity of the set. Thus a set is quite different from a *list*: that is, an ordered collection of elements, each of which may appear more than once. We give a more formal treatment of lists in Definition 4.18.

We finish this section by introducing a special set.

Definition 1.6. If X contains no elements then X is an *empty set*. ◇

Exercise 1.7. Suppose that X and Y are empty sets. Prove that $X = Y$. ◇

Thus there is at most one empty set. To obtain at least one, we require another axiom.

Axiom 1.8 (Empty set)**.** *There is an empty set, denoted $\varnothing$.* □

[1] There are various formal definitions of the natural numbers: for example see [8, page 44].

1.9. **Subsets.**

Definition 1.10. Suppose that X and Y are sets. Suppose that for any $x \in X$ we have $x \in Y$. Then we say that X is a *subset* of Y. We denote this by writing $X \subset Y$.

If X is a subset of Y, and not equal to Y, then we say that X is a *proper subset* of Y. ◇

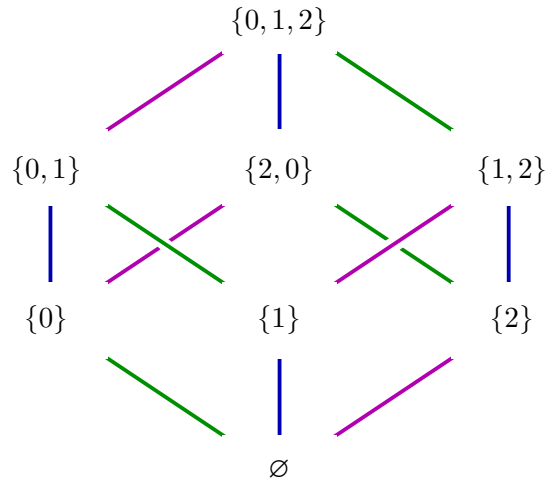


FIGURE 1.11. The subsets of $X = [\![3]\!] = \{0, 1, 2\}$. In this *Hasse diagram* we connect a subset A to a subset B by an ascending arc if $A \subset B$ and B has exactly one more element than A.

Exercise 1.12. Suppose that X is a set.

(1) Prove that the empty set $\varnothing$ is a subset of X.
(2) Prove that X is a subset of X. ◇

Notation 1.13. For any natural number n, we define

$$[\![n]\!] = \{0, 1, 2, 3, \ldots, n-1\} = \{k \in \mathbb{N} \mid k < n\}$$

That is, $[\![n]\!]$ is the subset of the natural numbers which are less than n. ◇

For example, $[\![0]\!]$ is another name for the empty set. Also the set $X = \{0, 1, 2\}$ mentioned in Section 1.1 is equal to $[\![3]\!]$. We display all eight subsets of X in Figure 1.11.

Exercise 1.14. Sketch a proof that $[\![n]\!]$ has 2^n subsets. ◇

Definition 1.15. Suppose that X is a set with exactly one element. Then we call X a *singleton* set. ◇

For example, among the subsets of $[\![3]\!]$ there are exactly three singleton sets, namely $\{0\}$, $\{1\}$, and $\{2\}$.

We now state a useful lemma.

Lemma 1.16. *Suppose that X and Y are sets. Then the following are equivalent.*

- $X = Y$.
- $X \subset Y$ *and* $Y \subset X$.

Proof. Suppose that $X = Y$. By the second part of Exercise 1.12 we have $X \subset X$. Thus $X \subset Y$ and $Y \subset X$.

Suppose that $X \subset Y$ and $Y \subset X$. Then every element of X belongs to Y and every element of Y belongs to X. Thus $X = Y$ by the axiom of extension. □

2. Power sets, specification, and functions

2.1. **Power sets.** Exercise 1.14 leads us to the following.

Axiom 2.2 (Power set)**.** *Suppose that X is a set. Then there exists a set whose elements are exactly the subsets of X.* □

The set given by the axiom is called the *power set* of X. We denote this set by $\mathcal{P}(X)$. Power sets grow frighteningly quickly. For example, by Exercise 1.14, the number of elements in $\mathcal{P}(\mathcal{P}([\![n]\!]))$ is 2^{2^n}.

2.3. **Specification.** Recall that in Notation 1.13 we wrote

$$[\![n]\!] = \{k \in \mathbb{N} \mid k < n\}$$

and asserted that $[\![n]\!]$ is a set. To justify this, we need the following.

Axiom 2.4 (Specification)**.** *Suppose that X is a set. Suppose that $S(x)$ is a property*[2] *which may or may not hold for elements of X. Then there is a subset $Y \subset X$ whose elements are exactly those elements x of X for which $S(x)$ holds.* □

Note in the above the requirement to "specify" the containing set X.

Notation 2.5. Suppose that X is a set. Suppose that $S(x)$ is a property. Then we write

$$\{x \in X \mid S(x)\}$$

to denote the subset given by Axiom 2.4. ◇

[2]The precise definition of *property* belongs to a first course in set theory.

So, for example, in the notation $[\![3]\!]$ the containing set is $\mathbb{N}$ and the property $S(k)$ is $(k < 3)$.

Remark 2.6. Building sets in this way is so common that we typically will not invoke the axiom of specification explicitly. ◇

2.7. **Functions.** Here is our first, slightly informal, definition of *function*.

Definition 2.8. A *function* $f\colon X \to Y$ consists of the following.

- A set X, called the *domain*.
- A set Y, called the *codomain*.
- A *rule* f which, for every $x \in X$, gives an element $f(x) \in Y$. ◇

We do not give a precise meaning to the word "rule"; in Definition 4.16 we give a more formal definition of functions. For two functions to be equal, they must have the same domains, codomains, and rules.

The notation $f\colon X \to Y$ is pronounced "f is a function from X to Y". If $X = Y$ then we may say that "f is a function on X".[3]

Here is the simplest kind of function.

Definition 2.9. Suppose that X is a set. The *identity function*

$$\mathrm{Id}_X\colon X \to X$$

is defined by the rule $\mathrm{Id}_X(x) = x$. ◇

Non-identity functions are required to actually "do something". One way to record this is as follows.

Definition 2.10. Suppose that X and Y are sets. Suppose that $f\colon X \to Y$ is a function. Suppose that $x \in X$ and $y \in Y$ are elements so that $y = f(x)$. Then we say that y is the *image* of x under f. ◇

For the next example we take $\mathbb{R}$ to be the set of real numbers and $\mathbb{R}_{\geq 0}$ to be the set of non-negative real numbers.

Example 2.11. Consider the following functions.

(1) $f\colon \mathbb{R} \to \mathbb{R}$ where $f(x) = x^2$
(2) $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ where $g(x) = x^2$
(3) $h\colon \mathbb{R} \to \mathbb{R}_{\geq 0}$ where $h(x) = x^2$
(4) $k\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ where $k(x) = x^2$

These functions have differing domains or codomains (or both). Thus all four functions are distinct. ◇

The next example shows that the rule for a function need not be algebraic.

[3]Some authors instead say that $f\colon X \to X$ is a "function from X to itself".

Lemma 2.12. *Suppose that $X = \{0, 1, 2\}$. Then there are 27 functions on X.*

Proof. Suppose that f is a function on X. Thus the domain and codomain are both equal to X. We can specify the rule for f by listing, in order, the elements $f(0)$, $f(1)$, and $f(2)$. There are three possibilities for each, and no other restrictions. Thus the number of possible rules is $3 \times 3 \times 3 = 27$. □

Exercise 2.13.

(1) Explicitly list all functions on $[\![2]\!]$.
(2) For any $n \in \mathbb{N}$, count the number of functions on $[\![n]\!]$. Give a proof that your answer is correct in the style of Lemma 2.12. ◇

3. Types of functions, cardinality and counting, Cantor's theorem and Russell's paradox

3.1. **Injections, surjections, bijections.** For the next three definitions we suppose that X and Y are sets and that $f : X \to Y$ is a function.

Definition 3.2. Suppose that f has the following property:

- For every x and x' in X, if $f(x) = f(x')$ then $x = x'$.

Then we call f *injective*. ◇

Definition 3.3. Suppose that f has the following property:

- For every y in Y, there is some x in X with $f(x) = y$.

Then we call f *surjective*. ◇

Definition 3.4. Suppose that f is both injective and surjective. Then we call f *bijective*. ◇

For example, the identity function Id_X is bijective. As shown by the next exercise, the status of a function – injective, surjective, both, or neither – depends not only on its rule but also on its domain and codomain.

Exercise 3.5. For each of the functions in Example 2.11, determine if it is injective, surjective, both, or neither. Give short justifications of your answers. ◇

Definition 3.6. Suppose that $f : X \to X$ is a bijection. Then we also call f a *permutation* of X. ◇

Exercise 3.7. Suppose that n is a natural number. Count, with proof, the number of permutations of $[\![n]\!]$. ◇

3.8. **Cardinality.** We have already done some counting in the exercises. We now give the formal terminology.

Definition 3.9. Suppose that X and Y are sets. Suppose that there is some bijection $f\colon X \to Y$. Then we say that X and Y have the same *cardinality*. When this happens we write $|X| = |Y|$. $\diamond$

Definition 3.10. Suppose that X is a set. We say that X is *finite* if it has the same cardinality as $[\![n]\!]$ for some $n \in \mathbb{N}$. In this case we write $|X| = n$ and we say that X has cardinality equal to n. $\diamond$

In the above definition we asked only for the existence of *some* $n \in \mathbb{N}$ with $[\![n]\!]$ in bijection with X. This leaves open the possibility that there are many such natural numbers. However, using the pigeonhole principle, in Exercise 17.11(4), we show that this does not happen.

3.11. **Cantor's theorem.** We now state and prove our first theorem.

Theorem 3.12 (Cantor's theorem)**.** *Suppose that X is a set. Suppose that $\mathcal{P}(X)$ is its power set. Suppose that $f\colon X \to \mathcal{P}(X)$ is a function. Then f is not a surjection.*

Proof. For a contradiction, suppose that f is a surjection. That is, for all $A \in \mathcal{P}(X)$ there is some $a \in X$ so that $f(a) = A$. We use f, and the axiom of specification, to build a subset of X as follows.

$$\mathcal{C} = \{x \in X \mid x \notin f(x)\}$$

That is, $\mathcal{C}$ is the subset of X consisting of those elements x which do not lie in the subset $f(x)$. Since f is surjective, there is some element of X, call it c, so that $f(c) = \mathcal{C}$. There are now two possibilities[4]: either c is an element of $\mathcal{C}$, or it is not.

Suppose that c is an element of $\mathcal{C}$. Thus, by the defining property of $\mathcal{C}$, we have that $c \notin f(c)$. Since $f(c) = \mathcal{C}$, we deduce that $c \notin \mathcal{C}$, a contradiction.

Suppose instead that $c \notin \mathcal{C}$. Since $f(c) = \mathcal{C}$ we deduce that $c \notin f(c)$. Thus, by the defining property of $\mathcal{C}$, we have that $c \in \mathcal{C}$, another contradiction.

As both possibilities lead to contradiction, we deduce that our original assumption (the surjectivity of f) was incorrect. This completes the proof. $\square$

Exercise 3.13. Suppose that X is a set. Suppose that $\mathcal{P}(X)$ is its power set. Prove that there is no injection $g\colon \mathcal{P}(X) \to X$. $\diamond$

[4]Here we use the *law of the excluded middle*.

As a consequence of Theorem 3.12, we have that if X is a set, then there is no bijection from X to $\mathcal{P}(X)$. One often abbreviates Cantor's theorem by saying that, for each set X is "smaller than" its power set. This is an intuitive notion for finite sets, as $n < 2^n$ for all natural numbers. However sets which are not finite are more difficult to understand.

Definition 3.14. Suppose that X is a set. We say that X is *infinite* if it is not finite. That is, X does not have the same cardinality as any set $[\![n]\!]$, for any $n \in \mathbb{N}$. ◇

For example, $\mathbb{N}$ is infinite; this follows from the *pigeonhole principle*. See Exercise 17.11(3).

Remark 3.15. From Cantor's theorem we have that no set X is in bijection with its power set. This also applies when X is infinite. For example, $\mathbb{N}$ is not in bijection with $\mathcal{P}(\mathbb{N})$. The latter power set is infinite; its singletons are in bijection with $\mathbb{N}$. However, Cantor tells us that $\mathcal{P}(\mathbb{N})$ is not in bijection with (and also does not inject into) $\mathbb{N}$. Thus $\mathcal{P}(\mathbb{N})$ is "larger" than $\mathbb{N}$. From this we deduce that there are "different sizes of infinities". ◇

3.16. **Russell's paradox.** This section is not examinable. The axiom of specification (Axiom 2.4) is sometimes called the axiom of "restricted comprehension". This is because it was historically preceded by (what is now called) the axiom of "unrestricted comprehension".

Axiom 3.17 (Unrestricted comprehension)**.** *Suppose that $S(x)$ is a property. Then there is a set Y whose elements are exactly those x for which $S(x)$ holds.* □

Unrestricted comprehension is much more powerful than specification; it eliminates the pesky requirement to "specify" the containing set X in advance. It therefore allows us to make many more sets. In fact, too many.

Theorem 3.18 (Russell's paradox)**.** *The axiom of unrestricted comprehension leads to contradiction.*

Proof. We consider the property $S(X)$ which states that X is not an element of itself. In symbols we have $S(X) = (X \notin X)$.

The axiom of unrestricted comprehension now gives us the set $R = \{X | X \notin X\}$. That is, R is the set of sets which do not contain themselves (as elements). Since R is itself a set, it either contains itself (as an element) or it does not.

Suppose that R contains itself as an element. Then, by the definition of R, we have that $S(R)$ holds. Thus $R \notin R$. Thus R does not contained itself as an element. This is a contradiction.

Suppose instead that R does not contains itself as an element. Then $S(R)$ holds. So by the definition of R it contain R as an element. This is also a contradiction, and completes the proof. □

At the heart of both Cantor's theorem and Russell's paradox there is a kind of "self-reference" or "iteration". Such ideas appear usefully in various parts of mathematics.

4. Ordered pairs, cartesian products, graphical relations, lists and strings

We now turn to the problem of giving a more formal definition of functions. Recall that the issue with Definition 2.8 was that the term "rule" was undefined.

4.1. Ordered pairs and cartesian products.

We first discuss unordered pairs.

Axiom 4.2 (Pair)**.** *Suppose that X and Y are sets. Then there is a set whose elements are exactly X and Y.* □

This set is written as $\{X, Y\}$. We call this the *unordered pair* of X and Y.

Definition 4.3. Suppose that X and Y are sets. Suppose that $x, x' \in X$ and $y, y' \in Y$ are elements. Then an *ordered pair* (x, y) contains x and y, in that order.[5] Two ordered pairs (x, y) and (x', y') are equal if and only if $x = x'$ and $y = y'$. ◇

If (x, y) is an ordered pair then we call x its first *entry* and y its second *entry*.

Axiom 4.4 (Ordered pairs)**.** *Suppose that X and Y are sets. Suppose that $x \in X$ and $y \in Y$. Then the ordered pair (x, y) exists.*[6] □

Remark 4.5. To obtain ordered triples, or quadruples, or higher, we could add further axioms. Or, we could define $(x, y, z) = ((x, y), z)$. Or, we could define $(x, y, z) = (x, (y, z))$.

We do not make a choice here. Instead we boldly assert that the choice of formalisation is not important; we will rely on our intuition to carry us though any difficulties. ◇

Definition 4.6. Suppose that X and Y are sets. Then the collection of all ordered pairs with first entry from X and second from Y is

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}$$

[5]Informally, ordered pairs are identical to lists of length two. However, if we allowed this in our formal definitions, then we would have a circularity. Namely, lists are defined in terms of functions (Definition 4.18), functions are defined in terms of ordered pairs (Definition 4.16), and finally ordered pairs are lists.

At a basic level it seems that there are three kinds of definitional scheme: unfounded, circular, or by infinite regress. Mathematicians have a strong preference for unfounded definitions, which they dignify with the name "axiomatic". Whether we can so avoid paradox remains to be seen.

[6]There are various constructions of ordered pairs as sets. Perhaps the most common is due to Kuratowski [1921]. He takes $(x, y) = \{\{x\}, \{x, y\}\}$. Note that this definition of ordered pair relies on Axiom 4.2.

and is called the *cartesian product* of X and Y. ◇

Axiom 4.7. *Suppose that X and Y are sets. Then the cartesian product $X \times Y$ is a set.*[7] □

The notation $X \times Y$ is very similar to the multiplicative notation for numbers. Here is one justification of that similarity.

Exercise 4.8. Sketch a proof that $[\![m]\!] \times [\![n]\!]$ has cardinality $m \times n$. ◇

There is a special case of the cartesian product when $X = Y$. Here we use the notation X^2 for $X \times X$. This notation is again justified by Exercise 4.8.

Remark 4.9. For example, addition and multiplication of natural numbers give functions from $\mathbb{N}^2$ to $\mathbb{N}$. ◇

Exercise 4.10. Suppos that X is a set. Show that $X \times \varnothing = \varnothing \times X = \varnothing$. ◇

4.11. **Graphical relations and functions.**

Definition 4.12. A *relation* R from X to Y consists of the following.

- A set X, called the *domain*.
- A set Y, called the *codomain*.
- A subset of $X \times Y$. ◇

If $X = Y$ then we say that R is "a relation on X". If (x, y) is an element of R we may denote this by writing xRy.

The empty set is a subset of $X \times Y$; this gives the *empty relation*. Likewise $X \times Y$ is a subset of itself; this gives the *universal relation*. When $X = Y$ then the set

$$\{(x, y) \in X^2 \mid x = y\}$$

gives the *identity relation* on X. Here is a more interesting example.

Example 4.13. Taking $X = Y = \mathbb{R}$ we have $X \times Y = \mathbb{R}^2$. The *unit circle* is

$$S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

Since S^1 is a subset of $\mathbb{R}^2$ it is also a relation on $\mathbb{R}$. ◇

Definition 4.14. Suppose that X and Y are sets. A relation G from X to Y is *graphical* if for every $x \in X$ there is exactly one $y \in Y$ so that (x, y) lies in G. ◇

[7] If we rely on the Kuratowski constructions of ordered pairs then the existence of $X \times Y$ follows from the axioms of union and power set.

If $X = Y = \mathbb{R}$ then $X \times Y = \mathbb{R}^2$ is the usual plane. In this case there is a pictorial interpretation of Definition 4.14; namely, a relation $G \subset \mathbb{R}^2$ on $\mathbb{R}$ is graphical if and only if every vertical line in $\mathbb{R}^2$ meets G in exactly one point.

Example 4.15. For example, of the following relations

$$P = \{(x, y) \in \mathbb{R}^2 \mid x^2 = y\} \quad Q = \{(x, y) \in \mathbb{R}^2 \mid x = y^2\}$$

the first is graphical and the second is not. ◇

Definition 4.16. A *function* $f : X \to Y$ is a relation G, from X to Y, which is graphical. If (x, y) lies in G then we write $f(x) = y$. ◇

Note that in this definition the function f again has a domain and a codomain (which it inherits from the underlying relation G).

As an example, the "upwards parabola" P of Example 4.15 gives the function $f : \mathbb{R} \to \mathbb{R}$ with underlying relation $\{(x, x^2) \mid x \in \mathbb{R}\}$. On the other hand the "sideways parabola" Q of Example 4.15 and the circle S^1 of Example 4.13 are not graphical and so do not give functions.

4.17. **Lists.**

Definition 4.18. Suppose that $\mathcal{A}$ is a set. Suppose that n is a natural number. A *list* L is a function $L : [\![n]\!] \to \mathcal{A}$. We call n the *length* of the list L. We also say that the elements of L are *taken from* $\mathcal{A}$. ◇

Since lists are functions, two lists are equal if and only if they

- have the same length,
- have elements taken from the same set, and
- have the same elements in the same order.

Notation 4.19. Suppose that C is a list of length n taken from the set $\mathcal{A}$. Then we may write either $C = (a_k)_{k=0}^{n-1}$ or

$$C = (a_0, a_1, a_2, \ldots, a_{n-1})$$

That is, we may write C as a sequence of elements of $\mathcal{A}$, preceded and succeeded by parentheses. ◇

Accordingly, the following lists are all distinct.

$$(0, 1, 2) \quad (0, 2, 1) \quad (1, 2, 0) \quad (0, 0, 0, 1, 2)$$

The first three lists are distinguished by having different last elements. Also, the last list has length five while the others have length three.

Lists of real numbers are often called *vectors*. According to Definition 4.18, the length of a list is always some natural number. In other treatments "infinite" lists – that is, functions from $\mathbb{N}$ – are allowed. These infinite lists are often called *sequences*.

4.20. **Strings.** Suppose that $\mathcal{A}$ is a set, here called an *alphabet*. The elements of $\mathcal{A}$ are called *letters* (or *characters*). A list $C = (a_i)$, consisting of letters from $\mathcal{A}$, is often called a *string* (or a *word*) over $\mathcal{A}$. When elements of the alphabet $\mathcal{A}$ are represented by single symbols, then the representation of the string w may omit the parentheses, the commas, and the spaces. Since strings are lists, they have a length: that is, the number of letters in the word.

Definition 4.21. Over any alphabet $\mathcal{A}$ there is a unique string $\epsilon_{\mathcal{A}}$ of length zero; this is called the *empty string*. ◇

Definition 4.22. Suppose that $\mathcal{B} = \{0, 1\}$ is our alphabet. We call the elements of $\mathcal{B}$ *bits*. Strings over $\mathcal{B}$ are called *binary strings*. ◇

As examples of binary strings we have $\epsilon_{\mathcal{B}}$ (of length zero), 00 (of length two), 0101010001 (of length ten), and so on.

Exercise 4.23.

(1) Determine the number of binary strings of length five.
(2) Determine the number of binary strings of length n.
(3) Determine the number of binary strings of length n having exactly one bit which is 1.
(4) Determine the number of binary strings of length n having exactly two bits which are 1. ◇

5. New sets from old

5.1. **The mathematicians' "and" and "or".** In informal mathematical proofs, we are somewhat free in our usage of natural language. For example if P and Q are sentences, then as usual we have that "P and Q" holds if and only if both P and Q hold.

However when writing proofs, even informally, in our use of the word "or" we depart somewhat from natural language. We instead assume that "P or Q" holds if and only if *at least* one of P and Q hold. In particular we do *not* require that exactly one of P and Q holds.

As an example consider the following sentences.

$$P = (0 < 1 \text{ as natural numbers}), \quad Q = (0 = 1 \text{ as natural numbers})$$

So P holds and Q does not. Thus "P and Q" does not hold while "P or Q" does hold. Also, both "P and P" and "P or P" hold while neither "Q and Q" nor "Q or Q" holds.

In the above (and everywhere in mathematics) we necessarily define and discuss our mathematical language in terms of natural language. Thus we have defined the mathematical "and" and "or" partly in terms

of their meanings in natural language. This is unavoidable; the feeling of circularity so produced also appears to be unavoidable.

5.2. **Union.**

Axiom 5.3. *Suppose that $\mathbb{X}$ is a set whose elements are again sets. Then the collection of all elements of elements of $\mathbb{X}$ forms a set.* □

This set is denoted

$$\bigcup_{X \in \mathbb{X}} X = \{z \mid z \in X \text{ for some } X \in \mathbb{X}\}$$

and is called the *union* over $\mathbb{X}$. We next concentrate on a special case of Axiom 5.3.

Notation 5.4. Suppose that X and Y are sets. Then we may form the union of X and Y as

$$X \cup Y = \{z \mid z \in X \text{ or } z \in Y\}$$

To see that this is a set we form $\mathbb{X} = \{X, Y\}$ using the axiom of pairing. We then take the union over $\mathbb{X}$ using the axiom of unions. ◇

Example 5.5.

(1) If m and n are natural numbers then $[\![m]\!] \cup [\![n]\!] = [\![\max(m, n)]\!]$.
(2) Suppose that $X = \{0, 1, 2\}$ and $Y = \{1, 2, 3\}$. Then $X \cup Y = \{0, 1, 2, 3\}$. ◇

Lemma 5.6. *Suppose that X and Y are sets. Then $X \subset X \cup Y$.*

Proof. To prove that $X \subset X \cup Y$ we must show that every element of X is an element of $X \cup Y$. So fix any element $x \in X$. Then x lies in X or it lies in Y. Thus x lies in $X \cup Y$. Since this holds for every x in X we are done. □

The following lemma likewise depends on the close connection between the union operator and the word "or".

Lemma 5.7. *Suppose that X, Y, and Z are sets. Then we have the following.*

(1) $X \cup \varnothing = X$ (identity)
(2) $(X \cup Y) \cup Z = X \cup (Y \cup Z)$ (associativity)
(3) $X \cup Y = Y \cup X$ (commutativity)
(4) $X \subset Y$ if and only if $X \cup Y = Y$ (absorption)
(5) $X \cup X = X$ (idempotent)

Exercise 5.8. Provide the proof of Lemma 5.7 ◇

5.9. **Intersection.**

Definition 5.10. Suppose that $\mathbb{X}$ is a set whose elements are again sets. The *intersection* over $\mathbb{X}$ is the collection

$$\bigcap_{X \in \mathbb{X}} X = \{x \mid x \in X \text{ for every } X \in \mathbb{X}\}$$

In the special case where $\mathbb{X} = \{X, Y\}$ we instead write

$$X \cap Y = \{z \mid z \in X \text{ and } z \in Y\}$$

for the *intersection* of X and Y. ◇

We could not prove that unions are sets directly; instead we needed a new axiom. The situation for intersections is simpler, because intersections are smaller than unions.

Lemma 5.11. *Suppose that $\mathbb{X}$ is a set whose elements are again sets. Then the intersection over $\mathbb{X}$ is a set.*

Proof. Let $\mathcal{X} = \bigcup_{X \in \mathbb{X}} X$ be the union over $\mathbb{X}$. We define the property $S(x)$ as follows:

$$S(x) = (\text{for every } X \in \mathbb{X}, \text{ we have } x \in X)$$

Then the intersection over $\mathbb{X}$ equals $\{x \in \mathcal{X} \mid S(X)\}$. This is a set by the axiom of specification. □

Notation 5.12. Suppose that X and Y are sets. Then we may form the intersection of X and Y as

$$X \cap Y = \{z \mid x \in X \text{ and } x \in Y\}$$

To see that this is a set we form $\mathbb{X} = \{X, Y\}$ using the axiom of pairing. We then take the intersection over $\mathbb{X}$ using Lemma 5.11.

If $X \cap Y = \varnothing$ then we say that X and Y are *disjoint*. ◇

There is a formal symmetry between union and intersection that mirrors the informal symmetry between the words "or" and "and". Because of this, the proofs of the following lemmas are modelled, extremely closely, on those of Lemmas 5.6 and 5.7. We leave their writing out to the reader.

Lemma 5.13. *Suppose that X and Y are sets. Then $X \cap Y \subset X$.* □

Lemma 5.14. *Suppose that X, Y, and Z are sets. Then we have the following.*

(1) $X \cap \varnothing = \varnothing$ (annihilator)
(2) $(X \cap Y) \cap Z = X \cap (Y \cap Z)$ (associativity)
(3) $X \cap Y = Y \cap X$ (commutativity)
(4) $X \subset Y$ if and only if $X \cap Y = X$ (absorption)
(5) $X \cap X = X$ (idempotent) □

5.15. **Set difference.**

Definition 5.16. Suppose that X and Y are sets. The collection

$$X - Y = \{x \in X \mid x \notin Y\}$$

is called the *set-theoretic difference* of X and Y. ◇

$X - Y$ is also called, in some sources, the "relative complement of Y in X".

Lemma 5.17. *Suppose that X and Y are sets. Then $X - Y$ is a set.*

Proof. This follows from the axiom of specification, using X as the set and

$$S(x) = (x \notin Y)$$

as the property. □

The proof of the next lemma is immediate from the definitions.

Lemma 5.18. *Suppose that X is a set. Then we have the following.*

(1) $X - \varnothing = X$
(2) $\varnothing - X = \varnothing$
(3) $X - X = \varnothing$.

In particular, if X is not empty, then we have $X - \varnothing \neq \varnothing - X$. Thus, as with subtraction of numbers, set-theoretic difference is not commutative.

5.19. **The boolean algebra of sets.** The operations of union, intersection, and set-theoretic difference are related to each other in various ways. Here we mention only some of the best known.

Lemma 5.20 (Distributive laws)**.** *Suppose that X, Y, and Z are sets. Then we have the following.*

- $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$
- $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$

Exercise 5.21. Provide the proof of Lemma 5.20. ◇

Lemma 5.22 (De Morgan's laws)**.** *Suppose that X, Y, and Z are sets. Then we have the following.*

- $X - (Y \cup Z) = (X - Y) \cap (X - Z)$
- $X - (Y \cap Z) = (X - Y) \cup (X - Z)$

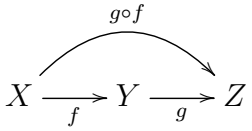
Exercise 5.23. Provide the proof of Lemma 5.22. ◇

6. New functions from old

6.1. **Composition.**

Definition 6.2. Suppose that X, Y, and Z are sets. Suppose that $f\colon X \to Y$ and $g\colon Y \to Z$ are functions. Then we define their *composition* to be the function $g \circ f\colon X \to Z$ given by $(g \circ f)(x) = g(f(x))$. ◇

The expression $g \circ f$ is pronounced as "g composed with f". Here is a "diagram" representing the composition.



Lemma 6.3. *Suppose that X, Y, Z, and W are sets. Suppose that $f\colon X \to Y$, $g\colon Y \to Z$, and $h\colon Z \to W$ are functions. Then we have*

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

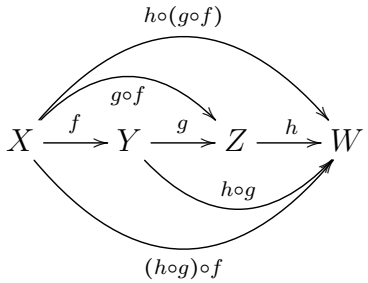
That is, composition of functions is *associative*.

Proof of Lemma 6.3. We note that $h \circ (g \circ f)$ and $(h \circ g) \circ f$ have the same domain (X) and codomain (W). Thus it only remains that they have the same rule. Suppose that x is any element of X. We compute as follows:

$$\begin{aligned} (h \circ (g \circ f))(x) &= h(g(f(x))) && \text{definition of } \circ \text{ twice} \\ &= ((h \circ g) \circ f)(x) && \text{definition of } \circ \text{ twice} \end{aligned}$$

Since this holds for all x in X, we deduce that $h \circ (g \circ f) = (h \circ g) \circ f$, as desired. □

The above proof can be reproduced in a more diagrammatic form.



Definition 6.4. Suppose that X and Y are sets. Suppose that $f\colon X \to Y$ and $g\colon Y \to X$ are functions. If $g \circ f = \mathrm{Id}_X$ then we say that g is a *left inverse* for f. Similarly we say that f is a *right inverse* for g.

If $g \circ f = \mathrm{Id}_X$ and $f \circ g = \mathrm{Id}_Y$ then we say that g is an *inverse* for f. ◇

Lemma 6.5. *Suppose that X and Y are non-empty sets. Suppose that $f\colon X \to Y$ is a function. Then we have the following.*

(1) f is injective if and only if f has a left inverse.
(2) f is surjective if and only if f has a right inverse.
(3) f is bijective if and only if f has an inverse. In this case the inverse is unique.

Proof. We prove the forward and backwards directions of the first statement. The rest are left as exercises.

