Math 402: Group Theory

S. Paul Smith

Department of Mathematics, Box 354350, Univ. Washington, Seattle, WA 98195
E-mail address: smith@math.washington.edu
Contents

Chapter 0. In the beginning... 5
  1. The language of sets 5
  2. Functions 11
  3. Writing Mathematics 17

Chapter 1. Integers 21
  1. The natural numbers 21
  2. Division and remainders 22
  3. The Euclidean Algorithm 23

Chapter 2. Groups 27
  1. Permutations 27
  2. Groups and Examples 32
  3. The groups $\mathbb{Z}/d$ 40
  4. Isomorphisms 42
  5. Subgroups 50
  6. Cosets and Lagrange’s Theorem 55
  7. Cyclic groups 57
  8. The product of two groups 60
  9. Finite abelian groups 63
  10. Putting together some of the ideas in this chapter 65

Chapter 3. Homomorphisms between groups 67
  1. Definition 67
  2. Normal subgroups 68
  3. The quotient group construction 69
  4. ??? 73
  5. Conjugacy 75

Chapter 4. Actions of groups on sets 79
  1. Actions 79
  2. The dihedral groups 80
  3. Orbits and stabilizers 82
  4. Applications to finite groups 84
  5. Fractional linear transformations 87

Chapter 5. Small Groups 91
  1. Groups of order 4 92
2. Groups of order 6 92
3. Groups of order 8 92
4. Groups of order 9 92
5. Groups of order 10 92
6. Groups of order 12 92

Chapter 6. The symmetric groups, again 95
1. Conjugacy classes 95
2. Generating $S_n$ 98
CHAPTER 0

In the beginning...

Every area of human enquiry requires its own particular language. That language is built up over a long period of time. It evolves in order to express the ideas that arise in the area. New words are created as new concepts develop. These languages are often impenetrable to outsiders. Usually there are good reasons for this. Mathematics is no exception in this regard.

Two fundamental concepts in mathematics are the notions of sets and functions. In order to state the definitions and results in any branch of mathematics we use the language of set theory. We use the English language too. In Spain though, they use the Spanish language together with the language of sets. It might be an enjoyable exercise for you to look at a mathematics book written in Spanish and observe that there is a lot in it you can understand without speaking Spanish simply because Spanish and English mathematicians use the same set-theoretic language, or notation.

Notation. We will use the standard notations of set theory. The notation is important. It is designed so we can say things briefly and precisely. It is part of the language of mathematics. You must learn to use it, and use it correctly. You must not, for example, confuse the symbols $\in$ and $\subseteq$; the first is used for elements, the second for subsets; it is ok to say $2 \in \mathbb{Z}$, but not ok to say $\{2\} \in \mathbb{Z}$. the first of these reads “$2$ is an element of $\mathbb{Z}$, and the second of these reads “the set consisting of $2$ is an element of $\mathbb{Z}$; it is ok to write $\{2\} \subseteq \mathbb{Z}$ because this reads “the set consisting of $2$ is a subset of $\mathbb{Z}$.

It is common practice to use upper case letters for sets and lower case letters for elements. Of course, we can not always follow this practice—for example, the subsets $E = \{\text{even integers}\}$ and $O = \{\text{odd integers}\}$ of the integers are themselves elements of the two element set $\{O, E\}$.

1. The language of sets

1.1. A set is a collection of things. Not necessarily mathematical things: there is the set of US presidents, past and present, the set of people who are alive, the set of dogs, the set of single women with red hair that own a house in Paris, the set of letters in the Greek alphabet, the set of popsicles, the set of New Zealand citizens, and so on. Mathematical examples include the set of whole numbers (called integers), the set of prime numbers, the set
of even numbers, the set of odd numbers, the set of squares, the set of lines in the plane \( \mathbb{R}^2 \), the set of pairs of distinct points in 3-space, and so on.