Suppose that f is injective. Since X is non-empty, we can fix some x_0 in X. We now define $g\colon Y \to X$ as follows.

$$g(y) = \begin{cases} x, & \text{if } f(x) = y \\ x_0, & \text{if there is no } x \in X \text{ with } f(x) = y \end{cases}$$

Suppose now that x is any element of X. Set $y = f(x)$. We compute as follows:

$$\begin{aligned} (g \circ f)(x) &= g(f(x)) && \text{definition of } \circ \\ &= g(y) && \text{definition of } y \\ &= x && \text{first line of definition of } g \\ &= \mathrm{Id}_X(x) && \text{definition of } \mathrm{Id}_X \end{aligned}$$

Since this holds for all x in X, we deduce that $g \circ f = \mathrm{Id}_X$, as desired.

Suppose that g is a left inverse for f. Suppose that x and x' are any elements of X. Suppose that $f(x) = f(x')$. We now compute as follows:

$$\begin{aligned} x &= \mathrm{Id}_X(x) && \text{definition of } \mathrm{Id}_X \\ &= (g \circ f)(x) && g \text{ is a left inverse for } f \\ &= g(f(x)) && \text{definition of } \circ \\ &= g(f(x')) && \text{because } f(x) = f(x') \\ &= (g \circ f)(x') && \text{definition of } \circ \\ &= \mathrm{Id}_X(x') && g \text{ is a left inverse for } f \\ &= x' && \text{definition of } \mathrm{Id}_X \end{aligned}$$

Since this holds for all x and x', we deduce that f is injective, as desired. □

Exercise 6.6. Provide the rest of the proof of Lemma 6.5. ◇

Lemma 6.7. *Suppose that X, Y, and Z are sets. Suppose that $f\colon X \to Y$ and $g\colon Y \to Z$ are functions. Then we have the following.*

(1) If f and g are injective then so is $g \circ f$.

(2) If f and g are surjective then so is $g \circ f$.

Exercise 6.8. Provide the proof of Lemma 6.7. ◇

6.9. **Iteration and orbits.** There is an important special case of function composition.

Definition 6.10. Suppose that X is a set. Suppose that f is a function on X. Then we define

(1) $f^{(0)} = \mathrm{Id}_X$ and
(2) for any $n \in \mathbb{N}$ we define $f^{(n+1)} = f \circ f^{(n)}$.

We say that $f^{(n)}$ is the n^{th} *iterate* of f. ◇

Sometimes iteration produces nothing new.

Definition 6.11. Suppose that X is a set. Suppose that f is a function on X. We say that $x \in X$ is a *fixed point* of f if $f(x) = x$. ◇

But often iteration produces many things.

Definition 6.12. Suppose that X is a set. Suppose that f is a function on X. Then for any $x \in X$ we define

$$\mathcal{O}_f(x) = \{f^{(n)}(x) \mid n \in \mathbb{N}\}$$

We call $\mathcal{O}_f(x)$ the *forward orbit* of x under f. ◇

In Definition 6.12 we used a new notation for sets. Here are the formalities.

Definition 6.13. Suppose that X and Y are sets. Suppose that $f\colon X \to Y$ is a function. Suppose that $Z \subset X$ is a subset. Then we write

$$f(Z) = \{f(z) \mid z \in Z\}$$

to denote the set

$$\{y \in Y \mid \text{there is some } x \in Z \text{ so that } f(x) = y\}$$

The set $f(Z)$ is called the *image* of Z under f. ◇

To complement images we have the following.

Definition 6.14. Suppose that X and Y are sets. Suppose that $f\colon X \to Y$ is a function. Suppose that $W \subset Y$ is a subset. Then

$$f^{-1}(W) = \{x \in X \mid f(x) \in W\}$$

is the *preimage* of Z under f. ◇

6.15. **Cantor–Schoerder–Bernstein.**

Theorem 6.16. *Suppose that X and Y are sets. Suppose that $f\colon X \to Y$ and $g\colon Y \to X$ are injections. Then there is a bijection $h\colon X \to Y$.* □

The proof, while elementary, is beyond the scope of this module. As a hint of the ideas involved, define $h = g \circ f$ and $k = f \circ g$. Both of these are injections by Lemma 6.7. We then divide each of X and Y into three pieces using the structure of orbits of h and of k. We then assemble these pieces to find the desired bijection $F\colon X \to Y$.

7. Reflexive, symmetric, and transitive relations, equivalence relations, order relations

7.1. **Reflexive, symmetric, and transitive relations.**

Definition 7.2. Suppose that X is a set. Suppose that $R \subset X^2$ is a relation on X.

- Suppose that, for all $x \in X$, we have xRx. Then we say that R is *reflexive*.
- Suppose that, for all $x, y \in X$, we have that xRy implies yRx. Then we say that R is *symmetric*.
- Suppose that, for all $x, y, z \in X$, we have that xRy and yRz implies xRz. Then we say that R is *transitive*. ◇

Example 7.3. All of the following are relations on $\mathbb{N}$.

- The identity relation is reflexive, symmetric, and transitive.
- The empty relation is symmetric and transitive, but not reflexive.
- The relation of less than or equal ($x \leq y$) is reflexive and transitive, but not symmetric.
- The relation ($|x - y| \leq 1$) is reflexive and symmetric, but not transitive.
- The relation ($x = y$ or $x + 1 = y$) is reflexive, but not symmetric or transitive.
- The relation of inequality ($x \neq y$) is symmetric but not reflexive or transitive.
- The relation of less than ($x < y$) is transitive, but not reflexive or symmetric.
- The relation ($x + 1 = y$) is not reflexive, symmetric, or transitive.

◇

Exercise 7.4. For each of the following relations on $\mathbb{N}$ determine if it is reflexive, symmetric, or transitive.

(1) xPy if $x + y = 2$
(2) xQy if $x - y \leq 2$
(3) xRy if $|x - y| \leq 2$ ◇

7.5. **Equivalence relations.**

Definition 7.6. Suppose that X is a set. A relation on X is an *equivalence relation* if it is reflexive, symmetric, and transitive. ◇

Remark 7.7. This definition is modelled on the properties of equality. So one is tempted to call equality of sets an equivalence relation. However this is not the case, because the collection of all sets is not a set. ◇

Here is a simple-sounding exercise that still gives a real feel for the nature of equivalence relations.

Exercise 7.8. Suppose that $X = [\![4]\!]$. List all equivalence relations on X. Verify that your list is *complete* (all equivalence relations appear) and *irredundant* (no equivalence relation appears twice). ◇

The same exercise, but with $X = [\![5]\!]$, is even more challenging. We leave this to the reader, together with the hint to look at sequence A000110 at the Online Encyclopedia of Integer Sequences [15].

The difficulty of these enumerations suggest that we require a more efficient way to think about equivalence relations. This will be provided in Section 8.

7.9. **Partial orders.**

Definition 7.10. Suppose that X is a set. Suppose that R is a relation on X. Then R is *antisymmetric* if, for all x and y in X we have

- if xRy and yRx then $x = y$. ◇

For example, the relation $\leq$ on $\mathbb{N}$ is antisymmetric.

Definition 7.11. Suppose that X is a set. Suppose that R is a relation on X. If R is reflexive, antisymmetric, and transitive, then R is a *partial order*. ◇

The identity relation on X is one example. Another is the relation of $\leq$ on $\mathbb{N}$. A third is the relation of divisibility on positive natural numbers. Here is an example which is closely related to that.

Example 7.12. Suppose that X is a set. Suppose that $\mathcal{P}(X)$ is the power set of X. The relation $P \subset Q$ on $\mathcal{P}(X)$ is reflexive and transitive by Definition 1.10. It is antisymmetric by Lemma 1.16. Thus it is a partial order. We call the pair $(\mathcal{P}(X), \subset)$ the *boolean poset* on X. ◇

The word "poset" stands for "partially ordered set".

7.13. **Total orders.**

Definition 7.14. Suppose that X is a set. Suppose that R is a relation on X. We say that R is a *total relation* if, for all x and y we have xRy or yRx. ◇

Definition 7.15. Suppose that X is a set. Suppose that R is a relation on X. We say that R is a *total order* if it is reflexive, antisymmetric, transitive, and total. ◇

That is, a total order is a partial order which is, additionally, total. Examples of total orders include "less than or equal" on the natural numbers (as well as the integers, the rationals, and the reals).

8. Partitions, equivalence classes, and quotients

We now begin to explain a much more convenient technique for dealing with equivalence relations.

8.1. **Partitions.**

Definition 8.2. Suppose that X is a set. Suppose that $\mathbb{P} \subset \mathcal{P}(X)$ has the following properties:

- If $P \in \mathbb{P}$ then P is non-empty.
- $X = \bigcup_{P \in \mathbb{P}} P$.
- If $P, Q \in \mathbb{P}$ then either $P = Q$ or $P \cap Q = \varnothing$.

Then we call $\mathbb{P}$ a *partition* of X. The elements $P \in \mathbb{P}$ are called the *parts* of the partition. ◇

Any non-empty set X has two special partitions. These are

- The partition into singletons: $\mathbb{I} = \{\{x\} \mid x \in X\}$.
- The partition with one part: $\mathbb{U} = \{X\}$.

Here are the five partitions of $[\![3]\!]$:

$$\{\{0\},\{1\},\{2\}\}\ \{\{0,1\},\{2\}\}\ \{\{0,2\},\{1\}\}\ \{\{0\},\{1,2\}\}\ \{\{0,1,2\}\}$$

The notation here, while correct, interferes with understanding. So we give the same information in a more efficient way.

$$0|1|2 \qquad 01|2 \qquad 02|1 \qquad 0|12 \qquad 012$$

Here we place a vertical bar "|" between parts. We list partitions with smaller parts before those with larger parts. As a consequence of this we list partitions with more parts before those with fewer parts. Since a part is a subset, inside a part we list its elements in terms of size. This is only for convenience; for example the partitions 01|23 and 10|32 of $[\![4]\!]$ are equal.

Exercise 8.3.

(1) Count the number of partitions of the set $X = [\![4]\!]$.
(2) Sketch a proof that the set $X = [\![n]\!]$ has at least $2^{n-1} - 1$ partitions.

◇

8.4. **Equivalence classes.**

Definition 8.5. Suppose that X is a set. Suppose that E is an equivalence relation on X. Suppose that x lies in X. We define the set

$$[x]_E = \{y \in X \mid xEy\}$$

to be the *equivalence class* of x. ◇

We have two trivial examples. In the identity relation on X equivalence classes are singletons. In the universal relation on X there is only one equivalence classes, namely X.

Exercise 8.6. Compute the equivalence classes for each of the equivalence relations you found in Exercise 7.8. ◇

Exercise 8.7. Here is a slightly non-mathematical problem. Suppose that X is the set of living people. We define an equivalence relation B on X where xBy if x and y are born in the same month. Count the number of equivalence classes for B. ◇

Notation 8.8. Suppose that X is a set. Suppose that E is an equivalence relation on X. Then we may denote the set of equivalence classes by[8]

$$X/E = \{[x]_E \mid x \in X\}$$

We call X/E the *quotient* of X by E. We also write $q_E : X \to X/E$ for the function $q_E(x) = [x]_E$. We call q_E the *quotient map* given by E. ◇

Lemma 8.9. *Suppose that X is a set. Suppose that E is an equivalence relation on X. Then the quotient X/E is a partition of X.*

Proof. We must verify that $\{[x]_E \mid x \in X\}$ is a partition of X.

Since E is reflexive we have that x lies in $[x]_E$. Thus $[x]_E$ is non-empty. Also, since this holds for all x we have that $X = \bigcup_{x \in X} [x]_E$.

Suppose that x and y lie in X. If $[x]_E$ is disjoint from $[y]_E$ then there is nothing left to show. So suppose that z lies in $[x]_E \cap [y]_E$. Thus xEz and yEz. By symmetry we have zEy. By transitivity we have xEy. Suppose now that w is any element in $[y]_E$. Thus yEw. Since xEy, by transitivity we have xEw. Thus $[y]_E \subset [x]_E$.

The same argument, swapping x and y throughout, proves that $[x]_E \subset [y]_E$. From Lemma 1.16 implies $[x]_E = [y]_E$, as desired. □

As a temporary piece of notation, we use $\Pi_X(E) = X/E$ to denote the function that takes equivalence relations on X to partitions of X. Lemma 8.9 can now be sharpened, as follows.

Corollary 8.10. *Suppose that X is a set. Then the function Π is a bijection.*

Proof. Suppose that $\mathbb{P}$ is a partition of X. We define a relation $Q = Q_{\mathbb{P}}$ on X where xQy if x and y lie in the same part of $\mathbb{P}$. Note that Q

[8] It is more correct here to write

$$X/E = \{P \in \mathcal{P}(X) \mid \text{there is some } x \in X \text{ so that } P = [x]_E\}$$

However the notation in 8.8 is much more common.

is reflexive, symmetric, and transitive, so Q is an equivalence relation. This gives a function $\Sigma(\mathbb{P}) = Q_{\mathbb{P}}$ from partitions to equivalence relations.

Suppose that E is an equivalence relation. Then $Q = \Sigma(\Pi(E))$ is also equivalence relation. By construction Q and E have the same equivalence classes. Thus xEy if and only if xQy. Thus $E = Q$. It follows that Σ is a left inverse for Π.

Suppose that $\mathbb{P}$ is a partition. Then $\mathbb{Q} = \Pi(\Sigma(\mathbb{P}))$ is also a partition. So the parts of $\mathbb{Q}$ are the equivalence classes of $\Sigma(\mathbb{P})$. But these are, by construction, the parts of $\mathbb{P}$. So $\mathbb{P} = \mathbb{Q}$. It follows that Σ is a right inverse for Π. By Lemma 6.5 we have that Π is a bijection. □

8.11. **Integers.** The construction of quotients is one of the jewels of modern mathematics: a natural way to construct new things out of old. Here we lay out our first example; it will not be the last.

Example 8.12. Suppose that $X = \mathbb{N}^2$. We define a relation $E \subset X^2 = (\mathbb{N}^2 \times \mathbb{N}^2)$ by

$$(p, q)E(r, s) \quad \text{if and only if} \quad p + s = r + q$$

Here the additions make sense because p, q, r, and s are natural numbers. ◇

Exercise 8.13. Prove that E, as given by Example 8.12, is an equivalence relation. ◇

With Exercise 8.13 in hand we may explore the equivalence classes of E. Suppose that p lies in $\mathbb{N}$. Then we have

$$[(p, 0)]_E = \{(p + k, k) \in \mathbb{N}^2 \mid k \in \mathbb{N}\}$$
$$[(0, p)]_E = \{(k, p + k) \in \mathbb{N}^2 \mid k \in \mathbb{N}\}$$

To "see" this, we draw $\mathbb{N}^2$ as a quarter plane, with x- and y-axis being the first and second copies of $\mathbb{N}$. The points of the equivalence class $[(p, q)]_E$ lie along the line, with slope one, through the point (p, q). See Figure 8.14.

The figure shows that the elements of a fixed equivalence class share a "mysterious" property; namely $(p, q)E(r, s)$ if and only if $p - q = r - s$. (We call this "mysterious" because subtraction is not defined for natural numbers.) We enshrine this observation in our next definition.

Definition 8.15. Let E be the equivalence relation on $\mathbb{N}^2$ given in Example 8.12. Then we write

$$\mathbb{Z} = \mathbb{N}^2/E$$

We call any element of $\mathbb{Z}$ an *integer*. Also, if k is a natural number then we write $k_{\mathbb{Z}} = [(k, 0)]_E$. We call $0_{\mathbb{Z}}$ and $1_{\mathbb{Z}}$ the *zero* and *one* of $\mathbb{Z}$. ◇

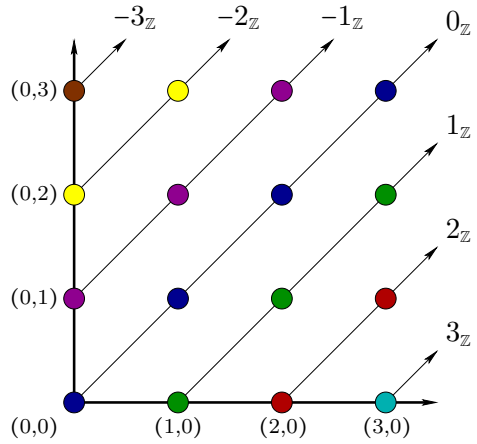


FIGURE 8.14. The equivalence classes of E (Example 8.12).

9. Well-defined functions and integer arithmetic

9.1. **Arithmetic in $\mathbb{Z}$.** We now would like to be able to add and multiply integers. Moreover, and this is rather the point of integers, we would also like to negate and subtract them. There is a natural choice for all of these operations, as well as a certain problem to be overcome.

Definition 9.2. Recall that $\mathbb{Z} = \mathbb{N}^2/E$ where E is the equivalence relation given in Example 8.12. We define three operations on $\mathbb{Z}$, as follows.

$$\begin{aligned} [(p,q)]_E +_{\mathbb{Z}} [(r,s)]_E &= [(p+r, q+s)]_E && \textit{addition} \\ [(p,q)]_E \times_{\mathbb{Z}} [(r,s)]_E &= [(pr+qs, ps+qr)]_E && \textit{multiplication} \\ -[(p,q)]_E &= [(q,p)]_E && \textit{negation} \end{aligned}$$

Here p, q, r, and s all lie in $\mathbb{N}$. ◇

Note that these definitions all operate in a similar fashion. They claim to define functions from $\mathbb{Z}^2$ to $\mathbb{Z}$, but in fact they are functions from $\mathbb{N}^2 \times \mathbb{N}^2$ to $\mathbb{Z}$. This is because they "look inside" the equivalence classes, choose ordered pairs, and compute. Only then do they form a new equivalence class.

Suppose, as an example, that we wish to compute $3_{\mathbb{Z}} +_{\mathbb{Z}} 4_{\mathbb{Z}}$. (Recall that $k_{\mathbb{Z}} = [(k,0)]_E$.) The two integers $3_{\mathbb{Z}}$ and $4_{\mathbb{Z}}$ both contain infinitely many ordered pairs. We might pick $(6,3)$ and $(6,2)$ to represent $3_{\mathbb{Z}}$ and $4_{\mathbb{Z}}$, respectively. We then form $[(12,5)]_E$ and note that this is $7_{\mathbb{Z}}$, as desired. However, what if we instead had picked $(10,7)$ and $(10,6)$ to represent $3_{\mathbb{Z}}$ and $4_{\mathbb{Z}}$? In that case we get $[(20,13)]_E$; luckily this is again equal to $7_{\mathbb{Z}}$. Being lucky is nice; however a more systematic method is available.

9.3. **Well-defined functions.**

Theorem 9.4. *Suppose that X and Y are sets. Suppose that E is an equivalence relation on X. Suppose that $f\colon X \to Y$ is a function. Suppose also that we have:*

$$(\text{INVAR}) \text{ for all } x, x' \in X, \text{ if } xEx' \text{ then } f(x) = f(x')$$

Then there is a unique function $f_E\colon X/E \to Y$ satisfying $f = f_E \circ q_E$.

Said another way, if the function $f\colon X \to Y$ satisfies the hypothesis (INVAR) then it *induces* a unique function $f_E\colon X/E \to Y$ with the desired property: $f = f_E \circ q_E$. Functions from $(X/E)^2$, and higher cartesian products, are induced in a similar way.

When mathematicians use Theorem 9.4 – the "universal property of quotients" – they often do not explicitly invoke it. Instead they

(1) define f (while pretending to define f_E),
(2) check the hypothesis (INVAR), and
(3) then declare that f_E is *well-defined*.

Proof of Theorem 9.4. We first build the function f_E and then prove it is unique.

Let Q be relation from X/E to Y which contains $([x]_E, y)$ exactly when

- there is some $x' \in X$ so that $f(x') = y$ and $x' \in [x]_E$.

Suppose that $[x]_E$ is any equivalence class. Note that $[x]_E$ is non-empty as it contains x. Let $y = f(x)$. Then $([x]_E, y)$ lies in Q.

Suppose that there is some $x \in X$ and $y', y'' \in Y$ so that $([x]_E, y')$ and $([x]_E, y'')$ lie in Q. By the defining property of Q there are x' and x'' in $[x]_E$ so that $f(x') = y'$ and $f(x'') = y''$. By the definition of $[x]_E$ we have xEx' and xEx''. By the symmetry of E we have $x'Ex$. By the transitivity of E we have $x'Ex''$. By the hypothesis (INVAR) we have $f(x') = f(x'')$, and so $y' = y''$. Thus Q is graphical, as desired. Taking $f_E = Q$ completes the construction.

Suppose that x is an element of X. Thus $(f_E \circ q_E)(x) = f_E([x]_E) = f(x)$ and so we have $f = f_E \circ q_E$.

Suppose that $f'_E\colon X/E \to Y$ is another function satisfying $f = f'_E \circ q_E$. Suppose that $[x]_E$ is an element of X/E. Then $f_E([x]_E) = f(x)$. Also, $f'_E([x]_E) = f'_E(q_E(x)) = f(x)$. Since f_E and f'_E have the same domain, codomain, and relation, they are equal. Thus f_E is unique. □

9.5. **Back to arithmetic.** We now show that the arithmetical operations on $\mathbb{Z}$ are induced by the operations on $\mathbb{N}^2$ given in Definition 9.2.

Lemma 9.6. *The arithmetical operations on $\mathbb{Z}$ are well-defined.*

Proof. Let E be the equivalence relation on $\mathbb{N}^2$ given in Example 8.12. Suppose that (p,q) and (p',q') are equivalent under the relation E. Suppose that (r,s) and (r',s') are equivalent under the relation E. Swapping as needed, we may assume that $p < p'$ and $r < r'$. Appealing to the solution to Exercise 8.13 we find that there are natural numbers P and R so that

$$(p+P, q+P) = (p', q') \quad \text{and} \quad (r+R, s+R) = (r', s')$$

To show that addition in $\mathbb{Z}$ is well-defined we compute as follows:

$$\begin{aligned}(p'+r', q'+s') &= (p+P+r', q+P+s') \\ &= (p+P+r+R, q+P+s+R) \\ &= (p+r+P+R, q+s+P+R)\end{aligned}$$

By the solution to Exercise 8.13 this last lies in $[(p+r, q+s)]_E$, as desired.

To show that multiplication in $\mathbb{Z}$ is well-defined we compute as follows:

$$\begin{aligned}(p'r'+q's', p's'+q'r') &= ((p+P)r'+(q+P)s', (p+P)s'+(q+P)r') \\ &= ((p+P)r'+(q+P)s', (p+P)s'+(q+P)r') \\ &= ((p+P)(r+R)+(q+P)(s+R), \\ &\quad (p+P)(s+R)+(q+P)(r+R)) \\ &= ((pr+qs+(p+q)R+(r+s)P+2PR, \\ &\quad ps+qr+(p+q)R+(r+s)P+2PR))\end{aligned}$$

By the solution to Exercise 8.13 this last lies in $[(pr+qs, ps+qr)]_E$, as desired.

To show that negation in $\mathbb{Z}$ is well-defined we compute as follows:

$$(q', p') = (q+P, p+P)$$

By the solution to Exercise 8.13 this last lies in $[(q,p)]_E$, as desired. □

Remark 9.7. With the operations of $\mathbb{Z}$ in place we can deduce their usual properties from the corresponding properties of addition and multiplication in $\mathbb{N}$. ◇

Recall that $k_{\mathbb{Z}} = [(k,0)]_E$. Here are two properties which are mysterious in some presentations, but which are transparent here.

Corollary 9.8. *We have the following:*

- $-0_{\mathbb{Z}} = 0_{\mathbb{Z}}$
- $(-1_{\mathbb{Z}}) \times_{\mathbb{Z}} (-1_{\mathbb{Z}}) = 1_{\mathbb{Z}}$

Proof. We obtain $-0_{\mathbb{Z}}$ by reversing the order of $(0,0)$. Since this does not change the ordered pair, we have $-0_{\mathbb{Z}} = 0_{\mathbb{Z}}$.

We now compute as follows:

$$\begin{aligned}
(-1_{\mathbb{Z}}) \times_{\mathbb{Z}} (-1_{\mathbb{Z}}) &= (-[(1,0)]_E) \times_{\mathbb{Z}} (-[(1,0)]_E) && \text{definition of } 1_{\mathbb{Z}} \\
&= [(0,1)]_E \times_{\mathbb{Z}} [(0,1)]_E && \text{definition of negation} \\
&= [(0 \times 0 + 1 \times 1, 0 \times 1 + 1 \times 0)]_E && \text{definition of } \times_{\mathbb{Z}} \\
&= [(1,0)]_E && \text{arithmetic in } \mathbb{N} \\
&= 1_{\mathbb{Z}} && \text{definition of } 1_{\mathbb{Z}}
\end{aligned}$$

and we are done. □

Here is a final corollary.

Corollary 9.9. *The function $\mathbb{N} \to \mathbb{Z}$ taking $k \mapsto k_{\mathbb{Z}} = [(k,0)]_E$ is injective and takes addition and multiplication in $\mathbb{N}$ to the corresponding operations in $\mathbb{Z}$.* □

With the integers safely defined, we may return to thinking of individual integers as numbers, rather than as equivalence classes. We also drop the subscripts from $+_{\mathbb{Z}}$ and $\times_{\mathbb{Z}}$, relying on context to remind us of the type of numbers we are adding or multiplying.

Challenge 9.10. Give a definition of $\mathbb{Q}$, the rational numbers, in the style of Example 8.12 and Definition 8.15. Use this to explain the claim that $1/2$ "equals" $2/4$. Prove that all non-zero elements of $\mathbb{Q}$ have a multiplicative inverse. Finally, for k in $\mathbb{Z}$ define $k_{\mathbb{Q}} = k/1$ in $\mathbb{Q}$. Prove that the function $k \mapsto k_{\mathbb{Q}}$ is injective and takes addition, multiplication, and negation in $\mathbb{Z}$ to the corresponding operations in $\mathbb{Q}$.

10. Modular arithmetic

10.1. **Multiples and divisors.**

Definition 10.2. Suppose that m and n are integers. We say that m is a *multiple* of n if there is an integer k so that $m = k \times n$. If m is a multiple of n then we also say that n *divides* m. ◇

Exercise 10.3. Show the following.

(1) The relation of divides on $\mathbb{Z}$ is reflexive and transitive, but not symmetric.
(2) Plus and minus one divide all integers.
(3) All integers divide zero.
(4) Zero divides only itself. ◇

10.4. **Modulus n.**

Definition 10.5. Suppose that a, b, and n are integers. We say that a and b are *congruent modulo* n if there is some integer k so that $a = b + kn$. That is, $a - b$ is a multiple of n. ◇

Thus congruence modulo n gives a relation on $\mathbb{Z}$.

Notation 10.6. Suppose that a, b, and n are integers. If a and b are congruent modulo n then we may write $a \equiv b \pmod{n}$. We call n the *modulus* of the congruence. ◇

Exercise 10.7. Prove that $a \equiv b \pmod{n}$ is an equivalence relation on $\mathbb{Z}$. ◇

Notation 10.8. The equivalence classes of $a \equiv b \pmod{n}$ are often called *congruence classes*. If a is an integer then we write $[a]_n$ for the congruence classes modulo n containing a. We also write

$$\mathbb{Z}/n\mathbb{Z} = \{[a]_n \mid a \in \mathbb{Z}\}$$

for the set of congruence classes.[9] Finally we use $q_n : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ for the quotient map. ◇

The notation $\mathbb{Z}/n\mathbb{Z}$ is pronounced "$\mathbb{Z}$ mod n $\mathbb{Z}$" or "$\mathbb{Z}$ modulo n $\mathbb{Z}$".

Exercise 10.9. Prove the following.

- The equivalence classes of $\mathbb{Z}/0\mathbb{Z}$ are singletons.
- The quotient $\mathbb{Z}/1\mathbb{Z}$ contains exactly one equivalence class.
- For any integer n we have $\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/(-n)\mathbb{Z}$. ◇

[9] The notation $\mathbb{Z}/n\mathbb{Z}$ comes from algebra, where we may form a "quotient" of a *group* (or *ring*) by one of its *subgroups* (or *ideals*).

We now count the number of elements of $\mathbb{Z}/n\mathbb{Z}$: that is, the number of equivalence classes of $a \equiv b \pmod{n}$.

Exercise 10.10. In light of Exercise 10.9 we may suppose that n is a positive integer.

- Suppose that m is an integer. Prove that $[m-n]_n = [m]_n = [m+n]_n$.
- Suppose that r and r' lie in $[\![n]\!]$. Prove that $r = r'$ if and only if $[r]_n = [r']_n$.
- Sketch a proof that $\mathbb{Z}/n\mathbb{Z}$ has exactly n elements: that is, $a \equiv b \pmod{n}$ has exactly n equivalence classes. ◇

As a concrete example we explore the equivalence classes of the integers modulo $n = 6$. Suppose that a lies in $\mathbb{Z}$. Then we have

$$[a]_6 = \{a + 6k \in \mathbb{Z} \mid k \in \mathbb{Z}\}$$

To "see" this, we draw $\mathbb{Z}$ along an logarithmic spiral in $\mathbb{C}$, the complex plane, as shown in Figure 10.11. The points of the equivalence class $[a]_6$ lie along the ray from the origin and through the point $\exp(a\frac{2\pi i}{6})$.

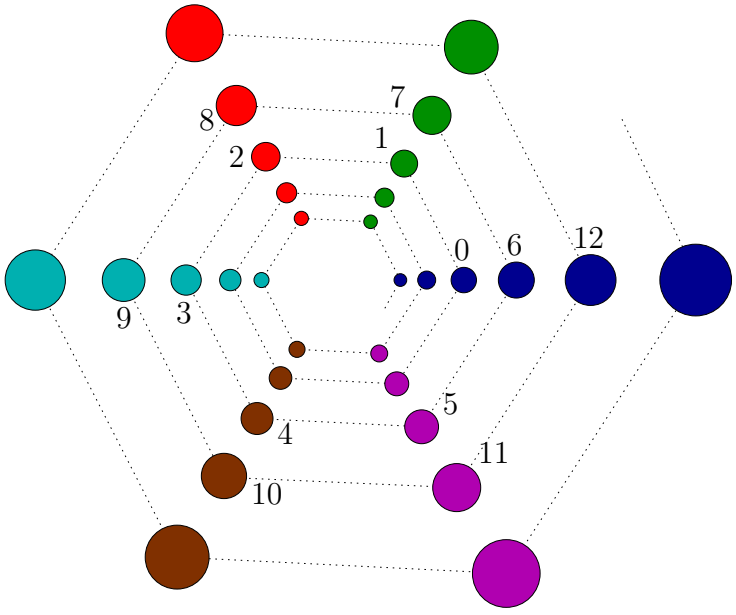


FIGURE 10.11. The equivalence classes of $a \equiv b \pmod{6}$ (Definition 10.5).

10.12. **Modular addition, multiplication, and negation.**

Definition 10.13. Suppose that n is an integer. We call the congruence classes $[0]_n$ and $[1]_n$ the *zero* and *one* of $\mathbb{Z}/n\mathbb{Z}$. We define the following

operations on $\mathbb{Z}/n\mathbb{Z}$.

$$\begin{aligned} [a]_n +_n [b]_n &= [a+b]_n && \textit{addition} \\ [a]_n \times_n [b]_n &= [a \times b]_n && \textit{multiplication} \\ -_n [a]_n &= [-a]_n && \textit{negation} \end{aligned}$$

Here a and b are integers. ◇

Lemma 10.14. *The arithmetical operations in $\mathbb{Z}/n\mathbb{Z}$, given in Definition 10.13, are well-defined.* □

The proofs are similar to, but simpler than, those of Lemma 9.6 and we omit them.

Corollary 10.15. *The quotient function $q_n : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ takes addition, multiplication, and negation in $\mathbb{Z}$ to the corresponding operations in $\mathbb{Z}/n\mathbb{Z}$.* □

Exercise 10.16.

(1) Find the unique integer between zero and six (inclusive) congruent to $3^{20} \pmod 7$.
(2) Find the unique integer between zero and six (inclusive) congruent to $2^{111} \pmod 7$.
(3) Find the unique integer between zero and six (inclusive) congruent to $2^{3^{20}} \pmod 7$. ◇

11. Booleans and their operators

11.1. **Booleans.** We introduce two new symbols T and F.

Definition 11.2. A *boolean* is an element of the set $\mathcal{B} = \{\mathrm{T}, \mathrm{F}\}$. ◇

One common interpretation of the symbols T and F is "true" and "false".

11.3. **Boolean operators.**

Definition 11.4. Suppose that n is a natural number. Suppose that $\mathcal{B} = \{\mathrm{T}, \mathrm{F}\}$. Recall that $\mathcal{B}^n$ is the cartesian product of $\mathcal{B}$ with itself, n times. We call a function $f : \mathcal{B}^n \to \mathcal{B}$ a *boolean operator*. The number n is called the *arity* of f. ◇

For example, the identity $\mathrm{Id}_{\mathcal{B}}$ is a boolean operator with arity one. There are two boolean operator with arity one which are *constant*; one always returns T and the other always returns F. Here is the final operator with arity one.

Definition 11.5. Suppose that P is a boolean. We define $(\neg P)$ via the following table.

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

We call $(\neg P)$ the *negation* of P. It is pronounced "not P". $\diamond$

Definition 11.6. Suppose that P and Q are booleans. We define the following operators; all have arity two.

P	Q	$(P \vee Q)$	$(P \wedge Q)$	$(P \to Q)$	$(P \leftrightarrow Q)$
T	T	T	T	T	T
T	F	T	F	F	F
F	T	T	F	T	F
F	F	F	F	T	T

The operators $\vee$, $\wedge$, $\to$, and $\leftrightarrow$ are called, respectively, the *disjunction*, the *conjunction*, the *implication*, and the *equivalence* operators. These are, in order, pronounced as "or", "and", "implies", and "is equivalent to" $\diamond$

Remark 11.7. There are twelve more boolean operators of arity two. It is an amusing exercise to list them all and give them their names. $\diamond$

We call the operators $\neg$, $\vee$, $\wedge$, $\to$, and $\leftrightarrow$ the *basic* boolean operators. We may compose these operators to obtain more complicated ones, with possibly higher arities.

Example 11.8. In these examples, the arity is two; this is because the arity counts the number of independent inputs to the operator.

- $((P \wedge P) \wedge (Q \wedge Q))$
- $((P \to (\neg P)) \to Q)$
- $((P \leftrightarrow (\neg P)) \leftrightarrow (Q \leftrightarrow (\neg Q)))$ $\diamond$

Suppose that f is a boolean operator. Any given composition of the basic operators, giving f, is called an *expression* for f. In an expression any sub-expression (beginning and ending with a pair of matching parentheses) is called a *clause*. A clause that does not contain parentheses is necessarily one of the basic operators. As an example, the expression

$$((P \wedge P) \wedge (P \vee P))$$

has three clauses of lengths five, five, and thirteen. This composition is a somewhat lengthy expression of the operator $\mathrm{Id}_{\mathcal{B}}$.

We note, but do not prove, that the five basic operators are a bit more than is strictly "needed".

Theorem 11.9. *Any boolean operator can be written as a composition of the operators* $\neg$ *and* $\vee$. □

As a few very special cases of the theorem, we note the following equalities.

$$(P \wedge Q) = (\neg((\neg P) \vee (\neg Q)))$$
$$(P \to Q) = ((\neg P) \vee Q)$$
$$(P \leftrightarrow Q) = ((P \to Q) \wedge (Q \to P))$$

Using these we can express $\leftrightarrow$ solely in terms of $\neg$ and $\vee$.

Exercise 11.10. Simplify the compositions given in Example 11.8. Then write each in terms of $\neg$ and $\vee$. ◇

11.11. **Boolean algebra.** The boolean operators $\vee$ and $\wedge$ have many features in common with the operations $\cup$ and $\cap$. The following are modelled on Lemmas 5.7, 5.14, 5.20, and 5.22

Lemma 11.12. *Suppose that* P, Q, *and* R *are booleans. Then we have the following.*

(1) $(P \vee \mathrm{T}) = \mathrm{T}$ *and* $(P \vee \mathrm{F}) = P$ *(identity)*
(2) $((P \vee Q) \vee R) = (P \vee (Q \vee R))$ *(associativity)*
(3) $(P \vee Q) = (Q \vee P)$ *(commutativity)*
(4) $(P \to Q) = T$ *if and only if* $(P \vee Q) = Q$ *(absorption)*
(5) $(P \vee P) = P$ *(idempotent)*

Exercise 11.13. Provide the proof of Lemma 11.12 ◇

We omit the remaining proofs.

Lemma 11.14. *Suppose that* P, Q, *and* R *are booleans. Then we have the following.*

(1) $(P \wedge \mathrm{T}) = P$ *and* $(P \wedge \mathrm{F}) = \mathrm{F}$ *(identity)*
(2) $((P \wedge Q) \wedge R) = (P \wedge (Q \wedge R))$ *(associativity)*
(3) $(P \wedge Q) = (Q \wedge P)$ *(commutativity)*
(4) $(P \to Q) = T$ *if and only if* $(P \wedge Q) = P$ *(absorption)*
(5) $(P \wedge P) = P$ *(idempotent)* □

Lemma 11.15 (Distributive laws)**.** *Suppose that* P, Q, *and* R *are booleans. Then we have the following.*

- $(P \wedge (Q \vee R)) = ((P \wedge Q) \vee (P \wedge R))$
- $(P \vee (Q \wedge R)) = ((P \vee Q) \wedge (P \vee R))$ □

Lemma 11.16 (De Morgan's laws)**.** *Suppose that* P *and* Q *are booleans. Then we have the following.*

- $(\neg(P \vee Q)) = ((\neg P) \wedge (\neg Q))$

- $(\neg(P \wedge Q)) = ((\neg P) \vee (\neg Q))$ □

11.17. **Tautologies.**

Definition 11.18. Suppose that n is a natural number. Suppose that $\mathcal{B} = \{\mathrm{T}, \mathrm{F}\}$. Suppose that $f : \mathcal{B}^n \to \mathcal{B}$ is a boolean operator. We call f a *tautology* if $f(x) = \mathrm{T}$ for all $x \in \mathcal{B}^n$. We call f an *antinomy* if $f(x) = \mathrm{F}$ for all $x \in \mathcal{B}^n$. ◇

For an example, consider $f(P) = (P \to P)$. This is a boolean operator of arity one. If $P = \mathrm{T}$ then we obtain $f(\mathrm{T}) = (\mathrm{T} \to \mathrm{T}) = \mathrm{T}$, according to Definition 11.6. If $P = \mathrm{F}$ then we obtain $f(\mathrm{F}) = (\mathrm{F} \to \mathrm{F}) = \mathrm{T}$. Thus f is a tautology. Similarly, $g(P) = (P \leftrightarrow (\neg P))$ is an antinomy.

Lemma 11.19. *Suppose that f and g are boolean operators of the same arity. Then $(f \leftrightarrow g)$ is a tautology if and only if $f = g$.* □

In particular, if f and g are given as different compositions of the basic operators, then we may freely replace the one expression by the other. Thus tautologies give us ways to simplify boolean operators expressed as compositions.

We note that Lemma 11.19, added to the lemmas of Section 11.11, give us a large supply of tautologies. Here are a few more, that seem to arise in practice.

$((\neg(\neg P)) \leftrightarrow P)$	double negation
$((P \to Q) \leftrightarrow ((\neg Q) \to (\neg P)))$	contraposition
$((P \to Q) \leftrightarrow ((\neg P) \vee Q))$	definition of implication
$((P \leftrightarrow Q) \leftrightarrow ((P \to Q) \wedge (Q \to P)))$	definition of equivalence

And here are some of the tautologies graced with the name "rules of deduction".

$(P \vee (\neg P))$	law of the excluded middle
$((P \wedge (P \to Q)) \to Q)$	modus ponens
$(((P \to Q) \wedge (Q \to R)) \to (P \to R))$	transitivity of implication
$(((\neg P) \to \mathrm{F}) \to P)$	argument by contradiction

12. Truth tables

Here we discuss *truth tables*; an algorithmic method for dealing with boolean operators expressed as a composition of the basic operators. Before we do so, we require an ordering on the set of binary strings of a fixed length.

12.1. **Ordering binary strings.** Suppose that n is a natural number. Suppose that u and v are binary strings of length n. Suppose that u_i is the i^{th} bit of u; we define v_i similarly.

We now say that $u \leq v$ if either

- $u = v$ or
- there is some k so that $u_i = v_i$ for $i < k$ and $u_k = 0$ while $v_k = 1$.

This is called the *lexicographic* order on binary strings of a fixed length. Said another way, $u \leq v$ if at the first position k where they disagree we have $u_k < v_k$.

Here is the lexicographic ordering of the binary strings of length three.

$$000 \leq 001 \leq 010 \leq 011 \leq 100 \leq 101 \leq 110 \leq 111$$

Here are a few examples in length five.

$$00111 \leq 01000 \qquad 01000 \leq 01101 \qquad 11110 \leq 11111$$

Exercise 12.2. Give the lexicographic order on binary strings of length four. ◇

We may transfer this to obtain an ordering on the elements of $\{\mathrm{T}, \mathrm{F}\}^n$ by making an identification of T with zero and F with one. Here is the resulting lexicographic order on strings over $\{\mathrm{T}, \mathrm{F}\}$ of length three.

$$\mathrm{TTT} \leq \mathrm{TTF} \leq \mathrm{TFT} \leq \mathrm{TFF} \leq \mathrm{FTT} \leq \mathrm{FTF} \leq \mathrm{FFT} \leq \mathrm{FFF}$$

12.3. **Lookup tables.** The *lookup table* for a boolean operator f of arity n is

- the list of strings ω of length n over $\{\mathrm{T}, \mathrm{F}\}$ and
- next to each ω in the list the boolean $f(\omega)$.

Examples have almost already appeared in Definitions 11.5 and 11.6.

12.4. **An algorithm to produce lookup tables.** Suppose that f is a boolean operator of arity n. Suppose that f is given to us as a composition of the basis operators. To be concrete, we use $f = ((P \vee Q) \to (Q \wedge P))$ as a running example. To form the lookup table for f we must compute $f(\mathrm{T}, \mathrm{T})$, $f(\mathrm{T}, \mathrm{F})$, $f(\mathrm{F}, \mathrm{T})$, and $f(\mathrm{F}, \mathrm{F})$ and record them, in that order. To save effort and space, we do not compute these one at a time. Instead we compute them together, in a *truth table*, as follows.

Algorithm 12.5.

(1) We make a table with five $(= 1 + 2^n)$ rows, with f in the first row. We leave all columns blank, except for the columns below

P and Q. Those we fill using the values for P and Q generated by lexicographic order.

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

(2) For each innermost clause (those that do not contain further parentheses), and for each row (so the values of P and Q are fixed), we use Definitions 11.5 and 11.6 to fill in the value of the clause.

$((P$	$\vee$	$Q)$	$\to$	$(Q$	$\wedge$	$P))$
T	**T**	T		T	**T**	T
T	**T**	F		F	**F**	T
F	**T**	T		T	**F**	F
F	**F**	F		F	**F**	F

The new entries in the table are in bold face.

(3) We now repeat the process for the clauses only containing innermost clauses, using the computed values of the innermost.

$((P$	$\vee$	$Q)$	$\to$	$(Q$	$\wedge$	$P))$
T	T	T	**T**	T	T	T
T	T	F	**F**	F	F	T
F	T	T	**F**	T	F	F
F	F	F	**T**	F	F	F

Again, the new entries in the table are in bold face.

(4) In general, we repeat the previous step; in its k^{th} iteration we assign values to clauses containing only clauses dealt with at step $k-1$ or earlier.

This completes the algorithm. $\diamond$

Here then is the lookup table for our running example.

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

We deduce that the boolean operator is equal to $(P \leftrightarrow Q)$.

Exercise 12.6. Use Algorithm 12.5 to prove that modus ponens

$$((P \wedge (P \to Q)) \to Q)$$

is a tautology. ◇

13. Quantifiers and their laws

13.1. **Quantifiers.** Here we explore two kinds of statements: the *universal* and the *existential*.

We begin with a running example from Euclid's *Elements* [5, Book 9, Proposition 20]. We give the statement in a slight nonstandard form.[10]

Theorem 14.2. *There are arbitrarily large primes.*

We postpone the proof. We rephrase Theorem 14.2 as follows:

(EUC) For every natural number n there exists a natural number p so that $p > n$ and p is prime.

We introduce a function $\textsc{Prime}\colon \mathbb{N} \to \{\mathrm{T}, \mathrm{F}\}$ defined by $\textsc{Prime}(n) = \mathrm{T}$ exactly when n is prime. With this definition we can rewrite our condition as follows.

(EUC) For every natural number n there exists a natural number p so that $((p > n) \wedge \textsc{Prime}(p))$.

To finish rewriting (EUC) we require a bit of notation and a definition.

Notation 13.2. Suppose that X is a set. We may abbreviate the phrase *for all* $x \in X$ as $\forall x \in X$. We may abbreviate the phrase *there exists* $x \in X$ as $\exists x \in X$.

In either case, if the set X is fixed throughout the discussion, we may simplify the notation to just $\forall x$ and $\exists x$. ◇

Note that $\forall$ is an upside-down "A" and $\exists$ is a backwards "E". These logical operators are called *quantifiers*. The fragment $\forall x$ may be pronounced as "for all x", "for every x", "for any x", "for each x", or similar. The fragment $\exists x$ may be pronounced as "there exists x", "for some x", "for at least one x", or similar.

Definition 13.3. Suppose that X is a set. Suppose that $S(x)$ is a property. Then the sentence $(\forall x \in X \;\; S(x))$ holds if and only if $S(x)$ holds for every x in X. On the other hand, the sentence $(\exists x \in X \;\; S(x))$ if and only if $S(x)$ holds for some x in X. Accordingly we call $\forall$ the *universal* quantifier and we call $\exists$ the *existential* quantifier. ◇

We can now finish rewriting our condition (with the background set understood to be the natural numbers).

(EUC) $(\forall n\, (\exists p\, ((p > n) \wedge \textsc{Prime}(p))))$

[10] This is because we wish to focus on quantifiers ranging over a fixed set (here the natural numbers) rather than quantifiers ranging over sets.

Thus do we replace clarity by concision. We justify this by noting that shorter expressions are far easier to manipulate formally.[11]

Exercise 13.4. Translate the following sentences into English and determine whether or not they hold. For each one the background set is the natural numbers.

- $(\forall p\,(\exists n\,((p > n) \wedge \textsc{Prime}(p))))$
- $(\exists n\,(\forall p\,((p > n) \wedge \textsc{Prime}(p))))$
- $(\exists p\,(\forall n\,((p > n) \wedge \textsc{Prime}(p))))$ ◇

As the exercise shows, the order of quantifiers is a delicate matter. That is, the sentences

$$(\forall x\,(\exists y\ S(x, y))) \qquad \text{and} \qquad (\exists y\,(\forall x\ S(x, y)))$$

are only very rarely equivalent.

13.5. **Negation and distribution for quantifiers.** There is a duality between the universal and existential quantifiers, as follows.

Theorem 13.6. *Suppose that X is our background set. Suppose that $S(x)$ is a property. Then*

- $(\neg(\forall x\ S(x)))$ *holds if and only if* $(\exists x\,(\neg S(x)))$ *holds.*
- $(\neg(\exists x\ S(x)))$ *holds if and only if* $(\forall x\,(\neg S(x)))$ *holds.*

Proof. Suppose that $(\neg(\forall x\ S(x)))$ holds. Thus $(\forall x\ S(x))$ does not hold. So $S(x)$ does not hold for some x in X. Thus $(\neg S(x))$ does hold for some x in X. That is, $(\exists x\,(\neg S(x)))$ holds.

Each step of the above argument reverses; thus the first equivalence is proven.

To prove the second, negate both sentences and recall that $(\neg(\neg P))$ is equivalent to P. □

Theorem 13.7. *Suppose that X is our background set. Suppose that $R(x)$ and $S(x)$ are properties. Then*

- $(\forall x\,(R(x) \wedge S(x)))$ *holds if and only if* $((\forall x\ R(x)) \wedge (\forall x\ S(x)))$ *holds.*
- $(\exists x\,(R(x) \vee S(x)))$ *holds if and only if* $((\exists x\ R(x)) \vee (\exists x\ S(x)))$ *holds.*

On the other hand, there are no "distributive laws" for $\forall$ over $\vee$ or for $\exists$ over $\wedge$. We omit the proofs.

[11] To see the truth of this, try multiplying one hundred and twelve by eighty-three *without* using some form of place notation.

Proof of Theorem 13.7. Suppose that $(\forall x\,(R(x) \wedge S(x)))$. Then for every x in X we have that $(R(x) \wedge S(x))$ holds. That is, for every x in X we have that $R(x)$ holds and $S(x))$ holds. So, for every x in X we have that $R(x)$ holds and, for every x in X, we also have that $S(x))$ holds. That is, $(\forall x\ R(x))$ holds as does $(\forall x\ S(x))$. Thus $((\forall x\ R(x)) \wedge (\forall x\ S(x)))$ holds.

Every step of the above argument reverses; thus the first equivalence is proven.

To prove the second, negate both sides, apply Theorem 13.6, and recall that $R(x)$ and $S(x)$ were arbitrary properties. □

Exercise 13.8.

- Prove that $(\forall x\,(R(x) \to S(x)))$ implies $((\forall x\ R(x)) \to (\forall x\ S(x)))$.
- Show by means of an example that the latter need not imply the former. ◇

14. Proof

14.1. **An example.**

Theorem 14.2. *There are arbitrarily large primes.*

The more usual statement is "There are infinitely many primes." The original statement in Euclid's Elements is "Prime numbers are more than any assigned multitude of prime numbers." [5, Book 9, Proposition 20]. But actually that is false: the actual "original" statement was written in Greek, probably on papyrus, perhaps within a few decades of 300 BCE. No trace of it remains. The oldest surviving complete copy of Euclid's Elements is instead written on parchment and dates to 888 CE [6].

Before giving the proof of Theorem 14.2, we require a definition and two preliminary results.

Definition 14.3. Suppose that p is a natural number greater than one. Then p is *prime* if it has exactly two positive divisors. ◇

Since p is divisible by one, and by itself, these are all of its (positive) divisors. Here are the needed "helper" statements.

Lemma 17.12. *Suppose that X is a finite set. Suppose that Y is a subset of X. Then Y is finite.*

Proposition 16.13. *Suppose that n is a natural number greater than one. Then n is divisible by some prime.*

Proof of Theorem 14.2. Recalling the discussion of Section 13.1, we begin by fixing a natural number n. We now must find a prime p larger than n.

Let $\mathcal{P}$ be the set of primes less than or equal to n. By Lemma 17.12, the set $\mathcal{P}$ is finite. So let $(p_i)_{i=0}^{k-1}$ be the list of elements of $\mathcal{P}$, say in order of size. We define N to be one, plus the product of these primes. That is:

$$N = 1 + p_0 \cdot p_1 \cdot p_2 \cdot \cdots \cdot p_{k-1}$$

Suppose that p lies in $\mathcal{P}$. Thus N is congruent to one modulo p. In particular N is not divisible by p.

By Proposition 16.13, the natural number N is divisible by some prime, say q. We deduce that q does not lie in $\mathcal{P}$. From the definition of $\mathcal{P}$ we deduce that q is greater than n. □

There are several things to note about the proof.

- The overall plan of the proof follows the previous discussion in Section 13.1.
- The existence of $\mathcal{P}$ follows from the axiom of specification applied to $[\![n+1]\!]$.
- The finiteness of $\mathcal{P}$ depends on Lemma 17.12. Although this lemma is "obvious" we have not proved it yet.
- Likewise, the claim that N is divisible by a prime depends on Proposition 16.13. Again we have not proved this yet.

Each of these points is interesting in its own right. Taking them altogether, they suggest that different proofs of Theorem 14.2 can have very different levels of detail.

However it is *not* the case that writing down a proof is "merely" a matter of being "formal enough". A proof may also require some act of imagination. As one piece of evidence for this we note that the definition of N, above, was a bit subtle. If this is not convincing enough, consider the following exercise.

Exercise 14.4. Prove that there are arbitrarily large primes that are congruent to 3 modulo four. ◇

14.5. **Proof, informally.** With our example in hand, we attempt an informal definition.

Definition 14.6. A *proof* is a convincing argument. ◇

This definition lays bare the central issue; for an argument to be convincing, there is some person that needs to be convinced. However, an argument that convinces one person may not convince another. In one respect this is obvious; somebody who does not read English will

not find the proof of Theorem 14.2, or indeed anything in these lecture notes, enlightening.

Next, there are rhetorical issues. For example, a writer may signal the difficulty of steps in a proof. Phrases such as "it is easy to see" or "a simple argument gives" are common. Less common, but still used are phrases like "it is a deep result that" or "the crucial step is". However, there is a blurry area between helpful advice on the one hand and misleading propaganda on the other. Both may be convincing, but we should try to avoid the latter.

Finally, there are subtle problems of representation. While writing these notes I have assumed that you, the reader, has an excellent background at A-levels maths. But perhaps my conception of A-level maths is incorrect. Or perhaps your preparation is stronger in some areas and weaker in others; as a result some of the arguments here may be more or less convincing.

Taking the above definition seriously leads us to a somewhat awkward conclusion; that "proof" is a social construct that lives somewhere between the minds of the writer and the reader.

14.7. **Formal proof.** We now try to minimise the sociological nature of "proof". (This section is not examinable.)

Definition 14.8. Suppose that $\mathcal{A}$ is a list of axioms. Suppose that $\mathcal{D}$ is a list of rules of deduction. Furthermore, suppose that there are algorithms that decide if a given statement lies in $\mathcal{A}$ (or not) and if a given statement can be deduced, by a given rule in $\mathcal{D}$, from a given list of other statements.

Suppose now that S is a statement. Then a *proof of S from $\mathcal{A}$ and $\mathcal{D}$* is a list $(S_i)_{k=0}^{n-1}$ of statements where:

- For each k, either
  - S_k is an axiom (lies in $\mathcal{A}$) or
  - S_k follows from the statements $(S_i)_{i<k}$ by one of the rules of deduction (from $\mathcal{D}$) and
- $S_{n-1} = S$. ◇

This more formal definition captures many features of mathematical practice. It is also a valuable tool in making proofs first-class mathematical objects. Done carefully, proofs written this way can be checked by computer, using tools such as *proof assistants*. Finally, and perhaps most significantly, Gödel started from the idea of formal proof and showed that (for very general $\mathcal{A}$ and $\mathcal{D}$) there are statements that can neither be proven nor disproven.

Definition 14.8 is the model for how mathematicians think of proofs. However it is not, for the most part, what they actually *do*. Mathematicians use diagrams, draw pictures, collect evidence (by hand or via computer), make conjectures, stare into space, and argue with each other a lot, often waving their hands about. It is perhaps possible to restate some of these activities in terms of Definition 14.8. But that is not the end goal of most mathematicians.

15. Patterns of proof

15.1. **Direct proof of implications.** Many mathematical statements are written in the form "if P then Q" where again P and Q are mathematical statements.

Definition 15.2. The statement 'if P then Q" is an the *implication* with *hypothesis* P and *conclusion* Q. ◇

Here is an example.

Lemma 15.3. *If n is an even integer then n^2 is an even integer.*

One way to prove an implication is *directly*. That is, we assume P, make various deductions, and finally deduce Q.

Proof of Lemma 15.3. Suppose that n is an even integer. So there is some integer k with $n = 2k$. Thus $n^2 = 4k^2 = 2 \times 2k^2$. Thus n^2 is even, as desired. □

Exercise 15.4. Give a direct proof of the implication "If n is an odd integer then n^2 is an odd integer." ◇

Definition 15.5. The *converse* of the implication "if P then Q" is the implication "if Q then P". ◇

In general if an implication holds, its converse may or may not hold.

Definition 15.6. We say that P and Q are *equivalent*, and we may write "P if and only if Q" when P implies Q and Q implies P. ◇

Note that to prove an equivalence we must prove a pair of implications.

Exercise 15.7. Give a direct proof of the following fragment of Lemma 5.14. Suppose that X and Y are sets. If $X \subset Y$ then $X \cap Y = X$. ◇

Exercise 15.8. Give a direct proof of the following implication: Suppose that X, Y, and Z are sets. Suppose that $f\colon X \to Y$ and $g\colon Y \to Z$ are bijections. Then $g \circ f$ is a bijection. ◇

15.9. **Proof by contraposition.**

Definition 15.10. The *contrapositive* of the implication "if P then Q" is the implication "if not Q then not P". ◇

As discussed in Section 11.17, the statement

$$((P \to Q) \leftrightarrow ((\neg Q) \to (\neg P)))$$

is a tautology. Thus to prove an implication it suffices to prove its contrapositive.

Suppose that k is an integer. We say that k is a *sum of two squares* if there are integers x and y so that $k = x^2 + y^2$.

Lemma 15.11. *Suppose that p is an integer. If $p \equiv 3 \pmod 4$ then p is not a sum of two squares.*

Proof. We prove the contrapositive, directly. Suppose that p is a sum of two squares. So there are integers x and y so that $p = x^2 + y^2$. There are now several cases: either x and y are both even, both odd, or have opposite parity.

Suppose that x and y are both even. So we have integers z and w so that $x = 2z$ and $y = 2w$. Thus $x^2 = 4z^2$ and $y^2 = 4w^2$ are both divisible by four. That is $x^2 \equiv y^2 \equiv 0 \pmod 4$. Thus $p \equiv x^2 + y^2 \equiv 0 \pmod 4$.

Suppose that x and y are both odd. So we have integers z and w so that $x = 2z+1$ and $y = 2w+1$. Thus $x^2 = 4z^2+4z+1$ and $y^2 = 4w^2+4w+1$ are both congruent to one modulo four. Thus $p \equiv x^2 + y^2 \equiv 2 \pmod 4$.

Finally suppose that x and y have opposite parity. Breaking the symmetry, we assume that x is even and y is odd. So we have integers z and w so that $x = 2z$ and $y = 2w + 1$. Using the same logic as above, we find that $p \equiv x^2 + y^2 \equiv 1 \pmod 4$.

In all cases we have that p is not congruent to three modulo four. □

Note that in the above proof the outermost pattern was a proof by contraposition. However, as an inner pattern we had a *proof by cases*. This is a way to organise work; we announce the complete list of cases and then dispose of each in order.

Exercise 15.12. Suppose that p is an integer. Prove that if $p \equiv 7 \pmod 8$ then p is not a sum of three squares. [Hint: if $p = x^2 + y^2 + z^2$ then consider x modulo four.] ◇

15.13. **Proof by contradiction.** We now have a result even older than Theorem 14.2. (This was known to the cult of Pythagoras [16]. However, it was likely also known a thousand years earlier [3].)

Theorem 15.14. *There is no rational number r so that $r^2 = 2$.*

This is a *non-existence statement*. To prove it, we use *proof by contradiction*. That is, we assume its negation and derive a contradiction. This is very similar to the tautology

$$(((\neg P) \to \mathrm{F}) \to P)$$

Proof of Theorem 15.14. Suppose, for a contradiction, that there is some rational number r so that $r^2 = 2$. By definition, we have that $r = p/q$ where p and q are integers, with q positive. Since $r^2 = (-r)^2$ we may also assume that p is positive.

We now give a *proof by infinite descent*. Take $p_0 = p$ and $q_0 = q$. Square both sides to find $p_0^2/q_0^2 = r^2 = 2$. Thus $p_0^2 = 2q_0^2$. Since p_0^2 is even, by the contrapositive of Exercise 15.4 we have that p_0 is even. We write $p_0 = 2p_1$ and find that $4p_1^2 = 2q_0^2$. Thus $q_0^2 = 2p_1^2$ and, again by Exercise 15.4, we have that q_0 is even. We write $q_0 = 2q_1$ and find that $4q_1^2 = 2p_1^2$. Thus $p_1^2 = 2q_1^2$.

Repeating the argument, we obtain sets $\{p_k\}_{k\in\mathbb{N}}$ and $\{q_k\}_{k\in\mathbb{N}}$ so that

$$p_k^2 = 2q_k^2 \qquad p_{k+1} = 2p_k \qquad \text{and} \qquad q_{k+1} = 2q_k$$

We deduce that $p_k > p_{k+1}$. Thus $\{p_k\}_{k\in\mathbb{N}}$ is a set of natural numbers with no least element. This contradicts the well-ordering principle (Theorem 16.12). □

Exercise 15.15. Prove that there is no rational number r so that $r^2 = 3$. ◇

Exercise 15.16. Prove that there is no rational number r so that $r^3 = 2$. ◇

15.17. **Proof by construction.**

Lemma 15.18. *There is a prime greater than 200.*

This is an *existence statement*. Note that Theorem 14.2 implies this, but without giving a explicit prime. In a *proof by construction* we give an explicit example and verify that it has the desired properties.

Proof of Lemma 15.18. We claim that $N = 211$ is prime. Suppose not; then N factors as a product of, say, $k \geq 2$ primes. If all of these primes are greater than $\sqrt{N} \sim 14.5$ then we conclude that $N > (\sqrt{N})^k \geq N$, a contradiction. Thus to prove N is prime it suffices to check if any of the primes less than 15 divide N.

The list of primes less than 15 is $(2, 3, 5, 7, 11, 13)$. For each we give its largest multiple which is still less than N.

$$210 = 2 \times 105 \quad 210 = 3 \times 70 \quad 210 = 5 \times 42$$

$$210 = 7 \times 30 \quad 209 = 11 \times 19 \quad 208 = 13 \times 16$$

Thus 211 is congruent to one modulo two, three, five, and seven. Similarly 211 is congruent to two modulo eleven and congruent to three modulo thirteen. In particular 211 is not divisible by any of the given primes, so we are done. □

The outermost pattern in the above is a proof by construction. However the inner steps are themselves proofs, using other patterns.

Exercise 15.19. Give a proof by construction of the following: there is a number less than 1000 having at least four distinct prime factors. ◇

Exercise 15.20. Prove that there is an injection from $\mathbb{Q}$ to $\mathbb{N}$. ◇

16. Induction and recursion

Very informally, mathematicians use *induction* to prove things and use *recursion* to define things.

16.1. **Induction.**

We now state the *principle of induction*.[12]

Axiom 16.2. *Suppose that $(P(n))_{n\in\mathbb{N}}$ is a sequence of sentences, one for each natural number. Suppose that*

(1) $P(0)$ holds and
(2) for all $k \in \mathbb{N}$, if $P(k)$ holds then $P(k+1)$ holds.

Then $P(n)$ holds for all n. □

In the above definition, part (1) is often called *the base case*. Part (2) is often called the *induction step*; also, in part (2) the sentence $P(k)$ is often called the *induction hypothesis*.

Proofs using Axiom 16.2 have a standard form, as follows.

Lemma 16.3. *Every natural number is either even or odd.*

Proof. Suppose that n is a natural number. Let $P(n)$ be the sentence "n is even or n is odd".

We proceed by induction. In the base case, we take $n = 0$. Since $0 = 2 \times 0$, the number zero is even. Thus $P(0)$ holds.

In the induction step we fix $k \in \mathbb{N}$. We assume the induction hypothesis $P(k)$. So k is even or odd. We must prove that $k+1$ is even or odd. There are two cases.

- Suppose that k is even. Thus there is a natural number ℓ so that $k = 2\ell$. So $k + 1 = 2\ell + 1$ is odd.
- Suppose that k is odd. Thus there is a natural number ℓ so that $k = 2\ell + 1$. So $k + 1 = 2\ell + 2 = 2(\ell + 1)$ is even.

Thus $k+1$ is odd or even. This implies (by the commutativity of "or") the sentence $P(k+1)$. Thus $P(k)$ implies $P(k+1)$, establishing the induction step.

By induction, $P(n)$ holds for all $n \in \mathbb{N}$. □

Exercise 16.4. Prove that every natural number is congruent to at least one of zero, one, or two, modulo three. ◇

[12] This can be proven from the axioms of set theory – for example, see [4, Theorem 4B].

16.5. **Recursion.** We now give one version of the *recursion theorem*, sometimes called *weak* or *linear* recursion.

Theorem 16.6. *Suppose that $F\colon \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is a function. Suppose that m is a natural number. Then there is a unique function $f\colon \mathbb{N} \to \mathbb{N}$ so that*

- $f(0) = m$ *and*
- *for all k we have $f(k+1) = F(f(k), k)$.* □

We omit the proof.

Example 16.7. With notation as in Theorem 16.6, we take $F(\ell, k) = \ell + 2k + 1$ and $m = 0$. The conclusion of the theorem gives a function $f\colon \mathbb{N} \to \mathbb{N}$ where

- $f(0) = 0$ and
- $f(n+1) = f(n) + 2n + 1$. ◇

We compute the first few values of f.

$$
\begin{aligned}
f(0) &= 0 \\
f(1) &= f(0) + 2 \times 0 + 1 = 1 \\
f(2) &= f(1) + 2 \times 1 + 1 = 1 + 3 = 4 \\
f(3) &= f(2) + 2 \times 2 + 1 = 1 + 3 + 5 = 9 \\
f(4) &= f(3) + 2 \times 3 + 1 = 1 + 3 + 5 + 7 = 16 \\
f(5) &= f(4) + 2 \times 4 + 1 = 1 + 3 + 5 + 7 + 9 = 25
\end{aligned}
$$

This leads us to the following.

Lemma 16.8. *The sum of the first n odd numbers equals n^2.*

Proof. The sum of the first $n+1$ odd numbers is given by adding the first $n-1$ odd numbers, and then adding $2n+1$ to the result. Thus the sum is exactly $f(n)$, as given by Example 16.7. Since $f(n)$ is defined by recursion, we give a proof by induction.

For the base case we have $f(0) = 0 = 0^2$, as desired.

For the induction step, we suppose that $f(k) = k^2$. We now compute as follows.

$$
\begin{aligned}
f(k+1) &= f(k) + 2k + 1 && \text{definition of } f \\
&= k^2 + 2k + 1 && \text{induction hypothesis} \\
&= (k+1)^2 && \text{algebra}
\end{aligned}
$$

Since the base case and the induction step hold we are done by induction. □

Exercise 16.9. Define $g(n)$ to be the sum of the first n natural numbers. Give a definition using recursion, compute the first several values, guess a polynomial equalling g, and prove your conjecture using induction. ◇

Notation 16.10. Suppose that $(a_k)_{k\in\mathbb{N}}$ is a sequence of numbers. We define the n^{th} *partial sum* recursively, as follows.

$$S(0) = 0$$
$$S(n+1) = S(n) + a_n$$

It is very common to represent $S(n)$ using the *summation notation*:

$$S(n) = \sum_{k=0}^{n-1} a_n$$ ◇

16.11. **The well-ordering principle.** Suppose that S is a subset of $\mathbb{N}$. An element $s \in S$ is a *least* element if, for all $t \in S$, we have $s \leq t$.

We are now in a position to pay off one of our debts. We prove the *well-ordering principle* by induction.

Theorem 16.12. *Suppose that S is a subset of $\mathbb{N}$. If S is non-empty, then S has a least element.*

Proof. We prove the contrapositive: that is, we suppose that S has no least element and we must prove that S is empty.

We now proceed by induction. Let $P(n)$ be the sentence "$S \cap [\![n]\!] = \varnothing$".

We first deal with the base case. Since $[\![0]\!] = \varnothing$ we have that $P(0)$ holds.

We now turn to the induction step. We prove this by contradiction. Suppose that $S \cap [\![n]\!]$ is empty *and* $S \cap [\![n+1]\!]$ is non-empty. Thus n lies in S, and is a least element. This contradicts our assumption that S has no least element. This completes the induction step.

We deduce that $P(n)$ holds for all n. Thus S is empty, as desired.

Suppose that s and s' are both least elements of S. Then $s \leq s'$ and $s' \leq s$. Thus $s = s'$, as desired. □

It is almost the case that the well-ordering principle implies the principle of induction. For a detailed and elementary discussion, see [11].

We are now equipped to pay off another of our debts [5, Book 7, Proposition 31].

Proposition 16.13. *Suppose that n is a natural number greater than one. Then n is divisible by some prime.*

Proof. Suppose that n is a natural number, greater than one. Let S be the set of divisors of n which are greater than one. That is,

$$S = \{k \in \mathbb{N} \mid k > 1 \text{ and } k \text{ divides } n\}$$

Note that n lies in S, so S is not empty.

Applying the well-ordering principle, let s be a least element of S. If s is prime we are done. So, for a contradiction, suppose that s is not prime. So s is divisible by some natural number t strictly between one and s. Since divisibility is transitive, we deduce that t lies in S. So s was not a least element of S, a contradiction. □

17. Strong induction, recursion, finite sets, and the pigeonhole principle

17.1. **Strong induction.** In the usual form of induction we only allowed ourselves to "look back one step". With the strong form we may look back as many steps as we like.

Axiom 17.2. *Suppose that $(P(n))_{n\in\mathbb{N}}$ is a sequence of sentences, one for each natural number. Suppose that*

- *for all $k \in \mathbb{N}$, if $P(\ell)$ holds for all $\ell < k$ then $P(k)$ holds.*

Then $P(n)$ holds for all n. □

Despite the name, strong induction is slightly *weaker* than weak induction. We again refer to [11] for a closely related discussion.

Stong induction gives us another "proof pattern". Here is an important example.

Theorem 17.3. *Suppose that n is a natural number, greater than zero. Then n has a unique factorisation into primes.*

Proof. Suppose that n is a natural number, greater than zero. We now assume that all natural numbers m, with $0 < m < n$, have a unique factorisation into primes.

We first prove that n has *at least one* prime factorisation. There are two cases: either n is prime or not.

Suppose that n is prime. Then it has the trivial factorisation (as itself) and we are done.

Suppose that n is not prime. Here there are two subcases: either $n = 1$ or $n > 1$. If $n = 1$ then n has the empty factorisation. Suppose instead that $n > 1$. Since n is not prime it factors as $n = a \cdot b$ where a and b are natural numbers strictly between one and n. As a and b are less than n, by the strong induction hypothesis both a and b have unique prime factorisations. Multiplying these together we obtain a prime factorisation for n.

We now prove that n has *exactly one* prime factorisation. Suppose that

$$p_1 \cdot p_2 \cdot \cdots \cdot p_k \qquad \text{and} \qquad q_1 \cdot q_2 \cdot \cdots \cdot q_\ell$$

are prime factorisations of n. We may assume that $p_i \le p_{i+1}$ and $q_i \le q_{i+1}$ for all i. If one of the factorisations is empty then both are empty, n equals one, and we are done.

So we may assume that $k, \ell \ge 1$. Suppose that $p_1 = q_1$. In this case the strong induction hypothesis applied to n/p_1 implies that $p_i = q_i$ for all $i > 1$; so we are done.

Suppose instead, for a contradiction, that $p_1 < q_1$. Then we have

$$n\left(\frac{p_1}{p_1} - \frac{p_1}{q_1}\right) = p_1(p_2 \cdot \cdots \cdot p_k - q_2 \cdot \cdots \cdot q_\ell) = (q_1 - p_1)q_2 \cdot \cdots \cdot q_\ell$$

All of these are natural numbers, and all are divisible by p_1. Since $0 < q_1 - p_1 < q_1 \le n$ the strong induction hypothesis gives us a unique prime factorisation

$$r_1 \cdot r_2 \cdot \cdots \cdot r_m$$

of $q_1 - p_1$. Thus

$$(q_1 - p_1)q_2 \cdot \cdots \cdot q_\ell = r_1 \cdot r_2 \cdot \cdots \cdot r_m \cdot q_2 \cdot \cdots \cdot q_\ell$$

and this factorisation is unique. As noted above, p_1 divides this. Thus p_1 equals one of the q_i or one of the r_i. However, since $p_1 < q_1$ we deduce that p_1 is not equal to any of the q_i. Thus p_1 equals one of the r_i. Thus p_1 divides $q_1 - p_1$. So p_1 divides q_1. However, since q_1 is prime (and not equal to p_1), this is a contradiction.

The case where $q_1 < p_1$ is argued similarly. This completes the strong induction and thus the proof. □

17.4. **Strong recursion.** There are many other versions of recursion; for example see [8, page 48]. Instead of giving a precise definition of "strong" recursion we give a very common example.

Definition 17.5. We define $f\colon \mathbb{N} \to \mathbb{N}$ by

- $f(0) = 0$,
- $f(1) = 1$, and
- $f(k+2) = f(k+1) + f(k)$ for all $k \in \mathbb{N}$.

We call $f(n)$ the n^{th} *Fibonacci number.* ◇

Note that the definition of $f(n+2)$ requires knowing the previous two values. Here are the first ten values of f.

n	0	1	2	3	4	5	6	7	8	9
$f(n)$	0	1	1	2	3	5	8	13	21	34

Proving facts about the Fibonacci numbers will typically require strong induction.

Exercise 17.6. Suppose that $f(n)$ is the n^{th} Fibonacci number, as defined in Definition 17.5. Prove that $f(n)$ is even if and only if n is divisible by three. ◇

Exercise 17.7.

(1) Determine the number of binary strings of length n where no adjacent bits are equal.
(2) [Hard.] Determine the number of binary strings of length n where no adjacent bits are both 0. ◇

17.8. **Finite sets and the pigeonhole principle.** A *dovecote* is a structure intended to house pigeons or doves. Part of its interior is typically divided into small *pigeonholes* to allow the pigeons to nest separately or in pairs. The *pigeonhole principle* states that

- if there are more pigeons than pigeonholes and
- if every pigeon is in some pigeonhole

then some pigeonhole contains at least two pigeons. The origins of the principle are, of course, unclear [2].

We give a slightly non-standard proof, starting with the following.

Theorem 17.9. *Suppose that n is a natural number. Suppose that $f\colon [\![n]\!] \to [\![n+1]\!]$ is a function. Then f is not a surjection.*

This statement is modelled on our version of Cantor's theorem (Theorem 3.12).

Proof of Theorem 17.9. We proceed by induction.

We begin with the base case. Here we must prove that there is no surjection from $[\![0]\!]$ to the singleton set $[\![1]\!]$. For a contradiction, suppose that there is some surjection $f\colon [\![0]\!] \to [\![1]\!]$. Thus there is some element a of $[\![0]\!]$ so that $f(a) = 0 \in [\![1]\!]$. However, by Notation 1.13 the set $[\![0]\!]$ is empty. Thus a does not exist, we have obtained a contradiction, and the base case is complete.

We now prove the contrapositive of the induction step. Suppose that $f\colon [\![n+1]\!] \to [\![n+2]\!]$ is a surjection. We now define a function $g\colon [\![n]\!] \to [\![n+1]\!]$.

$$g(k) = \begin{cases} f(k), & \text{if } f(k) < n+1 \\ f(n), & \text{if } f(k) = n+1 \text{ and } f(n) < n+1 \\ 0, & \text{if } f(k) = f(n) = n+1 \end{cases}$$

We now show that g is a surjection.

Suppose that ℓ is any element of $[\![n+1]\!]$; so $\ell < n+1$ and thus is an element of $[\![n+2]\!]$. Since f surjects $[\![n+2]\!]$ there is some $k \in [\![n+1]\!]$ so that $f(k) = \ell$. There are now two cases: either $k < n$ or $k = n$.

(1) Suppose that $k < n$. Since $f(k) = \ell < n+1$, the first line of the definition of g applies. Thus $g(k) = f(k) = \ell$ and we are done.
(2) Suppose that $k = n$. This does not lie in the domain of g. Instead, we again use the fact that f is surjective. So there is some $k' \in [\![n+1]\!]$ so that $f(k') = n+1$. Note that $f(n) = \ell < n+1 = f(k')$. Since f is a function, we deduce that k' is not equal to n. Since $f(k') = n+1$ and $f(n) < n+1$, the second line of the definition of g applies. Thus $g(k') = f(n) = \ell$ and we are done.

Thus g is a surjection, as claimed. This completes the induction step, and thus the proof of the theorem. □

Remark 17.10. In the proof above, the third line of the definition of g is only present to ensure that g is a function. It is not needed to show that g is a surjection. ◇

Exercise 17.11. Give proofs of the following, using Theorem 17.9 as needed.

(1) Suppose that n is a natural number. Prove that there is no injection $g\colon [\![n+1]\!] \to [\![n]\!]$.
(2) Suppose that m and n are natural numbers, with $m > n$. Prove that there is no injection $g\colon [\![m]\!] \to [\![n]\!]$.
(3) Suppose that n is a natural number. Prove that there is no bijection $g\colon \mathbb{N} \to [\![n]\!]$. [That is, $\mathbb{N}$ is infinite in the sense of Definition 3.14.]
(4) Suppose that X is a finite set. Then it has a unique cardinality (as given in Definition 3.9).
(5) Suppose that X is a finite set. Then a function $f\colon X \to X$ is an injection if and only if it is surjection. ◇

We now pay off our final (?) debt.

Lemma 17.12. *Suppose that X and Y are sets, with X finite.*

- *Suppose that $f\colon Y \to X$ is an injection. Then Y is finite.*
- *Suppose that $g\colon X \to Y$ is a surjection. Then Y is finite.*

Proof. We first suppose that $f\colon Y \to X$ is an injection. Since X is finite, it is in bijection with $[\![n]\!]$ for some $n \in \mathbb{N}$. So we can simplify the remainder of the proof by assuming $X = [\![n]\!]$. We can also assume that Y is a subset of $\mathbb{N}$.

We now build a function h by recursion. That is, we define a sequence of functions and take h equal to the last one.

In the base case we take $h_0 \colon [\![0]\!] \to Y$ to be the unique function with domain $[\![0]\!]$ and codomain Y.

For the recursive step, we assume that $h_k \colon [\![k]\!] \to Y$ has already been defined. Let $Y_k = h_k([\![k]\!])$ be the image of $[\![k]\!]$ under h_k. There are two cases.

- If $Y_k = Y$ then $h = h_k$ and the construction is done.
- Otherwise, let $y_k \in Y - Y_k$ be the least element, obtained from the well-ordering principle. We now define $h_{k+1} \colon [\![k+1]\!] \to Y$ by

$$h_{k+1}(\ell) = \begin{cases} h_k(\ell), & \text{if } \ell < k \\ y_k, & \text{if } \ell = k \end{cases}$$

This defines h_{k+1} and so completes the recursive step.

By induction each of the h_k is an injection. Thus by the pigeonhole principle (Exercise 17.11) the construction halts after at most n steps, giving h. Since the construction halts, h is surjective. Thus h is a bijection, and Y is finite.

We next suppose that $g \colon X \to Y$ is an surjection. We again replace X by $[\![n]\!]$ for some $n \in \mathbb{N}$. We now define a function $h \colon Y \to X$ by taking $h(y) = \min g^{-1}(y)$. That is, $h(y)$ is the smallest element in the preimage of y under g. Note that g is a left inverse for h, so h is injective. We now apply the previous statement to deduce that Y is finite. □

18. There is no Section 18

19. Groups: definitions and examples

In the final four weeks we give an introduction to elementary algebra, including groups, rings, and fields.

19.1. **Definition and notation.** Groups are avatars of symmetry; as such they arise throughout mathematics.

Definition 19.2. A *group* is a set G together with a function $G \times G \to G$ (here denoted as multiplication) satisfying the following properties.

(1) There is an element $e_G \in G$ so that for all $g \in G$ we have
$$e_G \cdot g = g \cdot e_G = g$$
(2) For every element $g \in G$ there is an element $h \in G$ so that
$$g \cdot h = h \cdot g = e_G$$
(3) For all $f, g, h \in G$ we have
$$(f \cdot g) \cdot h = f \cdot (g \cdot h)$$ ◇

Lemma 19.3. *Suppose that $(G, \cdot)$ is a group.*

(1) Suppose that $g \in G$ has the following property: for every h in G we have $g \cdot h = h \cdot g = h$. Then $g = e_G$.

(2) Suppose that $g \in G$ is an element. Suppose that h and h' are also elements of G, so that
$$g \cdot h = g \cdot h = g \cdot h' = g \cdot h' = e_G$$
Then $h = h'$.

Proof. For the first property we compute as follows.

$g = e_G \cdot g$	because e_G is an identity
$= e_G$	because g is an identity

So $g = e_G$, as desired.

For the second property we compute as follows.

$h = h \cdot e_G$	identity
$= h \cdot (g \cdot h')$	h' is an inverse for g
$= (h \cdot g) \cdot h'$	associativity
$= e_G \cdot h'$	h is an inverse for g
$= h'$	identity

So $h = h'$, as desired. □

When discussing a pair of groups, say $(G, \cdot_G)$ and $(H, \cdot_H)$, we may distinguish their group operations by adding subscripts.

Notation 19.4. There are two common notations for groups: additive and multiplicative.

(1) In additive notation, we use $+$ or $+_G$ for the group operation, we use 0 or 0_G for the identity element, and we use $-g$ for the *negative* of $g \in G$.
(2) In multiplicative notation, we use $\cdot$ or $\cdot_G$ for the group operation, we use 1 or 1_G for the identity element, and we use g^{-1} for the *inverse* of $g \in G$.

In both notations, the third property says that the group multiplication is *associative*. ◇

Note that, by Lemma 19.3, there is a unique identity element in G. Also, inverses in G exist and are unique. Thus there is a bijection $i\colon G \to G$ taking elements to their inverses.

For example, $(\mathbb{Z}, +)$ is a group: zero is the identity element and negation gives additive inverses.

Exercise 19.5. Suppose that $(G, \cdot)$ is a group (with multiplicative notation).

(1) Prove that $e_G^{-1} = e_G$.
(2) Suppose that $g \in G$ is an element. Prove that $(g^{-1})^{-1} = g$. ◇

19.6. **Simple examples.**

Definition 19.7. Suppose that $(G, \cdot)$ is a group. Then we call the cardinality $|G|$ the *order* of the group $(G, \cdot)$. ◇

A *trivial group* $(G, \cdot)$ has order one. Thus its only element is the identity e_G. The group multiplication has the single equality $e_G \cdot e_G = e_G$.

The subset $\{\pm 1\}$ of the integers, equipped with the usual multiplication, is a group of order two. Here 1 is the identity element and -1 is its own inverse. Also, the group operation is associative by Remark 9.7.

The group $(\mathbb{Z}, +)$ gives an additive example. Here 0 is the identity element and $-n$ is the negative of n. Also, the group operation is associative by Remark 9.7.

On the other hand, $(\mathbb{N}, +)$ has a zero as its identity element and satisfies associativity. But $(\mathbb{N}, +)$ is not a group because the only element with an additive inverse is zero itself.

In a similar, but slightly more subtle, fashion $(\mathbb{Q}, \times)$ is not a group. Here 1 serves as the identity element and multiplication is associative. However there is exactly one element, namely zero, which does not have a multiplicative inverse.

Example 19.8. Suppose that n is a positive natural number. We define $C_n = (\mathbb{Z}/n\mathbb{Z}, +_n)$ to be the *cyclic group of order* n. Corollary 10.15 implies that

- $[0]_n$ is the identity element,
- $[-k]_n$ is the negative of $[k]_n$, and
- $+_n$ is associative.

By Exercise 10.10 the group C_n has order n. ◇

The term "cyclic" is explained in Section 20.12.

19.9. **Dihedral groups.** Before giving the next example, we need a few definitions.

Definition 19.10. Suppose that $\theta \in \mathbb{R}$ is a real number. The function $R_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$R_\theta(x, y) = (\cos(\theta) \cdot x - \sin(\theta) \cdot y,\ \sin(\theta) \cdot x + \cos(\theta) \cdot y)$$

We call R_θ a *rotation* through the angle θ. ◇

One way to remember this is to connect R_θ to the rotation matrix, as follows.

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\theta) \cdot x - \sin(\theta) \cdot y \\ \sin(\theta) \cdot x + \cos(\theta) \cdot y \end{pmatrix}$$

Lemma 19.11. *For all real numbers η and θ we have the following.*

(1) $R_{2\pi} = R_0 = \mathrm{Id}_{\mathbb{R}^2}$.
(2) $R_\eta \circ R_\theta = R_{\eta+\theta}$. □

We omit the proofs; they follow from the usual trigonometric identities.

Definition 19.12. Suppose that $\theta \in \mathbb{R}$ is a real number. The function M_θ is given by

$$M_\theta(x, y) = (\cos(\theta) \cdot x + \sin(\theta) \cdot y,\ \sin(\theta) \cdot x - \cos(\theta) \cdot y)$$

We call M_u a *reflection* about the line with angle $\theta/2$. ◇

One way to remember this is to connect M_θ to the reflection matrix, as follows.

$$\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\theta) \cdot x + \sin(\theta) \cdot y \\ \sin(\theta) \cdot x - \cos(\theta) \cdot y \end{pmatrix}$$

Lemma 19.13. *For all real numbers η and θ we have the following.*

(1) $M_\eta \circ M_\theta = R_{\eta-\theta}$.
(2) $R_\eta \circ M_\theta = M_{\eta+\theta}$.
(3) $M_\eta \circ R_\theta = M_{\eta-\theta}$. □

We omit the proofs; they follow from the usual trigonometric identities.

Definition 19.14. Suppose that n is a positive natural number. Let $\Theta(n)$ be the set

$$\Theta(n) = \left\{ \frac{2\pi k}{n} \;\middle|\; k \in [\![n]\!] \right\}$$

That is, $\Theta(n)$ is the set of multiples of $2\pi/n$ less than 2π. Let D_{2n} be the set

$$D_{2n} = \{R_\theta, M_\theta \mid \theta \in \Theta(n)\}$$

We call $(D_{2n}, \circ)$ the *dihedral group* of order $2n$. ◇

Exercise 19.15. Prove that $(D_{2n}, \circ)$ is in fact a group. ◇

Exercise 19.16. Suppose that $n = 4$. The dihedral group D_{2n} of order eight is the group of symmetries of the square $[-1, 1]^2$ in the plane $\mathbb{R}^2$. Draw the square in the plane with decorations breaking all symmetry; then draw the image of the square under all elements of D_{2n}. ◇

20. Multiplication tables, commutativity, subgroups, intersections, generators, and powers

We now write simply G to represent a group, omitting the mention of the group multiplication.

20.1. **Multiplication tables.** Suppose that G is a group. Suppose that g is an element of G. Then we define two functions on G, as follows.

$$\begin{aligned} L_g &: G \to G & L_g(h) &= g \cdot h \\ R_g &: G \to G & R_g(h) &= h \cdot g \end{aligned}$$

These are *left multiplication* and *right multiplication* by g.

Lemma 20.2. *For every $g \in G$ the functions L_g and R_g are bijections.*

Proof. The function $L_{g^{-1}}$ is a two-sided inverse for L_g. So, by Lemma 6.5, the function L_g is a bijection. The proof for R_g is similar. □

The *multiplication table* for a group G has rows and columns indexed by the elements of G and has (g, h) entry being $g \cdot h$. Here is the multiplication table for the group $(\mathbb{Z}/4\mathbb{Z}, +)$; we write k instead of $[k]_4$ to simplify the notation.

	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

Lemma 20.2 has a nice consequence, which we might call the "sudoko theorem".

Corollary 20.3. *In the multiplication table for $(G, \cdot)$, every row and every column is a permutation of the elements of G.*

Proof. The row for g is the image of the function L_g while the column for h is the image of the function R_h. □

20.4. **Commutative groups.**

Definition 20.5. Suppose that G is a group. Suppose that $g, h \in G$ satisfy $g \cdot h = h \cdot g$. Then we say that g and h *commute*. If all pairs have this property then we call the group $(G, \cdot)$ *commutative*.[13] If there is a pair of elements that do not commute then we call the group *non-commutative*. ◇

As examples, $(\mathbb{Z}, +)$ is commutative, as are the cyclic groups C_n. However, many groups are not commutative.

Exercise 20.6. Suppose that $n \geq 3$. Find a pair of elements g and h in D_{2n} which do not commute. ◇

20.7. **Subgroups.**

Definition 20.8. Suppose that G is a group. Suppose that $H \subset G$ is a subset. Suppose also that

- $e_G \in H$,
- for all $g \in H$ we have g^{-1} in H, and
- for all $g, h \in H$ we have $g \cdot h$ in H.

Then we call H a *subgroup* of G and we write $H < G$. ◇

For example, $\{e_G\}$ is a subgroup of G; this is called the *trivial* subgroup. Also G is a subgroup of itself. If $H < G$ is a subgroup, not equal to $\{e_G\}$, then we call H a *non-trivial* subgroup. If $H < G$ is a subgroup, not equal to G, then we call H a *proper* subgroup.

Exercise 20.9. Find all subgroups of $(\mathbb{Z}/6\mathbb{Z}, +_6)$. ◇

[13] These are also called *abelian* groups. As Niels Henrik Abel's work, and indeed life, predate group theory we prefer the more informative name "commutative group".

Exercise 20.10. Suppose that n is a natural number greater than one. Prove that the subset of rotations is a non-trivial, proper subgroup of D_{2n} of order n. ◇

As a consequence, non-commutative groups may have commutative subgroups. The converse does not hold, as follows.

Lemma 20.11. *Suppose that G is a commutative group. Suppose that $H < G$ is a subgroup. Then H is commutative.* □

20.12. **Intersections of subgroups, generators, and powers.**

Lemma 20.13. *Suppose that G is a group. Suppose that $\mathcal{H}$ is a collection of subgroups of G. Then the intersection $\bigcap_{H \in \mathcal{H}} H$ is a subgroup of G.*

Proof. The identity e_G is an element of every subgroup. Thus e_G lies in H for all $H \in \mathcal{H}$. Thus the intersection $\bigcap_{H \in \mathcal{H}} H$ contains e_G.

Suppose that g and h lie in $\bigcap_{H \in \mathcal{H}} H$. Then they lie in H, for all $H \in \mathcal{H}$. Thus g^{-1} and $g \cdot h$ lie in H, for all $H \in \mathcal{H}$. Thus g^{-1} and $g \cdot h$ lie in $\bigcap_{H \in \mathcal{H}} H$, as desired. □

Definition 20.14. Suppose that G is a group. Suppose that $S \subset G$ is a subset of G. We define $\langle S \rangle$ to be the intersection of all subgroups of G that contain S. That is, let

$$\mathcal{H}_S = \{H < G \mid S \subset H\}$$

and define $\langle S \rangle = \bigcap_{H \in \mathcal{H}_S} H$. We call $\langle S \rangle$ the subgroup of G *generated* by S. ◇

If $H = \langle S \rangle$ then we call the elements of S *generators* for H. This is because every element of H may be obtained by multiplying together elements of S.

When $S = \{g\}$ is a singleton we may simply write $\langle g \rangle$ instead of writing $\langle S \rangle$. In this case g is a *generator* of $\langle g \rangle$.

Example 20.15. The singleton set $\{1\}$ generates the group $(\mathbb{Z}, +)$. On the other hand, the additive group $(\mathbb{Q}, +)$ does not have a finite generating set. ◇

We call any subgroup (with a single generator) a *cyclic subgroup*. Finally, the *order* of an element $g \in G$ is the cardinality of $\langle g \rangle$. For example, the order of e_G is one. The order of the rotation $R_{2\pi/n}$ is n and the order of the reflection M_θ is two, for any $\theta \in \mathbb{R}$.

Definition 20.16. Suppose that G is a group. Suppose that $g \in G$ is an element. For n in $\mathbb{N}$ we define g^n by recursion, as follows.

$$g^0 = e_G$$
$$g^{n+1} = g^n \cdot g$$

Finally, for $n < 0$ in $\mathbb{Z}$ we define $g^n = (g^{-1})^{-n}$. We call the elements g^n the *powers* of g. ◇

Induction proves the following "laws of exponents".

Lemma 20.17. *Suppose that G is a group. Suppose that g is an element of G. For any integers m and n we have $g^{m+n} = g^m \cdot g^n$ and $g^{mn} = (g^m)^n$.* □

Exercise 20.18. Suppose that G is a group. Suppose that $g \in G$ is an element. Prove that the subgroup $\langle g \rangle$ contains exactly the powers of g. Deduce that $\langle g \rangle$ is a commutative subgroup of G. ◇

Exercise 20.19. [Hard.] Suppose that G is a group. Suppose that $g \in G$ is an element with finite order $k > 1$. Prove that

$$\langle g \rangle = \{g^i \mid i \in [\![k]\!]\}$$

Deduce that k is the least natural number so that $g^k = e_G$. ◇

Exercise 20.20. Suppose that n is a positive natural number. Show that the following sets generate the following groups.

(1) $S = \{R_{2\pi/n}\}$ generates the rotation subgroup of D_{2n}.
(2) $T = \{R_{2\pi/n}, M_{2\pi/n}\}$ generates the dihedral group D_{2n}. ◇

21. The symmetric group

21.1. Definition of Sym(n).

Definition 21.2. Suppose that X is a set. Let $\mathrm{Sym}(X)$ be the set of permutations of X: that is, bijections from X to itself. We define the product $\circ$ on $\mathrm{Sym}(X)$ to be the composition of functions. We call $(\mathrm{Sym}(X), \circ)$ the *symmetric group* on X. ◇

Exercise 21.3. Suppose that X is a set. Prove that $\mathrm{Sym}(X)$ is a group. ◇

21.4. **Two- and one-line notation.** In a slight abuse of notation we write $\textsc{Sym}(n)$ for the group $\textsc{Sym}(\llbracket n \rrbracket)$. There are several notations for permutations in $\textsc{Sym}(n)$. We will give three; these are the *two-line*, *one-line*, and *cycle* notations. Two-line notation is extremely clear but not very concise. Cycle notation is more expressive.

Definition 21.5. Suppose that σ is a permutation of $\llbracket n \rrbracket$. The *two-line* notation for σ is the array of numbers with two rows and n columns, as follows. The first row is the elements of $\llbracket n \rrbracket$, in order. The second row is the elements of the list $(\sigma(i))_{i=0}^{n-1}$, in order. In particular the number in the second line $\sigma(i)$ appears immediately beneath the number i in the first line. ◇

Here are two examples of two-line notation from $\textsc{Sym}(5)$.

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 0 & 3 & 2 & 4 \end{pmatrix}$$

These represent, respectively, the identity and the permutation σ with $\sigma(0) = 1$, $\sigma(1) = 0$, $\sigma(2) = 3$, $\sigma(3) = 2$, and $\sigma(4) = 4$. Note that, following Definition 6.11, the identity has five fixed points, while σ has only one.

Exercise 21.6. Show that the permutation σ (given immediately above) has order two. ◇

For any permutation in $\textsc{Sym}(n)$, the first line of its two-line notation is always the same; thus we may omit the first line to obtain the following.

Definition 21.7. Suppose that σ is a permutation of $\llbracket n \rrbracket$. The *one-line* notation for σ is the second row of the two-line notation for σ. ◇

The one-line notations for the examples above are as follows.

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 3 & 2 & 4 \end{pmatrix}$$

21.8. **Composing permutations.** Recall that $\textsc{Sym}(n)$ is a group written multiplicatively. Thus, for permutations σ and τ, we will write simply $\sigma\tau$ for $\sigma \circ \tau$. In particular, we stress that $\sigma\tau$ is again a permutation, and $(\sigma\tau)(i) = \sigma(\tau(i))$. That is, we first act by τ and then by σ.

Exercise 21.9. Let σ and τ be the permutations in $\textsc{Sym}(6)$ with one-line notation as follows.

$$\sigma = \begin{pmatrix} 1 & 2 & 0 & 4 & 5 & 3 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 0 & 3 & 2 & 5 & 4 \end{pmatrix}$$

Write out the one-line notations for the compositions $\sigma\tau$ and $\tau\sigma$. ◇

We will write, as usual, σ^k for the k^{th} power of σ (that is, composing σ with itself k times).

Exercise 21.10. Let $\sigma \in \text{SYM}(6)$ be the permutation with one-line notation as follows.

$$\begin{pmatrix} 1 & 2 & 0 & 4 & 5 & 3 \end{pmatrix}$$

Write out the one-line notations for all powers of σ. ◇

22. Computing with permutations

22.1. **Cycles.** Before giving cycle notation we discuss cycles in general.

Definition 22.2. Suppose that $\sigma \in \text{SYM}(n)$ is a permutation. The *orbit decomposition* $\mathcal{O}_\sigma$ is the set of orbits of σ. That is:

$$\mathcal{O}_\sigma = \{\mathcal{O}_\sigma(k) \mid k \in [\![n]\!]\}$$ ◇

Lemma 22.3. *Suppose that $\sigma \in \text{SYM}(n)$ is a permutation. Then $\mathcal{O}_\sigma$ is a partition of $[\![n]\!]$.*

Proof. Every $k \in [\![n]\!]$ lies in $\mathcal{O}_\sigma(k)$. Thus the union of the orbits is all of $[\![n]\!]$.

Suppose that a, b, and c lie in $[\![n]\!]$. Suppose that c lies in the intersection $\mathcal{O}_\sigma(a) \cap \mathcal{O}_\sigma(b)$. So there are powers p and q with $\sigma^p(a) = c$ and $\sigma^q(b) = c$. Thus $\sigma^{p-q}(a) = b$. Thus b lies in $\mathcal{O}_\sigma(a)$. Induction now implies that $\mathcal{O}_\sigma(b)$ is contained in $\mathcal{O}_\sigma(a)$. A similar argument gives the opposite inclusion, and Lemma 1.16 implies the equality. □

Definition 22.4. Suppose that $\sigma \in \text{SYM}(n)$ is a permutation. Suppose that k is a natural number, larger than one. We say that σ is a *k–cycle* if exactly one part $R \in \mathcal{O}_\sigma$ has cardinality k and all other parts are singletons.

We call the part R the *support* of σ. ◇

Two-cycles are also called *transpositions*.

Lemma 22.5. *Suppose that σ is a k–cycle with support $R \subset [\![n]\!]$. Suppose that a is any element of R. Then the list $(\sigma^i(a))_{i=0}^{k-1}$ contains every element of R, exactly once.*

Furthermore, σ has order k. □

The proof is similar to the solution of Exercise 20.19 and we omit it.

Notation 22.6. Suppose that $\sigma \in \text{SYM}(n)$ is a k–cycle. Suppose that R is the support of σ. Let a be the smallest element of R. Then the list

$$\begin{pmatrix} a & \sigma(a) & \sigma^2(a) & \dots & \sigma^{k-1}(a) \end{pmatrix}$$

(written without commas) is the *cycle* notation for σ. ◇

Exercise 22.7. Prove that cycles in $\textsc{Sym}(n)$, with disjoint supports, commute. [That is: suppose that σ and τ are cycles in $\textsc{Sym}(n)$ with supports R_σ and R_τ. Suppose that $R_\sigma \cap R_\tau = \varnothing$. Prove that $\sigma \circ \tau = \tau \circ \sigma$.] ◇

Exercise 22.8. Write the four-cycle $\sigma = (0\ 1\ 2\ 3)$ as a composition of transpositions. ◇

Exercise 22.9. Suppose that $\sigma \in \textsc{Sym}(n)$ is a k–cycle. Prove that σ is equal to a composition of $k-1$ transpositions. ◇

Notation 22.10. Suppose that σ is any permutation in $\textsc{Sym}(n)$. We order the parts in $\mathcal{O}_\sigma$ by their smallest element. For each we write the corresponding cycle (even those of length one), in that order. ◇

As an example, consider the permutation σ with two-line notation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 1 & 2 & 0 & 4 & 3 & 6 & 5 & 7 & 8 \end{pmatrix}$$

This has one 3–cycle $(0\ 1\ 2)$, two 1–cycles $(3\ 4)$ and $(5\ 6)$, and two fixed points (7) and (8). Thus the cycle notation for σ is

$$(0\ 1\ 2)(3\ 4)(5\ 6)(7)(8)$$

It is very common to omit the fixed points (cycles of length one). This gives

$$(0\ 1\ 2)(3\ 4)(5\ 6)$$

for the above example.

22.11. **An algorithm to find cycle notation.**

Algorithm 22.12. We are given the following:

- a natural number n and
- a permutation $\sigma \in \textsc{Sym}(n)$.

We take C to be the empty string and W to be the empty set. While the set $[\![n]\!] - W$ is non-empty, we do the following.

(1) If the string C is empty, or ends with a right parenthesis, then we append a left parenthesis to C.
(2) If the last thing written was a left parenthesis, then we let a be the smallest element of $[\![n]\!] - W$. We append a to the string C and add a copy of a to the set W.
(3) If the last thing written was a number, say b, then we set $c = \sigma(b)$ and we do exactly one of the following.
   (a) If c lies in the set W then we append a right parenthesis to the string C.

(b) If c is not in the set W then we append c to the string C and add a copy of c to W.

Once the algorithm is complete, we may optionally erase all cycles of length one. ◇

Induction and Lemmas 22.3 and 22.5 imply the following.

Lemma 22.13. *When Algorithm 22.12 halts, the string C is the cycle notation for σ.* □

23. Homomorphisms and isomorphisms

23.1. **Definitions.**

Definition 23.2. Suppose that G and H are groups. Suppose that $\phi: G \to H$ is a function. We say that ϕ is a *homomorphism* if, for all $g, g' \in G$, we have

$$\phi(g \cdot g') = \phi(g) \cdot \phi(g')$$ ◇

Lemma 23.3. *Suppose that $\phi: G \to H$ is a homomorphism. Then, for any $g \in G$ and for any $k \in \mathbb{Z}$ we have $\phi(g^k) = (\phi(g))^k$.*

It follows that homomorphisms send cyclic subgroups to cyclic subgroups.

Proof of Lemma 23.3. When $k = 0$ the claim reduces to $\phi(e_G) = e_H$. To prove this, we compute as follows.

$$\begin{aligned}
\phi(e_G) &= e_H \cdot \phi(e_G) && \text{identity in } H \\
&= ((\phi(e_G))^{-1}\phi(e_G))\phi(e_G) && \text{inverse in } H \\
&= (\phi(e_G))^{-1}(\phi(e_G)\phi(e_G)) && \text{associativity in } H \\
&= (\phi(e_G))^{-1}\phi(e_G \cdot e_G) && \phi \text{ is a homomorphism} \\
&= (\phi(e_G))^{-1}\phi(e_G) && \text{identity in } G \\
&= e_H && \text{inverse in } H
\end{aligned}$$

When $k = -1$ the claim reduces to $\phi(g^{-1}) = (\phi(g))^{-1}$. To prove this, we compute as follows.

$$\begin{aligned}
\phi(g)\phi(g^{-1}) &= \phi(g \cdot g^{-1}) && \phi \text{ is a homomorphism} \\
&= \phi(e_G) && \text{inverse in } G \\
&= e_H && \text{first property (above)}
\end{aligned}$$

The uniqueness of inverses (Lemma 19.3) gives the desired equality.

Induction on k establishes the claim in general. □

Lemma 23.4. *Suppose that $\phi\colon G \to H$ is a homomorphism. Suppose that $\psi\colon H \to K$ is a homomorphism. Then $\psi \circ \phi\colon G \to K$ is a homomorphism.*

Proof. Suppose that g and g' lie in G. Then we compute as follows.

$$\begin{aligned}
(\psi \circ \phi)(g \cdot g') &= \psi(\phi(g \cdot g')) && \text{definition of composition} \\
&= \psi(\phi(g) \cdot \phi(g')) && \phi \text{ is a homomorphism} \\
&= \psi(\phi(g)) \cdot \psi(\phi(g')) && \psi \text{ is a homomorphism} \\
&= (\psi \circ \phi)(g) \cdot (\psi \circ \phi)(g') && \text{definition of composition}
\end{aligned}$$

Thus $\psi \circ \phi$ is a homomorphism □

Definition 23.5. Suppose that $\phi\colon G \to H$ is a bijective homomorphism. Then we call ϕ an *isomorphism* and we write $G \cong H$. ◇

Exercise 23.6. Suppose that $\mathcal{G}$ is a set of groups. Prove that the relation $G \cong H$ on $\mathcal{G}$ is an equivalence relation. ◇

23.7. **Examples of homomorphisms.**

Example 23.8. Suppose that n is an integer. Define the function $\phi_n\colon \mathbb{Z} \to \mathbb{Z}$ by $\phi_n(k) = n \cdot k$. Then ϕ is an injective homomorphism from (the additive group) $\mathbb{Z}$ to itself.

To prove this, we work with additive notation for $\mathbb{Z}$. Suppose that k and ℓ are integers. We compute as follows.

$$\begin{aligned}
\phi(k + \ell) &= n \cdot (k + \ell) && \text{definition of } \phi \\
&= n \cdot k + n \cdot \ell && \text{distributivity in } \mathbb{Z} \\
&= \phi(k) + \phi(\ell) && \text{definition of } \phi\text{, twice}
\end{aligned}$$ ◇

Exercise 23.9. Suppose that n is a positive natural number. Define the function $\phi_n\colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ by $\phi_n(k) = [k]_n$. Show that ϕ is a surjective homomorphism. ◇

Exercise 23.10. Suppose that G is a group. Suppose that g is an element of G. Then the function $\mathrm{pow}_g\colon \mathbb{Z} \to G$ defined by $\mathrm{pow}_g(n) = g^n$ is a homomorphism. Furthermore, the image of pow_g is $\langle g \rangle$, the subgroup generated by g. ◇

Proposition 23.11. *The usual inclusions $\mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ are all injective homomorphisms of additive groups.* □

Exercise 23.12. Suppose that m, n, and p are positive integers. Suppose that $m = p \cdot n$. Prove the following.

- The function $\phi\colon \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ defined by $\phi([k]_n) = [p \cdot k]_m$ is a well-defined homomorphism.

- The function $\psi\colon \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ defined by $\psi([k]_m) = [k]_n$ is a well-defined homomorphism.

$\diamond$

Example 23.13. Let $\mathbb{C}^\times = \mathbb{C} - \{0\}$ be the *punctured plane*. Recall that $(\mathbb{C}, +)$ is an additive group. Also $(\mathbb{C}^\times, \cdot)$ is a multiplicative group. The exponential function $\exp\colon (\mathbb{C}, +) \to (\mathbb{C}^\times, \cdot)$ is a homomorphism. This follows from the "law of exponents". The proof requires more analysis than we have at hand. $\diamond$

23.14. **Examples of isomorphisms.** Isomorphism measures a certain kind of "sameness". Suppose that P is a property of groups: that is, P is stated in terms of the axioms of groups. Suppose that G and H are isomorphic groups. Then $P(G)$ holds if and only if $P(H)$ holds.

We now give several examples of this.

Definition 23.15. Suppose that $\mathbb{C}$ is the complex plane. Suppose that n is a positive natural number. An n^{th} *root of unity* is any complex number in $\mathbb{C}$ satisfying the equation $z^n - 1 = 0$. These are

$$\zeta_n^k = \exp(2\pi i k/n) = \cos(2\pi k/n) + i \sin(2\pi k/n)$$

where k is an integer. $\diamond$

Exercise 23.16. Suppose that n, p, and k are positive integers. Show that $\zeta_{pn}^{pk} = \zeta_n^k$. $\diamond$

We now define $U_n = \{\zeta_n^k \mid k \in \mathbb{Z}\}$ to be the multiplicative group of n^{th} roots of unity.

Exercise 23.17. Prove that U_n is a group. $\diamond$

Lemma 23.18. *Suppose that $\phi\colon \mathbb{Z}/n\mathbb{Z} \to U_n$ is defined by $\phi([k]_n) = \zeta_n^k$. Then ϕ is a well-defined isomorphism.*

Proof. Note that if k and ℓ are integers with $k \equiv \ell \pmod{n}$ then $\zeta_n^k = \zeta_n^\ell$. Thus ϕ is well-defined (see Theorem 9.4). That ϕ is an homomorphism follows from the laws of exponents (Lemma 20.17). Finally, ϕ is a bijection by Exercise 10.10. $\square$

Exercise 23.19. Rephrase Exercise 23.12 in terms of the groups U_n. $\diamond$

Let E_n be the subgroup of rotations inside the dihedral group D_{2n}. The following has a proof similar to that of Lemma 23.18.

Lemma 23.20. *Suppose that $\phi\colon \mathbb{Z}/n\mathbb{Z} \to E_n$ is defined by $\phi([k]_n) = R_{2\pi k/n}$. Then ϕ is a well-defined isomorphism.* $\square$

Thus the cyclic group $\mathbb{Z}/n\mathbb{Z}$, the rotation group E_n, and the group of roots U_n are all isomorphic.

Exercise 23.21. Prove that $\textsc{Sym}(3)$ is isomorphic to the dihedral group D_6. ◇

Exercise 23.22. Prove that if $n > 3$ then $\textsc{Sym}(n)$ is not isomorphic to a dihedral group. ◇

24. Inversions and parity of permutations

24.1. **Inversions.** Suppose that n is a natural number. We define

$$\mathcal{D}(n) = \{\{a, b\} \subset [\![n]\!] \mid a \neq b\}$$

to be the set of elements in $\mathcal{P}([\![n]\!])$ with cardinality two. For example, $\mathcal{D}(0)$ and $\mathcal{D}(1)$ are empty, $\mathcal{D}(2)$ is a singleton, $\mathcal{D}(3)$ has three elements, and $\mathcal{D}(4)$ has six elements.

Exercise 24.2. Prove that $\mathcal{D}(n)$ has cardinality $n(n-1)/2$. ◇

Suppose that σ is a permutation. Suppose that $\{a, b\} \in \mathcal{D}(n)$. So a and b lie in $[\![n]\!]$ and have $a \neq b$. Let $a' = \sigma(a)$ and $b' = \sigma(b)$. Note that a' and b' again lie in $[\![n]\!]$. Also, as σ is a bijection, we have that $a' \neq b'$. Thus σ induces a bijection on $\mathcal{D}(n)$, sending $\{a, b\}$ to $\{a', b'\}$.

Definition 24.3. Suppose that $\sigma \in \textsc{Sym}(n)$ is a permutation. Suppose that $a, b \in [\![n]\!]$ have $a \neq b$. Let $a' = \sigma(a)$ and $b' = \sigma(b)$. We say that σ has an *inversion* at $\{a, b\}$ if either

- $a < b$ and $b' < a'$ or
- $b < a$ and $a' < b'$. ◇

We define $\textsc{inv}_{\{a,b\}} \colon \textsc{Sym}(n) \to \{0, 1\}$ by

$$\textsc{inv}_{\{a,b\}}(\sigma) = \begin{cases} 1, & \text{if } \sigma \text{ has an inversion at } \{a, b\} \\ 0, & \text{otherwise} \end{cases}$$

We define $\textsc{inv} \colon \textsc{Sym}(\sigma) \to \mathbb{N}$ by

$$\textsc{inv}(\sigma) = \sum \textsc{inv}_{\{a,b\}}(\sigma)$$

where the sum is over the elements $\{a, b\}$ in $\mathcal{D}(n)$. So $\textsc{inv}(\sigma)$ is the total number of inversions of σ. As an example, the permutation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 0 & 3 & 2 & 4 \end{pmatrix}$$

has two inversions, at $\{0, 1\}$ and $\{2, 3\}$. The permutation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 2 & 0 & 4 & 3 \end{pmatrix}$$

has three inversions, at $\{0,2\}$, $\{1,2\}$, and $\{3,4\}$.

Exercise 24.4. Suppose that $\tau \in \mathrm{SYM}(n)$ is a transposition. Give a direct proof that $\mathrm{INV}(\tau)$ is odd. ◇

Exercise 24.5. Suppose that $\sigma \in \mathrm{SYM}(n)$ is a permutation. Prove that $\mathrm{INV}(\sigma^{-1}) = \mathrm{INV}(\sigma)$. ◇

Exercise 24.6.

(1) Find all σ in $\mathrm{SYM}(n)$ with $\mathrm{INV}(\sigma) = 0$.
(2) Find all σ in $\mathrm{SYM}(n)$ with $\mathrm{INV}(\sigma) = n(n-1)/2$. ◇

Another nice problem is to count the number of permutations in $\mathrm{SYM}(n)$ with exactly k inversions.

24.7. **Parity.**

Definition 24.8. We define the function

$$\mathrm{pari}\colon \mathrm{SYM}(n) \to \mathbb{Z}/2\mathbb{Z} \qquad \text{by} \qquad \mathrm{pari}(\sigma) = [\mathrm{INV}(\sigma)]_2$$

That is, the *parity* of σ is the parity of its number of inversions. ◇

We can now state and prove a theorem that is crucial in the definition of the *determinant* of an n–by–n square matrix.

Theorem 24.9. *The function* $\mathrm{pari}\colon \mathrm{SYM}(n) \to \mathbb{Z}/2\mathbb{Z}$ *is a homomorphism.*

Proof. Suppose that σ and τ are permutations in $\mathrm{SYM}(n)$. We must prove that

$$\mathrm{pari}(\tau \circ \sigma) = \mathrm{pari}(\tau) + \mathrm{pari}(\sigma)$$

where the addition is taken modulo two.

Suppose that $\{a,b\}$ lies in $\mathcal{D}(n)$. Take $a' = \sigma(a)$ and $b' = \sigma(b)$. Using this notation, we now claim that

$$(+) \qquad \mathrm{INV}_{\{a,b\}}(\tau \circ \sigma) = \mathrm{INV}_{\{a',b'\}}(\tau) + \mathrm{INV}_{\{a,b\}}(\sigma)$$

modulo two. This claim proves the theorem, because

- INV is the sum of the functions $\mathrm{INV}_{\{a,b\}}$ and
- the function induced by σ on $\mathcal{D}(n)$ is a bijection.

To prove the claim we fix $\{a,b\}$ in $\mathcal{D}(n)$. We take $a' = \sigma(a)$ and $b' = \sigma(b)$. Also, we take $a'' = \tau(a')$ and $b'' = \tau(b')$. Suppose that

$$a < b \qquad a' < b' \qquad a'' < b''$$

In this case $\mathrm{INV}_{\{a,b\}}(\sigma)$, $\mathrm{INV}_{\{a',b'\}}(\tau)$, and $\mathrm{INV}_{\{a,b\}}(\tau \circ \sigma)$ are zero. Thus the equality $(+)$ holds.

If we switch any one of the inequalities above then exactly two of the counts switch from zero to one. Thus the equality $(+)$ continues to hold.

In general, every time we switch one of the inequalities exactly two of the counts change parity. Thus the equality (+) always holds. □

Corollary 24.10. *Suppose that σ is a permutation in* $\textsc{Sym}(n)$*. Then, in any expression of σ as a product of transpositions, the parity of the number of transpositions equals* $\operatorname{pari}(\sigma)$.

Proof. By Exercise 24.4, the parity of any transposition is odd. By Theorem 24.9, the parity of a product of permutations is the sum of their parities. Finally, the parity of a sum of odd numbers is the parity of the sum. □

Example 24.11. Suppose that σ is a k–cycle. Applying the solution to Exercise 22.9, we find that σ is a composition of $k-1$ transpositions. Thus, by Corollary 24.10, the parity of a k–cycle is $[k-1]_2$. ◇

25. Images, kernels, cosets, and quotients

Taking subgroups is one way to produce new groups from old. Here we develop the tools needed to produce new from old by taking quotients.

25.1. **Images and kernels.**

Definition 25.2. Suppose that $\phi\colon G \to H$ is a homomorphism. We define

$$\operatorname{imag}(\phi) = \phi(G) = \{\phi(g) \mid g \in G\}$$

to be the *image* of ϕ. ◇

Example 25.3. Consider the additive group $(\mathbb{Z}, +)$. Then the "multiplication map" $p_n\colon \mathbb{Z} \to \mathbb{Z}$ defined by $p_n(k) = n \cdot k$ is a homomorphism. The image of p_n is $\operatorname{image}(p_n) = \langle n \rangle = n\mathbb{Z}$, the subgroup of $\mathbb{Z}$ generated by n. ◇

Lemma 25.4. *Suppose that $\phi\colon G \to H$ is a homomorphism. Then the image $\phi(G)$ is a subgroup of H.*

Proof. By Lemma 23.3 we have $\phi(e_G) = e_H$. Thus e_H lies in the image of ϕ.

Suppose that h lies in the image of ϕ. Thus there is some g in G so that $\phi(g) = h$. By Lemma 23.3 we have $\phi(g^{-1}) = h^{-1}$. Thus h^{-1} lies in the image of ϕ.

Suppose that h and h' lie in the image of ϕ. Thus there are elements g and g' in G so that $\phi(g) = h$ and $\phi(g') = h'$. Since ϕ is a homomorphism we have that $\phi(g \cdot g') = h \cdot h'$. Thus $h \cdot h'$ lies in image of ϕ. □

We have already seen a very special case of Lemma 25.4 in Lemma 23.3.

Definition 25.5. We define

$$\ker(\phi) = \phi^{-1}(\{e_H\}) = \{g \in G \mid \phi(g) = e_H\}$$

to be the *kernel* of ϕ. ◇

Example 25.6. Consider the additive groups $(\mathbb{Z}, +)$ and $(\mathbb{Z}/n\mathbb{Z}, +_n)$. Then the quotient map $q_n \colon \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ is a homomorphism. The kernel of q_n is $\ker(p_n) = \langle n \rangle = n\mathbb{Z}$: the subgroup of $\mathbb{Z}$ generated by n. ◇

Lemma 25.7. *Suppose that $\phi \colon G \to H$ is a homomorphism. Then the kernel of ϕ is a subgroup of G.*

Proof. By Lemma 23.3 we have $\phi(e_G) = e_H$. Thus e_G lies in the kernel of ϕ.

Suppose that g lies in the image of ϕ. Thus $\phi(g) = e_H$. Thus

$$\begin{aligned} \phi(g^{-1}) &= \phi(g)^{-1} && \text{Lemma 23.3} \\ &= e_H^{-1} && g \text{ lies in the kernel} \\ &= e_H && \text{Exercise 19.5} \end{aligned}$$

So g^{-1} lies in the kernel of ϕ.

Suppose that g and g' lie in the kernel of ϕ. Thus

$$\begin{aligned} \phi(g \cdot g') &= \phi(g) \cdot \phi(g') && \phi \text{ is a homomorphism} \\ &= e_H \cdot e_H && g \text{ and } g' \text{ lie in the kernel} \\ &= e_H && \text{identity} \end{aligned}$$

So g lies in the kernel of ϕ. □

Example 25.8. The kernel of the parity homomorphism $\mathrm{pari} \colon \textsc{Sym}(n) \to \mathbb{Z}/2\mathbb{Z}$ is called the *alternating group* $\textsc{Alt}(n)$. That is, the elements of $\textsc{Alt}(n)$ are the permutations with even parity. ◇

Exercise 25.9. Suppose that $\phi \colon G \to H$ is a homomorphism. Prove the following.

- Suppose that K is a subgroup of G. Then the image $\phi(K)$ is a subgroup of H. (This is a generalisation of Lemma 25.4.)
- Suppose that K is a subgroup of H. Then the preimage $\phi^{-1}(K)$ is a subgroup of G. (This is a generalisation of Lemma 25.7.)

◇

Exercise 25.10. Suppose that ζ is a non-zero complex number. Let $Z = \langle \zeta \rangle$ be the cyclic subgroup of $(\mathbb{C}^\times, \cdot)$ generated by ζ. Let $\exp \colon (\mathbb{C}, +) \to (\mathbb{C}^\times, \cdot)$ be the homomorphism defined in Example 23.13. Discuss the subgroup $L = \exp^{-1}(Z)$. ◇

25.11. **Cosets.**

Definition 25.12. Suppose that G is a group. Suppose that H is a subgroup. Suppose that g is an element of G. We define two sets.

$$gH = \{gh \mid h \in H\} \qquad Hg = \{hg \mid h \in H\}$$

We call these, respectively, a *left* and a *right coset* of H inside of G. We call any $g' \in gH$ a *representative* of the coset. We use similar language for right cosets. ◇

There is a symmetry between left and right cosets. Accordingly, we focus on left cosets. Recall that $L_g : G \to G$ is left multiplication by g. We begin with a corollary of Lemma 20.2.

Lemma 25.13. *Suppose that gH and fH are left cosets of H in G. Then $L_{fg^{-1}} : gH \to fH$ is a bijection.* □

Thus, all cosets gH have the same cardinality.

Lemma 25.14. *Suppose that G is a group. Suppose that H is a subgroup. Then the set*

$$G/H = \{gH \mid g \in G\}$$

is a partition of G.

Said another way, "belonging to a common coset" is an equivalence relation on G. We call G/H the *quotient* of G by H.

Proof of Lemma 25.14. Note that e_G lies in H; thus g lies in gH. Thus the union of the left cosets is all of G.

Suppose that f lies in $gH \cap g'H$. Thus $f = gh$ and $f = g'h'$ for some elements h and h' of H. So $g' = gh(h')^{-1}$. Set $k = h(h')^{-1}$. Since H is a subgroup, k lies in H (Definition 20.8).

Suppose now that h is any element of H. Thus $gh = gkk^{-1}h = g'k^{-1}h$ lies in $g'H$. Thus gH is a subset of $g'H$. A similar argument shows that $g'H$ is a subset of gH. Thus $gH = g'H$ by Lemma 1.16. □

We now have enough machinery to state and prove a result of Jordan [9, page 166], generalising a result of Lagrange [10, page 202].

Theorem 25.15. *Suppose that G is a finite group. Suppose that $H < G$ is a subgroup. Then the order of H divides the order of G.*

Proof. The subgroup H is finite by the first half of Lemma 17.12. The quotient G/H is finite by the second half of Lemma 17.12. Applying Lemmas 25.13 and 25.14, as well as the representatives picked in the proof of Lemma 17.12, we build a bijection between G and $H \times G/H$. By Exercise 4.8 we have that $|G| = |H| \times |G/H|$. Thus $|H|$ divides $|G|$, as desired. □

The theorem of Lagrange is an important special case.

Corollary 25.16. *Suppose that G is a finite group. Suppose that $g \in G$ is an element. Then the order of g divides the order of G.* □

Remark 25.17. As a further application when p is prime the group $G = \mathbb{Z}/p\mathbb{Z}$ has

- exactly two subgroups and
- every non-zero element having order p.

In particular, any $k \in \mathbb{Z}/p\mathbb{Z}$ (other than zero) generates the entire group. ◇

25.18. **Normal subgroups.** This subsection is not examinable.

Definition 25.19. Suppose that G is a group. Suppose that $H < G$ is a subgroup. Suppose that, for all g, we have $gH = Hg$. Then we call H a *normal subgroup* and we write $H \triangleleft G$. ◇

Exercise 25.20. Suppose that G is a commutative group. Suppose that $H < G$ is a subgroup. Prove that H is normal in G. ◇

Suppose that G is a group and $H < G$ is a subgroup. Given cosets gH and $g'H$ we define

$$gH \cdot g'H = \{ghg'h' \mid h, h' \in H\}$$

In general, $gH \cdot g'H$ need not be a coset. However, if H is normal, then it is, and we have the following.

Lemma 25.21. *Suppose G is a group and that $H \triangleleft G$ is a normal subgroup. Then, for all g and g' in G we have $gH \cdot g'H = (gg')H$. It follows that G/H, together with the product operation on cosets, is a group.*

Proof. Since $g'H = Hg'$ we have that for any h there is some h'' so that $hg' = g'h''$. In particular, we have $ghg'h' = gg'h''h' \in (gg')H$. We deduce that $gH \cdot g'H = (gg')H$.

The coset $e_G H = H$ gives the identity element. The coset $g^{-1}H$ gives the inverse of gH. The product of cosets is again a coset by the above. Finally, associativity in G/H follows from associativity in G. □

Example 25.22. Lemma 19.13 implies that the rotation subgroup E_n is a normal subgroup of the dihedral group D_{2n}. The quotient D_{2n}/E_n is isomorphic to $\mathbb{Z}/2\mathbb{Z}$. The image of $g \in D_{2n}$ inside the quotient records whether or not g preserves the "orientation" of the plane $\mathbb{R}^2$. ◇

Example 25.23. Exercise 25.20 implies that the integers $\mathbb{Z}$ (equipped with addition) is a normal subgroup of the real numbers $\mathbb{R}$ (also equipped

with addition). Let $C = \mathbb{R}/\mathbb{Z}$ be the resulting quotient group. We define $S^1 \subset \mathbb{C}^\times$ to be the complex numbers of the form $\cos(\theta) + i\sin(\theta)$. It is an exercise to show that this is a subgroup of $\mathbb{C}^\times$. In fact, $C = \mathbb{R}/\mathbb{Z}$ is isomorphic to S^1 via the function taking $[t]_{\mathbb{Z}}$ to $\exp(2\pi i \cdot t)$. ◇

26. Rings and their arithmetic

In the next two sections we discuss two objects central in algebra: rings and fields.

26.1. **Rings.**

Definition 26.2. A *ring* R is a set (again denoted R) together with two operations $+$ and $\cdot$ (called *addition* and *multiplication*) with the following properties.

- $(R, +)$ is a commutative group with an additive identity 0_R.
- $(R, \cdot)$ is associative and has a multiplicative identity 1_R.
- Multiplication distributes over addition, on both sides. That is, for all $p, q, r \in R$ we have

$$p \cdot (q + r) = p \cdot q + p \cdot r$$
$$(q + r) \cdot p = q \cdot p + r \cdot p$$

◇

Note that, in the definition of a ring, we do *not* assume that $(R, \cdot)$ is a group. We also do not assume that the multiplication is commutative.

For (non-)examples of rings we consider the naturals, the integers, the rationals, the reals, and the complexes, all equipped with their usual addition and multiplication.

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

Addition and multiplication on $\mathbb{N}$ are commutative, associative, distributive, and have the usual identities. All of these are proved by induction. However, $\mathbb{N}$ is not a ring because it does not have additive inverses.

On the other hand, the integers, the rationals, the reals, and the complexes are rings. In each case, with a bit of work, the proofs reduce to the previous case.

Definition 26.3. Suppose that R is a ring. If the multiplication in R is commutative then we call R a *commutative ring*. ◇

The integers, the rationals, the reals, and the complexes are commutative rings. This essentially follows from the commutativity of multiplication in $\mathbb{N}$. Here is another, very useful, example of a commutative ring.

Definition 26.4. Suppose that R is a commutative ring. Suppose that x is a variable. Suppose that d is a natural number. We take x^d to be the *monomial* of degree d. As special cases, we have $x^0 = 1_R$ and $x^1 = x$. For any $r \in R$ we call the formal product $r \cdot x^d$ a *term*.

A *polynomial over* R is a finite formal sum of terms. In more detail, for any polynomial P there are finitely many coefficients $r_k \in R$ so that

$$P(x) = \sum_{k=0}^{d} r_k x^k = r_0 + r_1 x + r_2 x^2 + \cdots + r_d x^d$$

If r_d is non-zero then we say d is the *degree* of P.[14]

Addition of polynomials is done term-by-term. Multiplication is performed by multiplying all terms and then collecting coefficients of monomials of like degree. The result is the *polynomial ring* $R[x]$ (pronounced "R adjoin x"). ◇

We end this section with an example of a non-commutative ring.

Example 26.5. Let $M_n(\mathbb{C})$ be the set of n–by–n matrices with complex entries. Equipped with matrix addition and multiplication this forms a ring. When $n > 1$ this ring is not commutative; we verify this for $n = 2$ as follows.

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

We now compute the two possible products.

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Since $AB \neq BA$, the ring $M_2(\mathbb{C})$ is not commutative. ◇

26.6. **Arithmetic in rings.** Much of the usual arithmetic of the integers carries over to general rings. As examples, we have the following.

Lemma 26.7. *Suppose that R is a ring. Then, for any $r \in R$ we have $0_R \cdot r = r \cdot 0_R = 0_R$.*

[14] If all r_i are zero then we obtain the *zero polynomial*. The degree of the zero polynomial is either left undefined, or taken to be negative infinity.

Proof. We compute as follows.

$$\begin{aligned}
0_R \cdot r &= 0_R \cdot r + 0_R && \text{additive identity} \\
&= 0_R \cdot r + (0_R \cdot r + (-(0_R \cdot r))) && \text{additive inverse} \\
&= (0_R \cdot r + 0_R \cdot r) + (-(0_R \cdot r)) && \text{addition associative} \\
&= (0_R + 0_R) \cdot r + (-(0_R \cdot r)) && \text{distributive} \\
&= 0_R \cdot r + (-(0_R \cdot r)) && \text{additive identity} \\
&= 0_R && \text{additive inverse}
\end{aligned}$$

The other equality is obtained similarly. □

Lemma 26.8. *Suppose that R is a ring. Suppose that -1_R is the additive inverse of 1_R. Then for any $r \in R$ we have that $(-1_R)r$ is the additive inverse of r.*

Proof. We compute as follows.

$$\begin{aligned}
r + (-1_R)r &= (1_R)r + (-1_R)r && \text{multiplicative identity} \\
&= (1_R + (-1_R))r && \text{distributive} \\
&= 0_R \cdot r && \text{additive inverse} \\
&= 0_R && \text{Lemma 26.7}
\end{aligned}$$

Thus $(-1_R)r$ is the additive inverse of r. □

27. Fields: definitions and examples

Our next algebraic object is of particular importance in linear algebra.

27.1. **Fields.** For any ring R we define $R^\times = R - \{0_R\}$.

Definition 27.2. Suppose that R is a ring. If $(R^\times, \cdot)$ is a commutative group then we call R a *field*. ◇

That is, if R is a commutative ring where non-zero elements have multiplicative inverses, then R is a field. We now review the most common (non-)examples, all considered with their usual addition and multiplication.

- $\mathbb{N}$ is not a ring, so is not a field.
- $\mathbb{Z}$ is a commutative ring, but is not a field.
- $\mathbb{Q}$, $\mathbb{R}$, and CC are all fields.

27.3. **Finite fields.** Lemma 10.14 gives us our next family of rings.

Lemma 27.4. *$(\mathbb{Z}/n\mathbb{Z}, +_n, \cdot_n)$ is a commutative ring.* □

Among these we can find a family of fields.

Theorem 27.5. *Suppose that $n > 1$. Then $(\mathbb{Z}/n\mathbb{Z}, +_n, \cdot_n)$ is a field if and only if n is prime.*

Proof. We prove the contrapositive of the forward direction. Suppose that $n > 1$ is not prime. From Definition 14.3 we obtain natural numbers a and b so that $1 < a, b < n$ and $ab = n$. By Exercise 10.10 we know that the congruence classes $[a]_n$ and $[b]_n$ are not equal to $[0]_n$. On the other hand, we have $[a]_n \cdot_n [b]_n = [0]_n$. Thus $(\mathbb{Z}/n\mathbb{Z})^\times$ is not closed under multiplication. Thus $((\mathbb{Z}/n\mathbb{Z})^\times, \cdot_n)$ is not a group. Thus $\mathbb{Z}/n\mathbb{Z}$ is not a field.

We now prove the backwards direction. Suppose that n is prime. Suppose that $k \in [\![n]\!]$ is not zero. By Remark 25.17 the element k generates $\mathbb{Z}/n\mathbb{Z}$. That is, the multiplies of k give all elements of $\mathbb{Z}/n\mathbb{Z}$. So there is some ℓ in $\mathbb{N}$ so that

$$[k]_n \cdot_n [\ell]_n = [k \cdot \ell]_n = [1]_n$$

So $[\ell]_n$ is a multiplicative inverse for $[k]_n$.

All that remains is to prove that $[\ell]_n$ is not zero. For a contradiction, suppose that $[\ell]_n = [0]_n$. Then, by Lemma 26.7, we have that

$$[k]_n \cdot [\ell]_n = [k]_n \cdot [0]_n = [0]_n$$

Thus $[1]_n = [0]_n$. However this contradicts Exercise 10.10. □

One application of Theorem 27.5 is *solving congruences*.

Definition 27.6. A *linear congruence* is an equation of the form

$$ax \equiv b \pmod{n}$$

where a, b, and n are (fixed) integers and x is a variable ranging over the integers. The *solutions* to the congruence are the integers x making the equation hold. ◇

Example 27.7. Consider the following congruence.

$$2x \equiv 3 \pmod{5}$$

Trial and error tells us that $x = 4$ is one solution. Since we are working modulo 5 any integer of the form $4 + 5k$ is also a solution. In fact these are the only solutions.

To see this, we note that 5 is prime. Thus Theorem 27.5 promises us that 2 has a multiplicative inverse modulo 5. Trial and error shows us that 3 is the desired inverse. So we may multiply both sides of the

given congruence by 3 and not lose or gain any solutions. Thus the given congruence has the same set of solutions as

$$6x \equiv 9 \pmod 5$$

which has the same set of solutions as

$$x \equiv 4 \pmod 5$$

That is, the set of solutions is a congruence class.

$$[4]_5 = \{4 + 5k \mid k \in \mathbb{Z}\} \qquad \diamond$$

Some linear congruences do not have solutions; for example

$$2x \equiv 1 \pmod 8$$

has no integer solutions.

After proving Theorem 29.9 we will give a method to solve congruences which avoids trial and error.

Exercise 27.8. Find all solutions to the following linear congruences.

(1) $8x \equiv 1 \pmod 7$
(2) $3x \equiv 1 \pmod 7$
(3) $3x \equiv 2 \pmod 7$
(4) $3x \equiv 2 \pmod{13}$
(5) $6x \equiv 4 \pmod{26}$ $\diamond$

28. Greatest common divisors and the euclidean algorithm

28.1. **Greatest common divisors.**

Notation 28.2. Suppose that n is a natural number. Define

$$\mathcal{D}(n) = \{d \in \mathbb{N} \mid d \text{ divides } n\}$$

That is, $\mathcal{D}(n)$ is the set of divisors of n. $\diamond$

Note that $\mathcal{D}(n)$ is always non-empty (as it contains 1) and finite (if $n > 0$). Also, if $n > 0$ and d lies in $\mathcal{D}(n)$, then so does n/d. As examples we have

$$\mathcal{D}(100) = \{1, 2, 4, 5, 10, 20, 25, 50, 100\}$$
$$\mathcal{D}(36) = \{1, 2, 3, 4, 6, 9, 12, 18, 36\}$$

Definition 28.3. Suppose that a and b are natural numbers. We define

$$\gcd(a, b) = \begin{cases} 0 & \text{if } a = b = 0 \\ \max \mathcal{D}(a) \cap \mathcal{D}(b) & \text{otherwise} \end{cases}$$

We call $\gcd(a, b)$ the *greatest common divisor* of a and b. $\diamond$

For example, $\mathcal{D}(100) \cap \mathcal{D}(36) = \{1, 2, 4\}$ so $\gcd(100, 36) = 4$. Note that gcd is a function from $\mathbb{N}^2$ to $\mathbb{N}$. We can extend the definition to all of $\mathbb{Z}^2$ by defining $\gcd(a, b) = \gcd(|a|, |b|)$.

28.4. **Computing greatest common divisors by factoring.** Naturally, there is a direct relationship between the fundamental theorem of arithmetic (Theorem 17.3), giving prime factorisations, and computing greatest common divisors.

Theorem 28.5. *Suppose that a and b are natural numbers, both non-zero. Suppose that their prime factorisations are as follows.*

$$a = \prod p_i^{\alpha_i} \qquad b = \prod p_i^{\beta_i}$$

Then

$$\gcd(a, b) = \prod p_i^{\min(\alpha_i, \beta_i)} \qquad \square$$

As an example, suppose that $a = 100$ and $b = 36$. Note that $100 = 2^2 \cdot 3^0 \cdot 5^2$ and $36 = 2^2 \cdot 3^2 \cdot 5^0$. Thus $\gcd(100, 36) = 2^2 \cdot 3^0 \cdot 5^0 = 4$.

As a corollary of Theorem 28.5, if a and b are distinct primes, then $\gcd(a, b) = 1$.

Definition 28.6. Suppose that a and b are natural numbers. If $\gcd(a, b) = 1$ then we say that a and b are *relatively prime*. $\diamond$

Theorem 28.5 gives a quick and easy method for computing the greatest common divisor of a given a and b when at least one of them is reasonably small. On the other hand, when both a and b are large this method can be very slow. The following question has been outstanding for hundreds of years [7, Article 329].

Problem 28.7. Is there an efficient algorithm for factoring natural numbers?

Surprisingly, there is an efficient algorithm if one is allowed to use a quantum computer [14].

28.8. **The slow euclidean algorithm.** Before giving our first (slow) version of the euclidean algorithm, we require a bit of symmetry.

Lemma 28.9. *Suppose that a and b lie in $\mathbb{N}$. Then we have the following.*

(1) $\gcd(a, 0) = a$.
(2) $\gcd(a, 1) = 1$.
(3) $\gcd(a, b) = \gcd(b, a)$.
(4) $\gcd(a, b) = \gcd(a + b, b)$.

Proof. The proofs of the first three claims follow from the definitions.

To prove the last we note that d divides $\gcd(a,b)$ if and only if d divides a and divides b. This happens if and only if d divides $a+b$ and divides b. This happens if and only if d divides $\gcd(a+b,b)$.

We deduce that each of $\gcd(a,b)$ and $\gcd(a+b,b)$ divides the other. Thus they are equal. □

For example, subtracting gives $100-36=64$ and $64-36=28$. So, applying (4) twice, we have that $\gcd(100,36)=\gcd(64,36)=\gcd(28,36)$. Furthermore, applying (3) we can swap to find $\gcd(28,36)=\gcd(36,28)$. At this point we can again subtract. Continuing in this way produces the following table.

n	A_n	B_n	do
0	100	36	subtract
1	64	36	subtract
2	28	36	swap
3	36	28	subtract
4	8	28	swap
5	28	8	subtract
6	20	8	subtract
7	12	8	subtract
8	4	8	swap
9	8	4	subtract
10	4	4	subtract
11	0	4	swap
12	4	0	set $N=12$, return 4, and halt

By Lemma 28.9(4) and (3), the greatest common divisor of all of the pairs (A_n,B_n) are the same, and thus all equal to the last, giving $\gcd(100,36)=\gcd(4,0)=4$.

Here is the formal description of the slow algorithm; it is essentially the same as the algorithm given in [5, Book 7, Propositions 1 and 2].

Algorithm 28.10. We are given a pair of natural numbers a and b. We set $n=0$, $A_0=a$, and $B_0=b$.

While B_n is non-zero we do the following.

(1) If $A_n \geq B_n$, we *subtract*; that is, we take

$$A_{n+1}=A_n-B_n \qquad \text{and} \qquad B_{n+1}=B_n$$

(2) If $A_n < B_n$, we *swap*; that is, we take

$$A_{n+1}=B_n \qquad \text{and} \qquad B_{n+1}=A_n$$

(3) In either case, we add one to n.

If $B_n=0$, we set $N=n$, return A_N, and halt. ◇

Theorem 28.11. *Algorithm 28.10 terminates with* $A_N=\gcd(a,b)$.

Proof. Note that after a subtraction, we have $A_{n+1}+B_{n+1}=A_n<A_n+B_n$. Also, every swap (other than the last) is followed by a subtraction. Thus the algorithm halts by the well-ordering principle.

By parts (4) and (3) of Lemma 28.9, for every n we have

$$\gcd(A_{n+1}, B_{n+1}) = \gcd(A_n, B_n)$$

Thus, by induction, we have $\gcd(A_N, B_N) = \gcd(A_0, B_0)$. Finally, since $B_N = 0$, by Lemma 28.9(1) we have $\gcd(A_N, 0) = A_N$. Thus A_N is the greatest common divisor of a and b, as desired. □

Remark 28.12. Since $A_{n+1} + B_{n+1} = A_n < A_n + B_n$ after a subtraction, we perform at most $A_0 + B_0 = a + b$ subtractions in total. Since we never do two swaps in a row, the algorithm finishes after at most $2(a + b) + 1$ steps. ◇

29. The fast euclidean algorithm

In this section we give a much faster version of the euclidean algorithm.

29.1. **Division with remainder.** We reformulate Exercise 10.10.

Theorem 29.2. *Suppose that a and b are natural numbers with $b > 0$. Then there are unique natural numbers q and r so that $r < b$ and $a = qb + r$.*

Proof. Consider the set

$$\mathcal{R} = \{r \in \mathbb{N} \mid \text{there is some } q \in \mathbb{N} \text{ with } r = a - qb\}$$

Note that a lies in $\mathcal{R}$ (when $q = 0$) so $\mathcal{R}$ is non-empty. By the well-ordering principle we have that $\mathcal{R}$ has a smallest element, say r. Let q be the given natural number so that $r = a - qb$.

For a contradiction, suppose that $r \geq b$. Then $r - b \geq 0$. Also, $r - b = a - qb - b = a - (q + 1)b$. Thus $r - b$ lies in $\mathcal{R}$. Since $r - b < r$ we deduce that r was not the smallest element of $\mathcal{R}$, a contradiction.

Suppose that q' and r' are another pair of natural numbers with the same properties as q and r: namely $r' < b$ and $a = q'b + r'$. Thus r' lies in $\mathcal{R}$. Since r is a least element of $\mathcal{R}$ we have $r \leq r'$. Thus $r' - r$ is non-negative. Also $r' - r \leq r' < b$. Since $a = qb + r = q'b + r'$ we subtract and find that $r' - r = (q - q')b$. Thus $r - r'$ is a multiple of b. However, the only non-negative multiple of b less than b is zero. Thus $r = r'$, as desired. □

From this, induction, and Lemma 28.9 we deduce the following.

Corollary 29.3. *Suppose that a and b are natural numbers with $b > 0$. Suppose that q and r are given by Theorem 29.2. Then* $\gcd(a, b) = \gcd(b, a - qb) = \gcd(b, r)$. □

29.4. **The fast euclidean algorithm.** Before stating the fast euclidean algorithm we repeat our running example. Suppose that $a = 100$ and $b = 36$. Then $100 = 2 \cdot 26 + 28$. That is, division with remainder gives $q = 2$ and $r = 28$. This value of r appeared in step 3 of the slow euclidean algorithm. We also have, by Corollary 29.3, that $\gcd(100, 36) = \gcd(36, 28)$. Thus division with remainder replaces the first three steps of the slow algorithm with a single step. Here are all the steps of the fast algorithm, as a table.

n	A_n	B_n	q_n	r_n
0	100	36	2	28
1	36	28	1	8
2	28	8	3	4
3	8	4	2	0
4	4	0	–	–

By Corollary 29.3, the greatest common divisor of all of the pairs (A_n, B_n) are the same, and thus all equal to the last, giving $\gcd(100, 36) = \gcd(4, 0) = 4$. Note that the number of steps here counts the number of swaps in the slow algorithm. Also, the sum of the q_n counts the number of subtractions in the slow algorithm.

Here is the formal description of the fast algorithm.

Algorithm 29.5. We are given a pair of natural numbers a and b. We set $n = 0$, $A_0 = a$, and $B_0 = b$.

While B_n is non-zero we do the following.

(1) Let q_n and r_n be the numbers provided by division with remainder applied to A_n and B_n. We now *pivot forward*; that is, we take
$$A_{n+1} = B_n \qquad \text{and} \qquad B_{n+1} = A_n - q_n B_n$$
That is, $B_{n+1} = r_n$.

(2) Add one to n.

If $B_n = 0$, we set $N = n$, return A_N, and halt. ◇

The termination and correctness proof for Algorithm 29.5 is similar to that for Algorithm 28.10.

Theorem 29.6. *Algorithm 29.5 terminates with* $A_N = \gcd(a, b)$. □

Remark 29.7. To estimate the number of steps taken by Algorithm 29.5 it is convenient to assume that $0 < b \leq a$. Thus, in every step we have $B_n \leq A_n$ and thus $1 \leq q_n$. Thus we have
$$2r_n < r_n + B_n < r_n + qB_n = A_n$$
We deduce that $A_{n+2} < \frac{1}{2} A_n$. Since $1 \geq A_N$, it follows that N is at most $2 \cdot \log_2(a)$.

We also note, without proof, that the "most difficult" input to Algorithm 29.5 is when a and b are consecutive Fibonacci numbers. For such initial values division with remainder does not offer a speed-up. ◇

Analyses of the complexity of the fast euclidean algorithm date to at least 1811; Shallit [13] gives a detailed and accessible review of the history.

29.8. **Bézout's lemma.** We end with a consequence of the fast algorithm, sometimes called *Bézout's lemma*.[15]

Theorem 29.9. *Suppose that a and b are natural numbers. Then there are integers c and d so that $ad - bc = \gcd(a, b)$.*

To prove this, we need the *backwards euclidean algorithm*.

Algorithm 29.10. We are given a pair of natural numbers a and b. We perform Algorithm 29.5 to obtain the number N and the numbers $(q_n)_{n=0}^{N-1}$. We now set $n = N$, $C_N = 0$, and $D_N = (-1)^N$. While $n > 0$ we do the following.

(1) Given C_{n+1} and D_{n+1} we *pivot back*: that is, we take

$$C_n = q_n C_{n+1} + D_{n+1} \qquad \text{and} \qquad D_n = C_{n+1}$$

When $n = 0$ we return C_0 and D_0, and halt. ◇

Proof of Theorem 29.9. From Theorem 29.6 we have

$$A_N D_N - B_N C_N = (-1)^N A_N = (-1)^N \gcd(a, b)$$

With this as our base case we induct backwards, from N down to zero. We compute as follows.

$$\begin{aligned}
A_n D_n - B_n C_n &= A_n C_{n+1} - B_n(q_n C_{n+1} + D_{n+1}) && \text{definition of } C_n \text{ and } D_n \\
&= A_n C_{n+1} - q_n B_n C_{n+1} - B_n D_{n+1} && \text{algebra} \\
&= -A_{n+1} D_{n+1} + (A_n - q_n B_n) C_{n+1} && \text{definition of } A_{n+1} \\
&= -A_{n+1} D_{n+1} + B_{n+1} C_{n+1} && \text{definition of } B_{n+1} \\
&= (-1)(A_{n+1} D_{n+1} - B_{n+1} C_{n+1}) && \text{algebra} \\
&= (-1)(-1)^{n+1} \gcd(a, b) && \text{induction} \\
&= (-1)^n \gcd(a, b) && \text{algebra}
\end{aligned}$$

We deduce that $A_0 D_0 - B_0 C_0 = (-1)^0 \gcd(a, b) = \gcd(a, b)$. Taking $c = C_0$ and $D_0 = d$ finishes the proof. □

[15] The version we give predates Bézout's (more general) result by at least 155 years [1, Proposition 18, page 18].

Here is one way to understand the fast and the backwards euclidean algorithms. We define the *pivot matrices* P_n and the *Farey matrices* E_n as follows.

$$P_n = \begin{pmatrix} 0 & 1 \\ 1 & -q_n \end{pmatrix} \qquad E_n = \begin{pmatrix} A_n & C_n \\ B_n & D_n \end{pmatrix}$$

These matrices are related to each other as follows

$$\det(P_n) = -1 \qquad \det(E_n) = (-1)^n \gcd(a,b) \qquad E_{n+1} = P_n E_n$$

Induction implies that $E_N = P_{N-1}P_{N-2}\cdots P_1 P_0 E_0$. That is, we have the following.

$$\begin{pmatrix} \gcd(a,b) & 0 \\ 0 & (-1)^N \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -q_{N-1} \end{pmatrix} \cdots \begin{pmatrix} 0 & 1 \\ 1 & -q_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -q_0 \end{pmatrix} \begin{pmatrix} A_0 & C_0 \\ B_0 & D_0 \end{pmatrix}$$

Taking the determinant of both sides then yields Theorem 29.9.

We finish this section with two applications.

Exercise 29.11. Suppose that p is prime, and $a < p$ is a positive natural number. Use Theorem 29.9 to find a multiplicative inverse for a, modulo p. ◇

Exercise 29.12. Suppose that G is a cyclic group. Prove that all subgroups of G are again cyclic. ◇

Appendix A. Solutions to exercises

Solution 29.12. We prove this in the special case that $G = \mathbb{Z}$; the proof when G is finite is a consequence.

Suppose that $H < \mathbb{Z}$ is a subgroup. Let $H_{>0}$ be the positive elements of H. If this is empty then there is nothing to prove. So, appealing to the well-ordering principle, we take a to be a smallest element of $H_{>0}$.

We now claim that $H = \langle a \rangle$ (and so H is cyclic). To prove this, let b be any element of $H_{>0}$. By Theorem 29.9 there are integers c and d so that $ad - bc = \gcd(a,b)$. Thus $\gcd(a,b)$ lies in H. Since $\gcd(a,b) \leq a$ we deduce that $\gcd(a,b) = a$. Thus b is a multiple of a, as desired. □

Solution 1.7. Suppose that X and Y are empty sets. Then every element of X (there are none) belongs to Y. Similarly, every element of Y (there are none) belongs to X. Thus, by the axiom of extension, $X = Y$. □

Solution 1.12.

(1) To prove that $\varnothing \subset X$ we must check, for every $x \in \varnothing$, that x lies in X. Since there are no elements of $\varnothing$, no checks can (or need) be made. Thus all checks (again, there are none) are successful.

Thus $\varnothing \subset X$, as desired. [Such a statement, where nothing needs to be checked, is said to hold *vacuously*.]

(2) To prove that $X \subset X$ we must check, for every $x \in X$, that x lies in X. This holds, so we are done. □

Solution 1.14. Suppose that Y is a subset of $[\![n]\!]$. For each $k < n$, either k is in Y or it is not. All of these possibilities are *independent*: that is, for any collection of n two-fold choices there is such a subset Y. Thus there are $2 \times 2 \times \cdots \times 2 = 2^n$ subsets of $[\![n]\!]$. □

Solution 2.13.

(1) Recall that

$$[\![n]\!] = \{0, 1, 2, \ldots, n-1\}$$

is the set of the first n natural numbers. To represent a function $f\colon [\![2]\!] \to [\![2]\!]$ we list the images as lists of the form $(f(0), f(1))$. Here are all possibilities:

$$(0,0) \quad (0,1) \quad (1,0) \quad (1,1)$$

Note that there are two possibilities for the first entry, and two possibilities for the second, giving a total of $4 = 2 \times 2$ functions overall.

(2) Suppose that f is a function on $[\![n]\!]$. We can specify the rule for f by listing, in order, the elements $f(0)$, $f(1)$, $f(2)$, and so on until $f(n-1)$.[16] There are n possibilities for each, and no other restrictions. Thus the number of possible rules is n, multiplied by itself, n times. Thus the correct count is $n \times n \times n \times \ldots \times n = n^n$. □

Solution 3.5.

(1) The function $f\colon \mathbb{R} \to \mathbb{R}$, where $f(x) = x^2$, is not injective because $f(1) = f(-1)$. It is not surjective because -1 is not a square of a real number.

(2) The function $g\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$, where $g(x) = x^2$, is injective. For, suppose that x and y are non-negative reals. Suppose that $g(x) = g(y)$. Thus $x^2 = y^2$ and so $x^2 - y^2 = 0$. We factor and deduce that $(x+y)(x-y) = 0$. Note that, since x and y are non-negative, then $x + y = 0$ if and only if $x = y = 0$. In this case we are done. Suppose instead that $x + y > 0$. Dividing, we find that $x - y = 0$. Thus $x = y$, as desired.

On the other hand, g is not surjective because -1 is not a square of a positive real number.

[16] Any argument that mentions "and so on" in fact needs *induction*. We discuss this more carefully in Section 16.

(3) The function $h\colon \mathbb{R} \to \mathbb{R}_{\geq 0}$, where $h(x) = x^2$, is not injective, because $h(-1) = h(1)$. It is surjective because every non-negative real number has a real square root. A careful proof of this requires the *completeness* property of the real numbers.
(4) The function $k\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, where $k(x) = x^2$, is a bijection. It is injective for the same reason g is injective. It is surjective for the same reason h is surjective. □

Solution 3.7. Recall that

$$n! = n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1$$

is the *factorial* of n. We claim that the number of permutations of $[\![n]\!]$ is $n!$ (pronounced "n factorial".).

Suppose that $f\colon [\![n]\!] \to [\![n]\!]$ is a permutation. We may represent f as via the list $(f(0), f(1), \ldots, f(n-1))$. There are n possibilities for $f(0)$. Since f is a bijection, it is an injection. Thus we know that $f(1) \neq f(0)$. Thus there are $n-1$ possibilities for $f(1)$. By the same logic, we have that $f(2) \neq f(1)$ and $f(2) \neq f(0)$. Thus there are $n-2$ possibilities for $f(2)$. Continuing in this way, we have $n-k$ possibilities for $f(k)$. Thus, when we reach $f(n-1)$ there is only one possibility for it. We deduce that there are $n! = n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1$ permutations of $[\![n]\!]$. □

Solution 3.13. Suppose that X is a set. Suppose that $g\colon \mathcal{P}(X) \to X$ is an injection. We define a function $f\colon X \to \mathcal{P}(X)$ as follows.

Suppose that $a \in X$. Note that, as g is injective, there is at most one (and perhaps no) subset $A \subset X$ so that $g(A) = a$. That is, either $g(A) = a$ for some unique $A \subset X$ or there is no such A. In the former case we take $f(a) = A$. In the latter case we take $f(a) = \varnothing$. Since g is a function, we find that f is surjective. This contradicts Theorem 3.12. □

Solution 4.8. A solution here relies on our definition of $m \times n$.

A geometric definition takes $m \times n$ to be the number of unit squares in a m-by-n rectangle. The squares are indexed by ordered pairs from $[\![m]\!] \times [\![n]\!]$. This gives a bijection from $[\![m]\!] \times [\![n]\!]$ to the set of squares, and we are done. [This proof relies on our intuitions about the plane $\mathbb{R}^2$; however this is one of the things we were hoping to understand using cartesian products.]

A *recursive* definition instead assumes (of $\times$) that for all m we have:

(1) $m \times 0 = 0$ and
(2) for all n, we have $m \times (n+1) = m \times n + m$.

Now the proof that $[\![m]\!] \times [\![n]\!]$ has the correct cardinality requires induction, and the following lemma: If X and Y are disjoint finite sets then $|X \cup Y| = |X| + |Y|$. □

Solution 4.10. The set $A \times B$ consists of all ordered pairs (a, b) where a lies in A and b lies in B. Suppose that A is the empty set. Then there are no such elements a. Thus there are no such ordered pairs.

The argument when B is empty is similar. □

Solution 4.23.

(1) Here are two binary strings of length five:

$$11001 \quad 01010$$

In general, if $b_0b_1b_2b_3b_4$ is a binary string, then there are two possibilities for each of the bits b_i. Thus the number of binary strings of length five is $2^5 = 32$.

(2) The reasoning above generalises to any $n \in \mathbb{N}$; so there are 2^n binary strings of length n. In particular, there is exactly one binary string of length zero - the empty string $\epsilon_{\mathcal{B}}$.

(3) Suppose that w is a binary string of length n having exactly one bit which is 1. Thus the other bits are all zeros. There are five such strings of length five.

$$10000 \quad 01000 \quad 00100 \quad 00010 \quad 00001$$

In general, the string w is completely determined by the position (also called the *index*) of the non-zero bit. There are n choices for this position, thus there are n such binary strings.

(4) Suppose that w is a binary string of length n having exactly two bits which are 1. Thus the other bits are all zeros. There are ten such strings of length five.

$$\begin{array}{ccccc} 11000 & 10100 & 10010 & 10001 & 01100 \\ 01010 & 01001 & 00110 & 00101 & 00011 \end{array}$$

As in the previous exercise, w is determined by the positions of the non-zero bits. Suppose that these positions are i and j. Since i is a position in w, we have $0 \leq i < n$, and similarly for j. If we swap the roles of i and j we obtain the same string. Thus, to avoid counting this string twice, we introduce the assumption that $i < j$.

We have reduced the problem to that of counting the pairs (i, j) so that $0 \geq i < j < n$. We note that, if i is fixed, then there

are $n-i-1$ possibilities for j. Thus the number of such pairs is

$$\sum_{i=0}^{n-2} n-i-1 = (n-1)+(n-2)+\cdots+2+1 = \frac{n(n-1)}{2}$$

Checking our work with $n=5$ we obtain $(5\times 4)/2 = 10$. □

Solution 5.8. Each proof relies (in a slightly different way) on the similarity between

- how the union operator joins together sets and
- how the word "or" joins together properties.

We will also use Lemma 1.16 repeatedly; that is, to prove an equality of sets it suffices to prove that each contains the other.

(1) Taking $Y = \varnothing$ we apply Lemma 5.6 and deduce that $X \subset X \cup \varnothing$.
To prove the opposite inclusion, suppose that x lies in $X \cup \varnothing$. Thus x lies in X or the empty set. However, the empty set contains no elements (Definition 1.6) thus x lies in X. Thus $X \cup \varnothing \subset X$. By Lemma 1.16 we are done.

(2) Suppose that x lies in $(X \cup Y) \cup Z$. Then x lies in X or Y, or in Z. We deduce that x lies in X, or in Y or Z. Thus x lies in $X \cup (Y \cup Z)$. Thus $(X \cup Y) \cup Z \subset X \cup (Y \cup Z)$.
For the opposite inclusion, suppose that x lies in $X \cup (Y \cup Z)$. Then x lies in X, or in Y or Z. We deduce that x lies in X or Y, or in Z. Thus x lies in $(X \cup Y) \cup Z$. Thus $X \cup (Y \cup Z) \subset (X \cup Y) \cup Z$. By Lemma 1.16 we are done.

(3) Suppose that x lies in $X \cup Y$. Thus x lies in X or in Y. That is, x lies in X or x lies in Y. Thus x lies in Y or x lies in X. Thus x lies in Y or in X. Thus $X \cup Y \subset Y \cup X$.
The opposite inclusion is proved in the same way, swapping the roles of X and Y.

(4) Suppose that $X \subset Y$. Suppose that x lies in $X \cup Y$. Then x lies in X or in Y. Since $X \subset Y$ in either case x lies in Y. Thus $X \cup Y \subset Y$. The opposite inclusion follows from Lemma 5.6. From Lemma 1.16 we deduce that $X \cup Y = Y$.
Suppose that $X \cup Y = Y$. By Lemma 5.6 we have that $X \subset X \cup Y$. Since $X \cup Y$ we deduce that $X \subset Y$.

(5) By Lemma 5.6 we have that $X \subset X \cup X$.
For the opposite inclusion, suppose that x lies in X or in X. That is, x lies in X or x lies in X. Thus x lies in X. We deduce that $X \cup X \subset X$. From Lemma 1.16 we deduce that $X \cup X = X$. □

Solution 5.21. By Lemma 1.16 to prove an equality it suffices to check two inclusions.

Suppose that x lies in $X \cap (Y \cup Z)$. Thus x lies in X and x lies in Y or Z. Thus x lies in X and Y or x lies in X and Z. Thus x lies in $X \cap Y$ or x lies in $X \cap Z$. Thus x lies in $(X \cap Y) \cup (X \cap Z)$ as desired.

Suppose that x lies in $(X \cap Y) \cup (X \cap Z)$. Then we reverse the above steps to find that x lies in $X \cap (Y \cup Z)$, as desired.

The proof that union distributes over intersection is similar, and we omit it. □

Solution 5.23. By Lemma 1.16 to prove an equality it suffices to check two inclusions.

Suppose that x lies in $X - (Y \cup Z)$. Thus x lies in X and does not lie in Y or in Z. Since x is not in Y or in Z we deduce that x is not in Y and x is not in Z. Thus x lies in X and not in Y and not in Z. Thus x lies in X and not in Y and lies in X and not in Z. Thus x lies in $X - Y$ and lies in $X - Z$. This x lies in $(X - Y) \cap (X - Z)$, as desired.

Suppose that x lies in $(X - Y) \cap (X - Z)$. Reversing the above steps shows that x lies in $X - (Y \cup Z)$, as desired.

The proof of the second equality is similar, and we omit it. □

Solution 6.6. Suppose that f is surjective. Thus, for each $y \in Y$ there is some $x \in X$ so that $f(x) = y$. For each y we choose one such x and denote it $g(y)$.[17] Thus $g\colon Y \to X$ is a function. We now compute as follows:

$$\begin{aligned}
(f \circ g)(y) &= f(g(y)) && \text{definition of } \circ \\
&= y && \text{definition of } g(y) \\
&= \mathrm{Id}_Y(y) && \text{definition of } \mathrm{Id}_Y
\end{aligned}$$

Since this holds for all y in Y, we deduce that $f \circ g = \mathrm{Id}_Y$, as desired.

Suppose that g is a right inverse for f. Suppose that y is any element of Y. Let $x = g(y)$. We compute as follows:

$$\begin{aligned}
y &= \mathrm{Id}_Y(y) && \text{definition of } \mathrm{Id}_Y \\
&= (f \circ g)(y) && g \text{ is a right inverse for } f \\
&= f(g(y)) && \text{definition of } \circ \\
&= f(x) && \text{definition of } x
\end{aligned}$$

Since this holds for all y in Y, we deduce that f is surjective, as desired.

Suppose that f is bijective. Thus f is injective and surjective. By the above, f has both a left inverse, say g, and a right inverse, say g'. Note

[17] When Y is infinite, picking the elements $g(y)$ requires the *axiom of choice*.

that f is a right inverse for g and a left inverse for g'. Again applying the above, we deduce that g is surjective and g' is injective. Suppose that y is any element of Y. We compute as follows:

$$\begin{aligned}
g(y) &= g(\mathrm{Id}_Y(y)) && \text{definition of } \mathrm{Id}_Y \\
&= g((f \circ g')(y)) && g' \text{ is a right inverse for } f \\
&= (g \circ (f \circ g'))(y) && \text{definition of } \circ \\
&= ((g \circ f) \circ g')(y) && \text{associativity of } \circ \\
&= (g \circ f)(g'(y)) && \text{definition of } \circ \\
&= \mathrm{Id}_X(g'(y)) && g \text{ is a left inverse for } f \\
&= g'(y) && \text{definition of } \mathrm{Id}_x
\end{aligned}$$

Thus $g = g'$ is the desired inverse for f. Note also that if g'' is another inverse for f then the above computation (with g'' replacing g') shows that $g = g''$. Thus f has only one inverse, as desired.

Suppose that f has an inverse, say g. So g is both a right and left inverse for f. Thus by the above f is both injective and surjective. Thus f is bijective, as desired. □

Solution 6.8. Suppose that f and g are both injective. Suppose that x and x' lie in X. Suppose that $(g \circ f)(x) = (g \circ f)(x')$. Since g is injective, we deduce that $f(x) = f(x')$. Since f is injective, we deduce that $x = x'$, as desired.

Suppose that f and g are both surjective. Suppose that z lies in Z. Since g is surjective, there is some y in Y so that $g(y) = z$. Since f is surjective, there is some x is X so that $f(x) = y$. Thus $(g \circ f)(x) = z$, as desired. □

Solution 7.4.

(1) We list the solutions to $x + y = 2$ over $\mathbb{N}$. They are:

$$(0,2), (1,1), (2,0)$$

These are all of the elements of P. Since $(0,0)$ is not in P, we have that P is not reflexive. Since the formula defining P is symmetric in x and y, we have that P is symmetric. (This can also be proved directly, by checking the ordered pairs.) Since $(0,2)$ and $(2,0)$ lie in P but $(0,0)$ does not, P is not transitive.

(2) We list the solutions to $x - y \leq 2$ over $\mathbb{N}$. They are, for $n, k \in \mathbb{N}$:

$$(n+2, n), (n+1, n), (n, n), (n, n+1), \ldots, (n, n+k), \ldots$$

Since (n, n) lies in Q for all n, we have that Q is reflexive. Since $(0,3)$ lies in Q but $(3,0)$ does not, Q is not symmetric. Since $(4,2)$ and $(2,0)$ lie in Q but $(4,0)$ does not, Q is not transitive.

(3) We list the solutions to $|x - y| \leq 2$ over $\mathbb{N}$. They are, for $n \in \mathbb{N}$:

$$(n+2, n), (n+1, n), (n, n), (n, n+1), (n, n+2)$$

Since (n, n) lies in R for all n, we have that R is reflexive. Since the formula defining R is symmetric in x and y, we have that R is symmetric. Since $(4, 2)$ and $(2, 0)$ lie in Q but $(4, 0)$ does not, R is not transitive. □

Solution 7.8. Recall that an equivalence relation E on X is a subset of X^2. Since $X = [\![4]\!]$ we may visualise X^2 as a four-by-four grid of squares. We visualise E by filling in some of the squares of the grid – blank squares do not belong to E. Finally, we give each filled square a colour, and insist that filled squares in the same row or column receive the same colour. (The reason for this will be explained when we discuss *equivalence classes*.) All equivalence relations on $[\![4]\!]$, represented in this fashion, are drawn in Figure A.1.

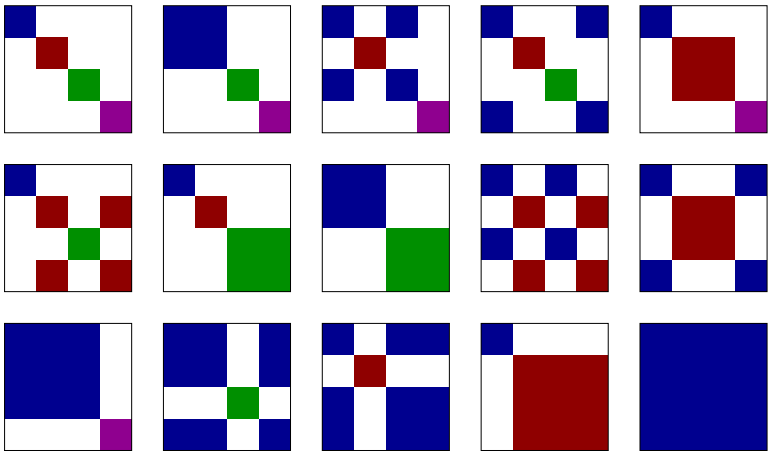


FIGURE A.1. Equivalence relations on $X = [\![4]\!]$. Drawing style inspired by Tilman Piesk [12] who presents the case of $X = [\![5]\!]$.

Each of the properties of an equivalence relation places a condition on the diagrams of Figure A.1.

- Since E is reflexive, the squares on the diagonal (running from the upper left to lower right in each diagram) are filled.
- Since E is symmetric, each diagrams has a reflection symmetry across the diagonal.
- Since E is transitive, there is a "Tetris" effect in the diagram. Suppose that the square (x, y), in row x and column y, is

filled. Then for every filled square in row y, say (y, z), the corresponding square, here (x, z), is filled in row x.

On the other hand, if a diagram follows these rules, then it encodes an equivalence relation. We now generate the diagrams as follows.

(1) The diagonal gives the identity relation.
(2) Given a diagram of an equivalence relation (say the identity) we can fill in a single additional square. We then use symmetry and transitivity to fill in the smallest number of squares needed to make the diagram that of an equivalence relation.

Starting with the identity relation the above procedure gives the next six diagrams. Starting with one of those, we obtain the next seven diagrams. Finally, starting with one of those we obtain the final diagram, of the universal relation.

Our list is complete because every equivalence is obtained from the identity relation by adding some number (perhaps zero) of squares. Our list is irredundant because all of the diagrams are distinct. □

Solution 8.3.

(1) We list the partitions of $X = [\![4]\!]$ by listing those with smaller parts before those with larger parts.

$$0|1|2|3 \quad 01|2|3 \quad 02|1|3 \quad 03|1|2 \quad 0|12|3$$

$$0|13|2 \quad 0|1|23 \quad 01|23 \quad 02|13 \quad 03|12$$

$$012|3 \quad 013|2 \quad 023|1 \quad 0|123 \quad 0123$$

If the largest part of a partition is a singleton, then all parts are singletons and we obtain $0|1|2|3$. To obtain more partitions from this, we combine any two; there are six ways to do this. After this we consider the partitions with two parts of size two; there are three of these. Finally, there are four partitions with a part of size three and one partition with a single part. Thus there are 15 partitions of $[\![4]\!]$.

(2) We count the partitions of $X = [\![n]\!]$ with exactly two parts, say P and Q. Note that $Q = X - P$ is determined by P. Thus it suffices to count the number of subsets of X which are non-empty and proper. There are $2^n - 2$ of these. However, these come in complementary pairs P and $X - P$. So dividing by two gives the desired lower bound. □

Solution 8.6. Each colour in each diagram in Exercise 7.8 corresponds to one equivalence class. As a consequence,

- the equivalence classes of the equivalence relations in the solution to Exercise 7.8 are exactly
- the parts of the partitions given in the solution to Exercise 8.3(1).

□

Solution 8.7. There are twelve equivalence classes – one for every month of the year. (Note that this assumes that in each month there is at least one person born in that month.)

As a further exercise: Suppose that D is the equivalence relation on the set of people where xDy if x and y are born on the same day of the same month. Count the number of equivalence classes for D. □

Solution 8.13. We recall that for $(p,q),(r,s) \in \mathbb{N}^2$ we have

$$(p,q)E(r,s) \quad \text{if and only if} \quad p+s=r+q$$

Since $p+q=p+q$ we deduce that E is reflexive. Since $p+s=r+q$ implies $r+q=p+s$ we deduce that E is symmetric.

Transitivity is more interesting. Suppose that $(p,q)E(r,s)$ and $(r,s)E(t,u)$. Thus $p+s=r+q$ and $r+u=t+s$. We add these equations to obtain

$$p+s+r+u=r+q+t+s$$

Addition in $\mathbb{N}$ is *cancellative* – that is $x+y=x+z$ if and only if $y=z$. Thus in the equation above we may cancel the terms (r and s) appearing on both sides. This gives $p+u=t+q$, as desired. □

Solution 10.3.

(1) Suppose that m, n, and ℓ are integers.

Since $m=1\times m$ we have that m divides m. Thus divides is reflexive.

Suppose that ℓ divides n which divides m. So there are integers k and h so that $n=k\times\ell$ and $m=h\times n$. Substituting, we have that $m=(h\times k)\times\ell$. Thus ℓ divides m. So divides is transitive.

Note that 1 divides 2, but 2 does not divide 1. Thus divides is not symmetric.

(2) Suppose that m is an integer. Then $m=m\times 1=(-m)\times(-1)$. Thus 1 and -1 divide m.

(3) Suppose that m is an integer. Then $0=0\times m$. Thus m divides zero.

(4) Suppose that m is an integer. Suppose that zero divides m. Thus there is some k so that $m=k\times 0$. Thus $m=0$. Thus zero divides only itself. □

Solution 10.7. Note that $a - a = 0 = 0 \times n$. Thus $a \equiv a \pmod{n}$. Thus the relation is reflexive.

Suppose that $a \equiv b \pmod{n}$. So n divides $a - b$. So there is an integer k so that $a - b = k \times n$. Thus $b - a = (-k) \times n$ and so n divides $b - a$. Thus $b \equiv a \pmod{n}$. Thus the relation is symmetric.

Suppose that $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$. Thus there are integers h and k so that

$$a - b = h \times n \quad \text{and} \quad b - c = k \times n$$

Adding these equations we have

$$(a - b) + (b - c) = h \times n + k \times n$$

and thus

$$a - c = (h + k) \times n$$

Thus $a \equiv c \pmod{n}$. Thus the relation is transitive and we are done. □

Solution 10.9.

- If $a \equiv b \pmod{0}$ then $a - b$ is a multiple of zero: that is, $a = b$. So we have $[a]_0 = \{a\}$ for any integer a.
- Suppose that a and b are integers. Thus $a = b + (a - b) \times 1$. Thus $a \equiv b \pmod{1}$, as desired.
- Suppose that $a \equiv b \pmod{n}$. So $a = b + kn$ for some integer k. Thus $a = b + (-k) \times (-n)$. We deduce that $a \equiv b \pmod{-n}$. □

Solution 10.10.

- Suppose that m is an integer. Note that $m - (m - n) = n = 1 \times n$. Thus $m - (m - n)$ is divisible by n. Thus $m \equiv m - n \pmod{n}$.

  Similarly, $m - (m + n) = -n = (-1) \times n$. Thus $m \equiv m + n \pmod{n}$, as desired.
- Suppose that r and r' lie in $[\![n]\!]$. If $r = r'$ then $r - r' = 0$. Since n divides 0 we have that $r \equiv r' \pmod{n}$, as desired.

  Suppose that $r \equiv r' \pmod{n}$. So n divides $r - r'$. So there is some integer k so that $r - r' = k \times n$. We have two cases: either $r \geq r'$ or $r' \geq r$.

  Suppose that $r \geq r'$. Then $r - r'$ is non-negative. Since n is positive, we deduce that k is non-negative. If $k = 0$ then $r = r'$ and we are done. If $k > 0$ then $r - r' = k \times n \geq n$. However, $r - r' \leq n - 1$, a contradiction.

  Suppose that $r' \geq r$. Thus $r - r'$ is non-positive. Since n is positive, we deduce that k is non-positive. If $k = 0$ then $r = r'$ and we are done. If $k < 0$ then $r - r' = k \times n \leq -n$. However, $r - r' \geq 1 - n$, a contradiction.

- Let

$$\mathcal{E} = \{[r]_n \mid r \in [\![n]\!]\}$$

By the above $\mathcal{E}$ has n elements. We claim that $\mathcal{E} = \mathbb{Z}/n\mathbb{Z}$: that is, these are all of the equivalence classes of the relation.

We first deal with the non-negative integers. Suppose that $m < n$. In this case m lies in $[\![n]\!]$ and we are done. Suppose now that $m \geq n$. Suppose also, for a contradiction, that the congruence class $[m]_n$ does not lie in $\mathcal{E}$. We may assume that m is the least such positive integer.[18] Thus $m - n$ is non-negative and less than m. By the above $[m]_n = [m-n]_n$. Thus $[m-n]_n$ does not lie in $\mathcal{E}$. This contradicts our assumption that m was minimal.

Now suppose that $m < 0$. Suppose for a contradiction that the congruence class $[m]_n$ does not lie in $\mathcal{E}$. We may assume that m is the greatest such negative integer. Note that $m + n > m$. By the above we have $[m+n]_n = [m]_n$. If $m + n \geq 0$ then we contradict the previous paragraph. If $m + n < 0$ then we contradict the assumed maximality of m. In either case we are done. □

Solution 10.16.

(1) Using the laws of exponents (for $\mathbb{N}$) and modular arithmetic we compute as follows:

$$\begin{aligned}
3^{20} &\equiv (3^2)^{10} \pmod 7 && \text{because } 20 = 2 \times 10 \\
&\equiv 9^{10} \pmod 7 && 3^2 = 9 \\
&\equiv 2^{10} \pmod 7 && 9 \equiv 2 \pmod 7 \\
&\equiv (2^5)^2 \pmod 7 && 10 = 5 \times 2 \\
&\equiv (32)^2 \pmod 7 && 2^5 = 32 \\
&\equiv 4^2 \pmod 7 && 32 \equiv 4 \pmod 7 \\
&\equiv 16 \pmod 7 && 4^2 = 16 \\
&\equiv 2 \pmod 7 && 16 \equiv 2 \pmod 7
\end{aligned}$$

[18] This requires the *well-ordering principle* proven later as Theorem 16.12.

(2) Using the laws of exponents and modular arithmetic we compute as follows:

$$\begin{aligned} 2^{111} &\equiv (2^3)^{37} \pmod 7 && \text{because } 111 = 3 \times 37 \\ &\equiv 8^{37} \pmod 7 && 2^3 = 8 \\ &\equiv 1^{37} \pmod 7 && 8 \equiv 1 \pmod 7 \\ &\equiv 1 \pmod 7 && 1^{37} = 1 \end{aligned}$$

(3) We note that

$$[2]_7 \qquad [4]_7 \qquad [8]_7 = [1]_7$$

are the equivalence classes of 2^1, 2^2, and 2^3, respectively. Thus

$$2^4 \equiv (2^3 \times 2) \equiv 1 \times 2 \equiv 2 \pmod 7$$

Thus $[2^k]_7$ depends only on the residue of k modulo 3. Said another way, we have

$$[2^k]_7 = \begin{cases} [1]_7, & \text{if } k \equiv 0 \pmod 3 \\ [2]_7, & \text{if } k \equiv 1 \pmod 3 \\ [4]_7, & \text{if } k \equiv 2 \pmod 3 \end{cases}$$

Since $3^{20} \equiv 0 \pmod 3$ we have $2^{3^{20}} \equiv 1 \pmod 7$. □

Solution 11.10.

- From the definition of $\wedge$ we deduce that we may replace $(P \wedge P)$ by P. Thus we have the equality of functions
$$((P \wedge P) \wedge (Q \wedge Q)) = (P \wedge Q)$$
The latter is equal to $(\neg((\neg P) \vee (\neg Q)))$.
- From the definition of $\to$ we deduce that we may replace $(P \to (\neg P))$ by $(\neg P)$. Thus we have the equality of functions
$$((P \to (\neg P)) \to Q) = ((\neg P) \to Q)$$
The latter is equal to $((\neg(\neg P)) \vee Q)$ which in turn equals $(P \vee Q)$.
- From the definition of $\leftrightarrow$ we deduce that we may replace $(P \leftrightarrow (\neg P))$ by F. Thus we have the equality of functions
$$((P \leftrightarrow (\neg P)) \leftrightarrow (Q \leftrightarrow (\neg Q))) = (\mathrm{F} \leftrightarrow \mathrm{F})$$
Thus this is the function (of arity two) that always equals T. □

Solution 11.13.

- The operator $(P \vee Q)$ gives T if either boolean is T. Thus $(P \vee \mathrm{T})$ gives T. On the other hand $(P \vee \mathrm{F})$ gives T if and only if $P = \mathrm{T}$.

- If all of P, Q, and R are F then both of the operators $((P \vee Q) \vee R)$ and $(P \vee (Q \vee R))$ give F. On the other hand, if any of P, Q, or R are T then both operators are also T.
- One of P or Q is T if and only if one of Q or P is T.
- We have that $(P \vee Q) = Q$ does *not* hold if and only if P = T and Q = F. But this is also the case for $(P \to Q)$.
- From the definition of $\vee$ we have $(\mathrm{T} \vee \mathrm{T})$ = T and $(\mathrm{F} \vee \mathrm{F})$ = F. □

Solution 12.2. We take the order on the strings of length three, form two copies, and prefix the strings in the first copy with 0 and those in the second with 1.

$$0000 \le 0001 \le 0010 \le 0011 \le 0100 \le 0101 \le 0110 \le 0111 \le$$
$$1000 \le 1001 \le 1010 \le 1011 \le 1100 \le 1101 \le 1110 \le 1111$$ □

Solution 12.6. We give the completed truth table.

$((P$	$\wedge$	$(P$	$\to$	$Q))\to$	$Q)$	
T	T	T	T	T	**T**	T
T	F	T	F	F	**T**	F
F	F	F	T	T	**T**	T
F	F	F	T	F	**T**	F

The final entries (in bold) are all T, thus the operator is a tautology. □

Solution 13.4.

- The given sentence is

$$(\forall p\,(\exists n\,((p > n) \wedge \textsc{Prime}(p))))$$

  which directly translates to

  for all p there exists an n so that $p > n$ and p is prime

  This does not hold. For consider $p = 4$. Then the sentence after the universal quantifier becomes

  there exists an n so that $4 > n$ and 4 is prime

  This does not hold, regardless of the n we consider, because 4 is not prime.
- The given sentence is

$$(\exists n\,(\forall p\,((p > n) \wedge \textsc{Prime}(p))))$$

  which directly translates to

  there exists an n such that for any p we have $p > n$ and p is prime

This does not hold. Suppose for a contradiction that such an n exists. Then inner clause holds for all p. So we may take $p = n$. We deduce that $n > n$; this is the desired contradiction.

- The given sentence is

$$(\exists p\,(\forall n\,((p > n) \wedge \textsc{Prime}(p))))$$

which directly translates to

there exists a p so that, for any n, we have $p > n$ and p is prime

This does not hold. Again, for a contradiction we suppose that such a p existed. Since the inner clause holds for all n, we may take $n = p$. We deduce that $p > p$; this is the desired contradiction. □

Solution 13.8.

- Suppose that

$$(\forall x\,(R(x) \to S(x)))$$

holds. Suppose further that $(\forall x\ R(x))$ holds. We must show that $(\forall x\ S(x))$ holds. So fix any x' in X. Thus $R(x')$ and $(R(x') \to S(x'))$ both hold. From modus ponens we deduce that $S(x')$ holds. As x' was arbitrary, we deduce that $(\forall x\ S(x))$ holds, as desired.

- We now show, by means of an example, that

$$((\forall x\ R(x)) \to (\forall x\ S(x)))$$

need not imply

$$(\forall x\,(R(x) \to S(x)))$$

Suppose that X is the set $\{0, 1\}$. Suppose that $R(x)$ is the property $(x = 0)$. Suppose that $S(x)$ is the property $(x = 1)$. Then the sentence

$$((\forall x\ R(x)) \to (\forall x\ S(x)))$$

translates to

> If every number in X equals zero, then every number in X equals one.

The hypothesis does not hold and neither does the conclusion. Therefore the implication holds (Definition 11.6). On the other hand, the sentence

$$(\forall x\,(R(x) \to S(x)))$$

translates to

> For any x in X, if $x = 0$ then $x = 1$.

Since $0 \neq 1$, this does not hold. □

Solution 14.4. Suppose that n is a positive integer. Let $\mathcal{P}$ be the set of primes less than or equal to n which are congruent to 3 modulo four. By Lemma 17.12 the set $\mathcal{P}$ is finite. So we may write $\mathcal{P}$ as a list $(p_i)_{i=0}^{k-1}$, in order of size. We define M to be the product of these primes. We now define N to be

$$N = \begin{cases} M+2, & \text{if } M \equiv 1 \pmod 4 \\ M+4, & \text{if } M \equiv 3 \pmod 4 \end{cases}$$

Thus N is congruent to 3 modulo four. Thus N is odd and so not divisible by two. Also, N is not divisible by any prime p in $\mathcal{P}$.

By Proposition 16.13, the integer N is divisible by some prime. Thus there is some prime q dividing N. So q does not lie in the set P. Also, since q is not equal to two, we deduce that q is odd.

Let (q_i) be the list of primes dividing N, ordered by size.[19] Note that none of the q_i lie in $\mathcal{P}$.

Suppose, for a contradiction, that all of the q_i are congruent to 1 modulo four. Then the same holds of their product, N. This contradicts the definition of N. Thus there is at least one prime q in the list (q_i) which is congruent to 3 modulo four. This completes the proof. □

Solution 15.4. Suppose that n is an odd integer. Thus there is an integer k so that $n = 2k+1$. We find that $n^2 = 4k^2+4k+1 = 2(2k^2+2k)+1$. Thus n^2 is odd, as desired. □

Solution 15.7. Suppose that $X \subset Y$. Thus, by Definition 1.10 we have that, every x in X also lies in Y. Recall from Definition 5.10 that we have

$$X \cap Y = \{z \mid z \in X \text{ and } z \in Y\}$$

Note that $X \cap Y \subset X$, by Lemma 5.13.

We now prove that $X \subset X \cap Y$. Suppose that x is any element of X. Since $X \subset Y$ we have that x lies in Y. Thus x lies in X and x lies in Y. So x lies in $X \cap Y$. Thus $X \subset X \cap Y$.

Since $X \cap Y \subset X$ and $X \subset X \cap Y$ Lemma 1.16 tells us that $X \cap Y = X$, as desired. □

Solution 15.8. Since f and g are bijections, they are both injections. By Lemma 6.7 we have that $g \circ f$ is an injection. Similarly, since f and g are bijections, they are both surjections. By Lemma 6.7 we have that $g \circ f$ is a surjection. Since $g \circ f$ is both an injection and a surjection, it is a bijection, as desired. □

Solution 15.12. Suppose that p is a sum of three squares: say $p = x^2 + y^2 + z^2$. We now break into cases depending on the residues of

[19] This step requires the *fundamental theorem of arithmetic* – see Theorem 17.3.

x, y, and z modulo four. For example, suppose that $x = 4w + b$ where b lies in $[\![4]\!]$. Thus $x^2 = 16w^2 + 8wb + b^2$. Taking this modulo eight we get either 0, 1, or 4. The same holds for y^2 and z^2. The sum of all three modulo eight is one of the following:

$$0 \equiv 0+0+0 \quad 1 \equiv 1+0+0 \quad 2 \equiv 1+1+0 \quad 3 \equiv 1+1+1$$
$$4 \equiv 4+0+0 \quad 5 \equiv 4+1+0 \quad 6 \equiv 4+1+1$$

In no case do we obtain a residue of seven modulo eight. □

Solution 15.15. Suppose that, for a contradiction, there is a rational number $r = p/q$ with $r^2 = 3$. Changing the signs of *both* p and q (if needed) we may assume that q is positive. Since $r^2 = (-r)^2$ we may further assume that p is positive.

We now descend. By hypothesis, $p^2 = 3q^2$. So p^2 is a multiple of three. So the same holds of p. Thus we may write $p = 3u$. Thus $q^2 = 3u^2$. We deduce that q is a multiple of three. Thus we may write $q = 3v$. Thus $r = u/v$ where $u < p$ and $v < q$. This completes the descent.

We have shown that if $r = p/q$ then there is a smaller pair of positive numbers u and v having the same ratio. Repeating this indefinitely contradicts the well-ordering principle (Theorem 16.12). □

Solution 15.16. Suppose that, for a contradiction, there is a rational number $r = p/q$ with $r^3 = 2$. Changing the signs of *both* p and q (if needed) we may assume that q is positive. Note that r^3 has the same sign as r itself. Thus $p > 0$.

We now descend. By hypothesis, $p^3 = 2q^3$. So p^3, and thus p, is even. Thus we may write $p = 2u$. So $p^3 = 8u^3 = 2q^3$. Thus $q^3 = 4u^2$. We deduce that q is even. Thus we may write $q = 2v$. Thus $q^3 = 8v^3 = 4u^2$. So, finally, $u^3 = 2v^3$. Thus $r = u/v$ where $u < p$ and $v < q$. This completes the descent.

We have shown that if $r = p/q$ then there is a smaller pair of positive numbers u and v having the same ratio. Repeating this indefinitely contradicts the well-ordering principle (Theorem 16.12). □

Solution 15.19. The smallest number with four distinct prime factors is $2\times3\times5\times7$ which equals 210. As this is less than 1000 we are done. □

Solution 15.20. Suppose that $r = \epsilon p/q$ where $\epsilon = \pm1$, where $p \geq 0$, where $q > 0$, where p and q share no common factors, and where $\epsilon = 1$ if $p = 0$. We define $f\colon \mathbb{Q} \to \mathbb{N}$ by $f(r) = 2^{\epsilon+1}3^p5^q$.

We now show that f is an injection. Suppose that $r = \epsilon p/q$ and $s = \delta u/v$ is a pair of rational numbers. Suppose that $n = f(r) = f(s)$. Since $q > 0$ the number n is greater than or equal to five. By Theorem 17.3

the number n has a unique prime factorisation. Thus we may recover ϵ, p, and q from n. Thus $\epsilon = \delta$, $p = u$, and $q = v$. Thus $r = s$, as desired. □

Solution 16.4. Suppose that n is a natural number. Let $P(n)$ be the sentence "n is congruent to zero, one, or two, modulo three".

We now deal with the base case, where $n = 0$. Since $0 = 3 \times 0$, the number zero congruent to zero modulo three. Thus $P(0)$ holds.

We now deal with the induction step. Suppose that k is a natural number. Suppose that $P(k)$ holds. Thus there are natural numbers ℓ and i so that $k = 3\ell + i$ and $i \leq 2$. Adding one we find that $k+1 = 3\ell+i+1$. If $i = 0$ or 1 we are done. If $i = 2$ then $k + 1 = 3\ell + 3 = 3(\ell + 1)$ and again we are done. Thus $P(k)$ implies $P(k + 1)$, establishing the induction step.

Thus we are done by induction. □

Solution 16.9. With notation as in Theorem 16.6, we take $F(k, \ell) = k + \ell$ and $m = 0$. This gives a function $f\colon \mathbb{N} \to \mathbb{N}$ where

- $f(0) = 0$ and
- $f(n + 1) = f(n) + n + 1$.

We compute the first few values of f.

$$
\begin{aligned}
f(0) &= 0 \\
f(1) &= f(0) + 0 + 1 = 1 = 1 \\
f(2) &= f(1) + 1 + 1 = 1 + 2 = 3 \\
f(3) &= f(2) + 2 + 1 = 1 + 2 + 3 = 6 \\
f(4) &= f(3) + 3 + 1 = 1 + 2 + 3 + 4 = 10 \\
f(5) &= f(4) + 4 + 1 = 1 + 2 + 3 + 4 + 5 = 15
\end{aligned}
$$

This leads us to the following.

Lemma A.2. *The sum of the first n natural numbers equals $\frac{(n+1)n}{2}$.*

Since $f(n)$ is defined by recursion, we give a proof by induction.

In the base case we have $f(0) = 0 = \frac{(0+1)0}{2}$

For the induction step, we suppose that $f(k) = \frac{(k+1)k}{2}$. We now compute as follows.

$$\begin{aligned} f(k+1) &= k+1+f(k) && \text{definition of } f \\ &= k+1+\frac{(k+1)k}{2} && \text{induction hypothesis} \\ &= \frac{2(k+1)}{2}+\frac{k(k+1)}{2} && \text{common denominator} \\ &= \frac{(k+2)(k+1)}{2} && \text{add and factor} \end{aligned}$$

Since the base case and the induction step hold we are done by induction. □

Solution 17.6. Let $P(k)$ be the sentence "$f(3k)$ is even while $f(3k+1)$ and $f(3k+2)$ are odd". From the table of values, given after Definition 17.5, we see that $f(0) = 0$ is even while $f(1) = f(2) = 1$ are odd. Thus $P(0)$ holds and the base case is obtained.

Suppose that $P(k)$ holds. We have the following.

$$\begin{aligned} f(3k+3) &= f(3k+2)+f(3k+1) \equiv 1+1 \equiv 0 \pmod 2 \\ f(3k+4) &= f(3k+3)+f(3k+2) \equiv 0+1 \equiv 1 \pmod 2 \\ f(3k+5) &= f(3k+4)+f(3k+3) \equiv 1+0 \equiv 1 \pmod 2 \end{aligned}$$

Thus $P(k+1)$ holds, and the induction step is obtained. □

Solution 17.7.

(1) Suppose that $S = b_0b_1b_2b_3\ldots b_{n-1}$ is a binary string of length n. Suppose also that, for all $i > 0$ we have that b_{i+1} is different from b_i. Then b_0 determines b_1, which determines b_2, and so on. Thus b_0 determines S. Thus there are exactly two possibilities for S: either S begins $01010\ldots$ or begins $10101\ldots$.

(2) Suppose that F_n is the set of binary strings of length n where no adjacent bits are both zero. Suppose that f_n is the number of such strings. We now gather a bit of data.

n	f_n	F_n
0	1	$\epsilon_{\mathcal{B}}$
1	2	0, 1
2	3	01, 10, 11
3	5	010, 011, 101, 110, 111
4	8	0101, 0110, 0111, 1010, 1011, 1101, 1110, 1111
5	13	01010, 01011, 01101, 01110, 01111, 10101, 10110, 10111, 11010, 11011, 11101, 11110, 11111

Equipped with this, we now claim that $f_{n+2} = f_{n+1} + f_n$ for all $n \geq 0$. [That is, the numbers f_n are the Fibonacci numbers, with indexing off by one.]

We prove this as follows. Suppose that F_n^0 is the set of those strings in F_n ending with a zero. Suppose that F_n^1 is the set of those strings in F_n ending with a one. For example, $F_4^0 = \{0110, 1010, 1110\}$ and $F_4^1 = \{0101, 0111, 1011, 1101, 1111\}$ contain three and five strings, respectively. Let $f_n^0 = |F_n^0|$ and $f_n^1 = |F_n^1|$. Note that $f_n = f_n^0 + f_n^1$.

Thus, to prove that $f_{n+2} = f_{n+1} + f_n$ it suffices to prove that $f_{n+2}^0 = f_n$ and $f_{n+2}^1 = f_{n+1}$. And to prove that it suffices to find bijections $F_{n+2}^0 \to F_n$ and $F_{n+2}^1 \to F_{n+1}$.

Suppose that w is a binary string in F_{n+2}^1. Thus the final bit of w is a 1. We delete this last bit to obtain the string w'. Note that w' lies in F_{n+1}. Furthermore, every element of F_{n+1} is obtained, exactly once, in this way.

Suppose that w is a binary string in F_{n+2}^0. So the last bit of w is a 0. Since adjacent zeros are not allowed, the second to last bit of w is a 1. We delete these two final bits to obtain w''. Note that w'' lies in F_n. Furthermore, every element of F_n is obtained, exactly once, in this way. □

Solution 17.11.

(1) Suppose that n is a natural number. Suppose that $g\colon [\![n+1]\!] \to [\![n]\!]$ is an injection. By Lemma 6.5 we have a left inverse $f\colon [\![n]\!] \to [\![n+1]\!]$ which is a surjection. This contradicts Theorem 17.9.

(2) Suppose that m and n are natural numbers, with $m > n$. Suppose that $g\colon [\![m]\!] \to [\![n]\!]$ is an injection. Note that $[\![n+1]\!]$ is a subset of

$[\![m]\!]$. Thus the restriction $h = g|[\![n+1]\!]$ is an injection of $[\![n+1]\!]$ to $[\![n]\!]$.[20] This contradicts (1).

(3) Suppose that n is a natural number. Suppose that $g\colon \mathbb{N} \to [\![n]\!]$ is a bijection. Thus g is an injection. Thus $h = g|[\![n+1]\!]$ is also an injection. This contradicts (1).

(4) Suppose that we are given bijections from X to $[\![m]\!]$ and to $[\![n]\!]$. By inverting and composing we obtain a bijection from $[\![m]\!]$ to $[\![n]\!]$. In particular we have injections from $[\![m]\!]$ to $[\![n]\!]$ and from $[\![m]\!]$ to $[\![n]\!]$. Thus, by (2) we have $m \leq n$ and $n \leq m$. Thus $m = n$.

(5) Suppose that X is a finite set. Since cardinality is unique, we may assume that $X = [\![n]\!]$ for some $n \in \mathbb{N}$. Suppose that $f\colon [\![n]\!] \to [\![n]\!]$ is a function.

Suppose that f is not a surjection. Thus $[\![n]\!]$ is non-empty (and so $n > 0$). Post-composing with a permutation we may assume that $n-1$ does not lie in the image of f. Restricting the codomain to $[\![n-1]\!]$, we obtain a function $f'\colon [\![n]\!] \to [\![n-1]\!]$. By (1) we have f' is not an injection, and thus neither is f.

For the opposite direction, suppose that f is not an injection. Thus $[\![n]\!]$ is non-empty (and so $n > 0$). Pre-composing with a permutation we may assume that $f(\ell) = f(n-1)$ for some $\ell < n-1$. Restricting the domain to $[\![n-1]\!]$, we obtain a function $f'\colon [\![n-1]\!] \to [\![n]\!]$. By Theorem 17.9 we have that f' is not a surjection, and thus neither is f. □

Solution 19.5.

(1) For all $g \in G$ we have $g \cdot e_G = e_G \cdot g = g$. So, taking $g = e_G$ we deduce that $e_G \cdot e_G = e_G$. Thus e_G is its own inverse, as desired.

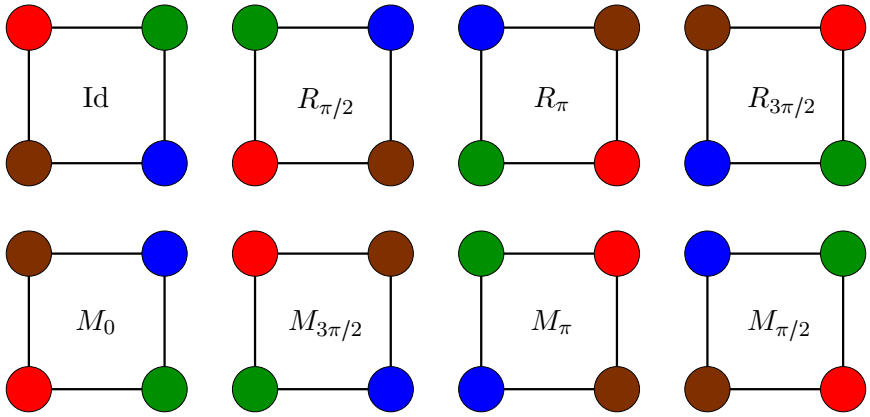
(2) By definition of g^{-1} we have $g \cdot g^{-1} = g^{-1} \cdot g = e_G$. Thus g is the inverse of g^{-1}. From the uniqueness of inverses (Lemma 19.3) we deduce that $g = (g^{-1})^{-1}$. □

Solution 19.15. We first show that function composition gives a function from $D_{2n} \times D_{2n}$ to D_{2n} Suppose that g and h lie in D_{2n}. Lemmas 19.11 and 19.13 allow us to rewrite $g \circ h$ as a single rotation or reflection. The lemmas also imply that $R_{2\pi+\theta} = R_\theta$ and $M_{2\pi+\theta} = M_\theta$. Thus the angle in the subscript can be assumed to lie in $[0, 2\pi)$. Finally, integer linear combinations of fractions with denominator n again have denominator n. Thus $g \circ h$ lies in D_{2n}.

[20] Suppose that $g\colon X \to Z$ is a function. Suppose that Z is a subset of X. Then we define $g|Z$, the *restriction of g to Z*, to be the function from Z to Y where $(g|Z)(z) = g(z)$.

Next, we note that $R_0 = \mathrm{Id}_{\mathbb{R}^2}$ gives the identity element of D_{2n}. Lemmas 19.11 and 19.13 imply that $R_\theta^{-1} = R_{-\theta}$ and $M_\theta^{-1} = M_\theta$. Lastly, function composition is associative by Lemma 6.3. □

Solution 19.16. Here are the asked-for decorated squares.



We label the centre of each square S with the group element in D_{2n} (with $n = 4$) that transforms the upper left square into S. □

Solution 20.6. We take $g = R_\eta$ and $h = M_0$. Then $g \circ h = M_\eta$ and $h \circ g = M_{-\eta}$. We take $\eta = 2\pi/n$. Thus M_η is the reflection in the line with angle π/n. As $n > 2$ this line has positive slope. On the other hand, $M_{-\eta}$ is the reflection in the line with angle $-\pi/n$; this line has negative slope. Thus M_η is not equal to $M_{-\eta}$. □

Solution 20.9. To simplify the notation below we write k for $[k]_6$. We list the subgroups of $(\mathbb{Z}/6\mathbb{Z}, +_6)$ by size.

$$\{0\} \qquad \{0,3\} \qquad \{0,2,4\} \qquad \mathbb{Z}/6\mathbb{Z}$$ □

Solution 20.10. Let $H \subset D_{2n}$ be the set of rotations. Note that $\mathrm{Id}_{\mathbb{R}^2} = R_0$ is a rotation, so lies in H. Suppose that g and h elements of H. Thus they are rotations in D_{2n}. By Lemma 19.11 we have that g^{-1} is again a rotation, as is $g \cdot h$. Thus H is a subgroup of D_{2n}.

Since $n > 1$ the rotation $R_{2\pi/n}$ lies in H and is not the identity. Thus H is a non-trivial subgroup. Since M_0 does not lie in H we have that H is a proper subgroup. Finally we note that $R_{2\pi k/n}$ equals $R_{2\pi\ell/n}$ if and only if $k \equiv \ell \pmod{n}$. Thus, by Exercise 10.10, the subgroup H has order n. □

Solution 20.18. Let $P_g \subset G$ be the set of powers of g. Note that g lies in $\langle g \rangle$, as does g^{-1}. By induction all powers of g lie in $\langle g \rangle$. That is, $P_g \subset \langle g \rangle$.

On the other hand, P_g is a subgroup of G. Thus $P_g \subset \langle g \rangle$. Thus $P_g = \langle g \rangle$.

Finally, suppose that g^m and g^n lie in $\langle g \rangle$. By Lemma 20.17 we have $g^m \cdot g^n = g^{m+n}$. Since $m+n = n+m$ in $\mathbb{Z}$, we deduce that $g^m \cdot g^n = g^n \cdot g^m$ as desired. □

Solution 20.19. We define a function $\text{pow}\colon \mathbb{Z} \to \langle g \rangle$ by $\text{pow}(n) = g^n$. Note that pow is surjective by Exercise 20.18. We must show that $\text{pow}|[\![k]\!]$ is a bijection. Note that $\langle g \rangle$ and $[\![k]\!]$ have the same cardinality. Thus by Exercise 17.11(5) it suffices to show that $\text{pow}|[\![k]\!]$ is injective.

Suppose, for a contradiction, that n and m lie in $[\![k]\!]$, have $n < m$, and have $\text{pow}(n) = \text{pow}(m)$. Let $\ell = m - n$. Thus $0 < \ell < k$. Since $g^n = g^m$ we may multiply both sides by g^{-n}, apply Lemma 20.17, and deduce $e_G = g^\ell$.

Suppose that N is any integer. By Exercise 10.10 there are integers q and r with $0 \le r < \ell$ and $N = q\ell + r$.

$$\begin{aligned}
g^N &= g^{q\ell+r} && \text{reducing } N \text{ modulo } \ell \\
&= g^{q\ell} \cdot g^r && \text{Lemma 20.17} \\
&= g^{\ell q} \cdot g^r && \text{arithmetic} \\
&= (g^\ell)^q \cdot g^r && \text{Lemma 20.17} \\
&= e_G^q \cdot g^r && \text{definition of } \ell \\
&= g^r && \text{Exercise 20.18}
\end{aligned}$$

Thus $\text{pow}(N) = \text{pow}(r)$. We deduce that the image $\text{pow}(\mathbb{Z})$ equals the image $\text{pow}([\![\ell]\!])$. This has cardinality at most ℓ. We deduce that $[\![\ell]\!]$ surjects $[\![k]\!]$, contradicting Theorem 17.9. This completes the proof that $\text{pow}|[\![k]\!]$ is bijective.

As a consequence, if $0 < \ell < k$ then g^ℓ is not equal to the identity. On the other hand, g^k is a power of g, so lies in $\langle g \rangle$ by Exercise 20.18. If $g^k = g^i$ for some $i \neq 0$ then we may define $\ell = k - i$ and proceed as above. We deduce that $g^k = e_G$, and k is the smallest such (positive) power. □

Solution 20.20.

(1) By induction, the k^{th} power of $R_{2\pi/n}$ is the element $R_{2\pi k/n}$. Thus the powers of $R_{2\pi/n}$ give the rotation subgroup and we are done by Exercise 20.18

(2) Since S is a subset of T, we deduce that $\langle T \rangle$ contains the rotation subgroup. Applying Lemma 19.13, we now obtain all reflections in D_{2n} as a product of M_0 and some rotation. □

Solution 21.3. We verify each property of Definition 19.2 in turn.

(1) Id_X (Definition 2.9) is the identity element.
(2) Permutations are bijections (Definition 3.6) so have inverses (Lemma 6.5).
(3) Function composition is associative (Lemma 6.3).

Thus $(\textsc{Sym}(X), \circ)$ is a group. □

Solution 21.6. We note that σ is not the identity, as $\sigma(0) = 1$. On the other hand, for all $k \in [\![5]\!]$ we have $\sigma^2(k) = k$. So $\sigma^2 = \mathrm{Id}_{[\![5]\!]}$. □

Solution 21.9. For each $i \in [\![6]\!]$ we write down $\sigma(\tau(i))$ and $\tau(\sigma(i))$ to obtain the one-line notations

$$\sigma\tau = \begin{pmatrix} 2 & 1 & 4 & 0 & 3 & 5 \end{pmatrix}, \quad \tau\sigma = \begin{pmatrix} 0 & 3 & 1 & 5 & 4 & 2 \end{pmatrix}$$

Since $\sigma\tau \neq \tau\sigma$ we deduce that the $\textsc{Sym}(6)$ is not commutative. □

Solution 21.10. The one-line notations for σ, σ^2, and σ^3 are

$$\begin{pmatrix} 1 & 2 & 0 & 4 & 5 & 3 \end{pmatrix} \qquad \begin{pmatrix} 2 & 0 & 1 & 5 & 3 & 4 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

Thus $\sigma^3 = \mathrm{Id}_{[\![6]\!]}$. We deduce that $\sigma^4 = \sigma$, $\sigma^5 = \sigma^2$, and generally, by induction, $\sigma^{3k+\ell} = \sigma^\ell$. Thus the one-line notations for all powers of σ (including negative ones) are given above. □

Solution 22.7. Suppose that k lies in $[\![n]\!]$. If k is not in $R_\sigma \cup R_\tau$ then we have

$$\sigma(\tau(k)) = \sigma(k) = k = \tau(k) = \tau(\sigma(k))$$

Otherwise (as the other case is similar) suppose that k lies in R_σ but not in R_τ. Take $\ell = \sigma(k)$. So ℓ lies in R_σ and thus not in R_τ. So we have

$$\sigma(\tau(k)) = \sigma(k) = \ell = \tau(\ell) = \tau(\sigma(k))$$

and the proof is complete. □

Solution 22.8. We have $\sigma = (0\ 1\ 2\ 3) = (0\ 3)(0\ 2)(0\ 1)$. □

Solution 22.9. Suppose that a is the smallest number in the support of the k–cycle σ. For all $i \in \mathbb{N}$ we define the numbers $a_i = \sigma^i(a)$ and the permutations $\tau_i = (a_0\ a_i)$. So $\tau_i = (a_0\ a_i)$ is a transposition for i (strictly) between 0 and k.

Informally, we deduce that

$$\sigma = (a_0\ a_{k-1})(a_0\ a_{k-2})\cdots(a_0\ a_3)(a_0\ a_2)(a_0\ a_1)$$

To make this more rigorous we define permutations η_j (for $j \geq 1$) recursively by

$$\eta_1 = \mathrm{Id}_{[\![n]\!]} \qquad \text{and} \qquad \eta_{j+1} = \tau_j \circ \eta_j$$

Induction on j gives us the following: η_j is a product of $j-1$ transpositions, η_j is a j–cycle, and $\eta_j^i(a_0) = a_i$ (for $i < j$). Thus $\eta_k = \sigma$ and we are done. □

Solution 23.6. Suppose that G lies in $\mathcal{G}$. Then Id_G is an isomorphism. So $G \cong G$ and the relation is reflexive.

Suppose that G and H lie in $\mathcal{G}$. Suppose that $\phi\colon G \to H$ is an isomorphism. Since ϕ is a bijection Lemma 6.5 provides an inverse, as a function on sets. Let $\psi\colon H \to G$ be this inverse. Note that ψ is also a bijection. Thus, to prove that ψ is an isomorphism, we must show it is a homomorphism.

To do this, suppose that h and h' lie in H. Suppose that $g = \psi(h)$ and $g' = \psi(h')$. Thus $\phi(g) = h$ and $\phi(g') = h'$. We now compute as follows.

$$\begin{aligned}
\psi(h \cdot h') &= \psi(\phi(g) \cdot \phi(g')) && \text{definition of } g \text{ and } g' \\
&= \psi(\phi(g \cdot g')) && \phi \text{ is a homomorphism} \\
&= \mathrm{Id}_G(g \cdot g') && \psi \text{ is inverse of } \phi \\
&= g \cdot g' && \text{identity} \\
&= \psi(h) \cdot \psi(h') && \text{definition of } g \text{ and } g'
\end{aligned}$$

Thus ψ is a homomorphism, and so an isomorphism, as desired. So $H \cong G$ and the relation is symmetric.

Suppose that G, H, and K lie in $\mathcal{G}$. Suppose that $\phi\colon G \to H$ and $\psi\colon H \to K$ are isomorphisms. Then $\psi \circ \phi$ is a bijection (Lemma 6.7) and a homomorphism (Lemma 23.4). Thus $\psi \circ \phi$ is an isomorphism. So $G \cong K$ and the relation is transitive. □

Solution 23.9. This is (part of) Corollary 10.15. □

Solution 23.10. This follows from (part of) Lemma 20.17 and from Exercise 20.18. □

Solution 23.12.

- Suppose that ℓ lies in the congruence class $[k]_n$. So there are integers a and b with $k = \ell + a \cdot n$. We compute as follows.

$$\begin{aligned}
pk &= p(\ell + an) \\
&= p\ell + a(pn) \\
&= p\ell + am \\
&\equiv p\ell \pmod{m}
\end{aligned}$$

  Thus ϕ is well-defined.

Now, for any k and ℓ we have $[p(k+\ell)]_m = [pk]_m + [p\ell]_m$. Thus ϕ is a homomorphism.

- Suppose that ℓ lies in $[k]_m$. So there are integers a and b with $k = \ell + a \cdot m$. Thus $k = \ell + a(pn) = k = \ell + (ap)n$. So $[k]_n = [\ell]_n$ and ψ is well-defined.

  Now, for any k and ℓ we have $[k+\ell]_n = [k]_n + [\ell]_n$. Thus ψ is a homomorphism.

□

Solution 23.16. From the definition we have $\zeta_{pn}^{pk} = \exp(2\pi ipk/pn) = \exp(2\pi ik/n) = \zeta_n^k$. □

Solution 23.17. We sketch the solution. Since $\zeta_n^k \cdot \zeta_n^\ell = \zeta_n^{k+\ell}$ the set U_n is closed under multiplication. Since $\zeta_n^0 = 1$ it is the identity in U_n. Also, ζ_n^{-k} is the inverse of ζ_n^k. Finally, multiplication in $\mathbb{C}$ is associative. □

Solution 23.19. Suppose that m, n, and p are positive integers. Suppose that $m = p \cdot n$. Recalling Exercise 23.16 we may rephrase Exercise 23.12 as follows.

- The function $\Phi: U_n \to U_m$ defined by $\Phi(g) = g$ is a homomorphism.
- The function $\Psi: U_m \to U_n$ defined by $\Psi(g) = g^p$ is a homomorphism. □

Solution 23.21. Consider the three points

$$z_0 = (1, 0) \qquad z_1 = (-1/2, \sqrt{3}/2) \qquad z_2 = (-1/2, -\sqrt{3}/2)$$

lying in $\mathbb{R}^2$. Suppose that g lies in D_6. Thus g permutes the points $\{z_i\}_{i=0}^2$. Thus induces a permutation of their subscripts: that is, an element $\phi(g)$ of $\textsc{Sym}(3)$. This gives the desired function ϕ. Note that ϕ sends the identity to the identity, sends the rotations to the three-cycles, and sends the reflections to the transpositions. Thus ϕ is a bijection.

We may write out the multiplication tables of $\textsc{Sym}(3)$ and D_6 (say) to show that ϕ is a homomorphism. □

Solution 23.22. We only sketch the proof. Throughout we assume that $n > 3$. The symmetric group has exactly one subgroup of *index two*: that is, whose order is exactly half the order of the symmetric group. This subgroup is the *alternating group*, and it is not commutative (and so not cyclic).

On the other hand, in a dihedral group we have the rotation subgroup. This is index two and cyclic. □

Solution 24.2. Note that $\mathcal{D}(n)$ is a subset of $\mathcal{D}(n+1)$. The cardinality of $\mathcal{D}(n+1) - \mathcal{D}(n)$ is n. The result now follows by induction. □

Solution 24.4. Suppose that τ is the transposition $(i\ j)$, where $i < j$. That is, $\tau(i) = j$ and $\tau(j) = i$ while all other elements of $[\![n]\!]$ are fixed points of τ. Suppose that k and ℓ lie in $[\![n]\!]$. If neither of k nor ℓ is equal to either of i or j then τ does not have an inversion at $\{k, \ell\}$.

Suppose now that $k < i$ or $k > j$. Then τ does not have an inversion at $\{k, i\}$ or at $\{j, k\}$.

Suppose now that $i < k < j$. In this case τ has inversions both at $\{i, k\}$ and at $\{k, j\}$. So k contributes two inversions. There are $j - i - 1$ such k.

Also, τ has an inversion at $\{i, j\}$. Thus $\text{INV}(\tau) = 2(j - i - 1) + 1$. This is odd, as desired. □

Solution 24.5. Suppose that $a, b \in [\![n]\!]$ have $a \neq b$. Let $a' = \sigma(a)$ and $b' = \sigma(b)$. Thus $a = \sigma^{-1}(a')$ and $b = \sigma^{-1}(b')$. We deduce that σ has an inversion at $\{a, b\}$ if and only if σ^{-1} has an inversion at $\{a', b'\}$. Since σ induces a bijection of $\mathcal{D}(n)$ this completes the proof. □

Solution 24.6.

(1) The only permutation in $\text{SYM}(n)$ with the desired property is the identity.

We prove this by induction. Suppose that $\sigma \in \text{SYM}(k+1)$ is a permutation without inversions. We deduce that $\sigma(k) = k$. So $\sigma|[\![k]\!]$ is again a permutation, and again has no inversions. By the induction hypothesis, $\sigma|[\![k]\!]$ is the identity in $\text{SYM}(k)$. Thus σ is the identity in $\text{SYM}(k+1)$ and we are done.

(2) Define $\Delta_n \in \text{SYM}(n)$ by $\Delta(i) = n - 1 - i$. We call Δ_n the *longest* element of $\text{SYM}(n)$. The only permutation in $\text{SYM}(n)$ with the desired property is Δ_n

We prove this by induction. Suppose that σ is any element of $\text{SYM}(k+1)$ with $(k+1)k/2$ inversions. We deduce that $\sigma(k) = 0$. We now define $\sigma' \in \text{SYM}(k)$ by $\sigma'(\ell) = \sigma(\ell) - 1$. By the induction hypothesis $\sigma' = \Delta_k$. Thus $\sigma = \Delta_{k+1}$. □

Solution 25.9.

- Since K is a subgroup, e_G lies in K. Thus $\phi(e_G) = e_H$ lies in $\phi(K)$.

  Suppose that h lies in $\phi(K)$. So there is some $g \in K$ so that $\phi(g) = h$. Thus g^{-1} lies in K. Thus

  $$\phi(g^{-1}) = \phi(g)^{-1} = h^{-1}$$

  lies in $\phi(K)$.

Suppose that h and h' lie in $\phi(K)$. So there are some $g, g' \in K$ so that $\phi(g) = h$ and $\phi(g') = h'$ Then $g \cdot g'$ lies in K. Thus

$$\phi(g \cdot g') = \phi(g) \cdot \phi(g') = h \cdot h'$$

lies in $\phi(K)$.

- Since K is a subgroup, e_H lies in K. Since $\phi(e_G) = e_H$ we deduce that e_G lies in $\phi^{-1}(K)$.

  Suppose that g lies in $\phi^{-1}(K)$. Let $h = \phi(g)$. This lies in K. Thus h^{-1} also lies in K. Since $\phi(g^{-1}) = h^{-1}$ we deduce that g^{-1} lies in $\phi^{-1}(K)$.

  Suppose that g and g' lie in $\phi^{-1}(K)$. Let $h = \phi(g)$ and $h' = \phi(g')$. These lie in K. Thus $h \cdot h'$ also lies in K. Since $\phi(g \cdot g') = h \cdot h'$ we deduce that $g \cdot g'$ lies in $\phi^{-1}(K)$. □

Solution 25.10. We write $\zeta = re^{i\theta}$ in polar form. Note that r is non-zero as ζ lies in $\mathbb{C}^\times$. Applying the law of exponents, we find that $z = \log(r) + i\theta$ is one preimage of ζ. Thus $m \cdot z$ is a preimage of ζ^m. Since $\exp(2\pi i n) = 1$ we also have that

$$m \cdot z + n \cdot (2\pi i)$$

is a preimage of ζ^m. As m and n vary over $\mathbb{Z}$ this gives all elements of the preimage. That is, there is a homomorphism from $(\mathbb{Z}^2, +)$ (with coordinate addition) to the subgroup $L = \exp^{-1}(Z)$. If $r \neq 1$, or if θ is not a rational multiple of π, then this homomorphism is an isomorphism between $(\mathbb{Z}^2, +)$ and L. □

Solution 25.20. Suppose that h lies in H and g lies in G. Then $gh = hg$, by commutativity. Thus $gH = Hg$ as desired. □

Solution 27.8. We solve the given linear congruences. In each case we give a few words of justification.

(1) The set of integers x solving $8x \equiv 1 \pmod 7$ equals the set solving $x \equiv 1 \pmod 7$ because $8 \equiv 1 \pmod 7$. That is, the desired answer is $[1]_7 = \{1 + 7k \mid k \in \mathbb{Z}\}$.

(2) We note that $3 \cdot 2 = 6 \equiv -1 \pmod 7$. So $3 \cdot (-2) \equiv 1 \pmod 7$. Since $7 + (-2) = 5$ we deduce that $3 \cdot 5 = 15 \equiv 1 \pmod 7$. That is, 2 and 5 are multiplicative inverses modulo 7. Thus the desired answer is $[5]_7$.

(3) To solve $3x \equiv 2 \pmod 7$ we multiply both sides by 5 (the inverse of 2, modulo 7) to obtain the congruence $15x \equiv 10 \pmod 7$. This has the same solutions as $x \equiv 3 \pmod 7$. So the desired answer is $[3]_7$.

(4) To solve $3x \equiv 2 \pmod{13}$ we note that

$$3 \cdot 9 = 27 = 1 + 26 \equiv 1 \pmod{13}$$

So we may multiply both sides of the original congruence to obtain $27x \equiv 18 \pmod{13}$. This has the same solutions as $x \equiv 5 \pmod{13}$. So the desired answer is $[5]_1 3$.

(5) We note that x solves $6x \equiv 4 \pmod{26}$ if and only if there is an integer k so that $6x = 4 + 26k$. The latter holds if and only if $3x = 2 + 13k$. This congruence was solved immediately above; thus the desired answer is $[5]_1 3$. □

Solution 29.11. Theorem 29.9 gives us integers c and d so that $ad - pc = 1$. Thus $ad \equiv 1 \pmod{p}$. That is, $[a]_n \cdot [d]_n = [1]_n$. Thus d is the desired multiplicative inverse. □

Solution 29.12. We prove this in the special case that $G = \mathbb{Z}$; the proof when G is finite is a consequence.

Suppose that $H < \mathbb{Z}$ is a subgroup. Let $H_{>0}$ be the positive elements of H. If this is empty then there is nothing to prove. So, appealing to the well-ordering principle, we take a to be a smallest element of $H_{>0}$.

We now claim that $H = \langle a \rangle$ (and so H is cyclic). To prove this, let b be any element of $H_{>0}$. By Theorem 29.9 there are integers c and d so that $ad - bc = \gcd(a, b)$. Thus $\gcd(a, b)$ lies in H. Since $\gcd(a, b) \leq a$ we deduce that $\gcd(a, b) = a$. Thus b is a multiple of a, as desired. □

Appendix B. Glossary

We give very short and informal explanations of some of the more common bits of mathematical jargon.

axiom – A mathematical statement assumed to hold, without further justification.

conjecture – A mathematical statement which is thought to hold.

corollary – A mathematical statement, following directly from a previously proved statement.

definition – A name for a mathematical object that already exists.

implication – A mathematical statement of the form "if P then Q". Here P and Q are themselves mathematical statements called the *hypothesis* and *conclusion* of the implication.

lemma – A mathematical statement, usually part of a larger result, followed by a proof.

proposition – A mathematical statement, usually not as substantial as a theorem, followed by a proof.

theorem – A mathematical statement followed by a proof.

Index

References

[1] Claude-Gaspar Bachet. *Problèmes plaisants & délectables, qui se font par les nombres*. Lyon, 1624. https://mdz-nbn-resolving.de/details:bsb10081407. [84]

[2] Rittaud Benoît and Albrecht Heeffer. The pigeonhole principle, two centuries before Dirichlet. *Mathematical Intelligencer*, 36(2):27–29, 2014. doi:10.1007/s00283-013-9389-1. [53]

[3] Yale Babylonian Collection, 2022. Images of YBC 07289. https://collections.peabody.yale.edu/search/Record/YPM-BC-021354. [45]

[4] Herbert B. Enderton. *Elements of set theory*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1977. [48]

[5] Euclid. *The thirteen books of Euclid's Elements translated from the text of Heiberg. Vol. I: Introduction and Books I, II. Vol. II: Books III–IX. Vol. III: Books X–XIII and Appendix*. Dover Publications, Inc., New York, 1956. Translated with introduction and commentary by Thomas L. Heath, 2nd ed. [39, 41, 50, 81]

[6] Euclid. *Elementa (MS. D'Orville 301)*. Stephanos the clerk, Byzantine Empire, 888. https://medieval.bodleian.ox.ac.uk/catalog/manuscript_4146. [41]

[7] Carl Friedrich Gauss. *Disquisitiones arithmeticae*. Yale University Press, New Haven, Conn.-London, 1966. Translated into English by Arthur A. Clarke, S. J. [80]

[8] Paul R. Halmos. *Naive set theory*. The University Series in Undergraduate Mathematics. D. Van Nostrand Co., Princeton, N.J.-Toronto-London-New York, 1960. [1, 2, 52]

[9] Camille Jordan. Mémoire sur le numbre des valeurs des fonctions. *Journal de l'École Polytechnique*, 22:113–194, 1861. https://gallica.bnf.fr/ark:/12148/bpt6k433691p/f171.item. [73]

[10] Joseph-Louis Lagrange. Suite des réflexions sur la résolution algébrique des équations. Section troisieme. De la résolution des équations du cinquieme degré & des degrés ultérieurs. *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Berlin*, pages 138–254, 1771. https://books.google.co.uk/books?id=_-U_AAAAYAAJ&pg=PA202. [73]

[11] Lars-Daniel Öhman. Are induction and well-ordering equivalent? *Math. Intelligencer*, 41(3):33–40, 2019. doi:10.1007/s00283-019-09898-4. [50, 51]

[12] Tilman Piesk. The 52 partitions of a 5-element set represented by matrices of equivalence relations, 2012. https://en.wikipedia.org/wiki/Equivalence_relation. [92]

[13] Jeffrey Shallit. Origins of the analysis of the Euclidean algorithm. *Historia Math.*, 21(4):401–419, 1994. doi:10.1006/hmat.1994.1031. [84]

[14] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, 1997. doi:10.1137/S0097539795293172. [80]

[15] N. J. A. Slone. Bell numbers, 2022. Entry A000110 in the Online Encyclopedia of Integer Sequences. http://oeis.org/A000110. [22]

[16] Kurt Von Fritz. The discovery of incommensurability by Hippasus of Metapontum. *Annals of Mathematics*, 46(2):242–264, 1945. doi:10.2307/1969021. [45]