Usually we use an upper case letter to denote a set and when specifying the set we use curly parentheses. For example, we could write

\[
A = \{ \text{living mothers} \} \\
P = \{ \text{prime numbers} \} \\
E = \{ \text{even numbers} \} \\
O = \{ \text{odd numbers} \}.
\]

Particularly important sets have special names and are denoted by special symbols. For example,

\[
\mathbb{N} = \{ \text{the natural numbers} \} = \{ 0, 1, 2, 3, \ldots \} \\
\mathbb{Z} = \{ \text{integers} \} = \{ \text{whole numbers} \} \\
= \{ \ldots, -2, -1, 0, 1, 2, \ldots \} \\
\mathbb{Q} = \{ \text{rational numbers} \} = \{ \text{fractions} \} \\
= \{ \frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0 \} \\
\mathbb{R} = \{ \text{real numbers} \} \\
\mathbb{C} = \{ \text{complex numbers} \}
\]

\( \mathbb{R}^2 \) = \{ \text{the points in the plane} \}.

We also improvise on these standard notations. Some examples are

\[
2\mathbb{N} = \{ \text{the even natural numbers} \} = \{ 0, 2, 4, 6, \ldots \} \\
\mathbb{Z}_{\leq 0} = \{ \text{the non-positive integers} \} = \{ \ldots, -3, -2, -1, 0 \ldots \} \\
3\mathbb{Z} = \{ \text{all integer multiples of 3} \} = \{ \ldots, -9, -6, -3, 0, 3, 6, 9, \ldots \} \\
\mathbb{R}_{\geq 0} = \{ \text{the non-negative real numbers} \} = \{ r \in \mathbb{R} \mid r \geq 0 \},
\]

Of particular importance is the \textbf{empty set}, denoted \( \emptyset \), that has no elements at all. It might seem a little odd to talk about the empty set and to have a special symbol for it, but think of the parallel with the symbol we use for zero, 0. It is quite interesting to read about the history of zero on Wikipedia. Check it out!

1.2. Elements. The things that belong to a set are called its \textbf{elements}. For example, the prime number 37 is an element of the set \( P \) above. The point \((3, -2)\) belongs to the set \( \mathbb{R}^2 \). The number \( \pi \) is an element the set \( \mathbb{R} \).

We usually use lower case letters for the elements of a set. When \( x \) is an element of a set \( X \), we write \( x \in X \), and read this as \( x \) is an element of \( X \) or \( x \) belongs to \( X \) or \( X \) contains \( x \) or, simply, \( x \) is in \( X \).

Before we can talk about a particular set we need a precise description of it. One possibility is to list its elements—for example, we can define the set \( V \) consisting of the vowels by writing \( V = \{ a, e, i, o, u, y \} \). If a set is finite, i.e., it does not have an infinite number of elements in it, it might be
possible to describe the set by explicitly writing out a list of all its members. The set of people now living on planet earth is finite, but we are not able to list all the elements in it. However, if a set is infinite, or even finite but extremely large, this is not possible, so we must find some short way to describe the set precisely.

A set is completely determined by its elements. Two sets are equal if and only if they have the same elements.

1.3. The symbol meaning "such that". We already mentioned the set of rational numbers \( \mathbb{Q} \). You already know what fractions are but let’s define them using set notation:

\[
\mathbb{Q} = \left\{ \frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0 \right\}.
\]

The vertical symbol \( | \) is read as “such that”. Thus, the mathematical sentence (0-1) should be read as follows: \( \mathbb{Q} \) is the set of all numbers \( \frac{a}{b} \) such that \( a \) and \( b \) are integers and \( b \) is not zero.

Another common notation for “such that” is the colon. Using the colon the above sentence would be

\[
\mathbb{Q} = \left\{ \frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0 \right\}.
\]

I prefer the symbol \( | \) to the symbol : because it is more visible.

By using the symbol \( | \) we can give shorter definitions of sets. For example, if \( a \in \mathbb{Z} \), we write \( a\mathbb{Z} \) for the set of integer-multiples of \( a \) and can write this succinctly as

\[
a\mathbb{Z} := \{ax \mid x \in \mathbb{Z}\}.
\]

It is helpful to introduce the following notation. If \( A \) and \( B \) are sets of numbers we introduce the notation \( A + B \) for the set consisting of all sums \( a + b \) such that \( a \) is in \( A \) and \( b \) is in \( B \). More succinctly,

\[
A + B = \{a + b \mid a \in A, b \in B\}.
\]

In a similar way we introduce the notation

\[
a - b\mathbb{Z} = \{a - bx \mid x \in \mathbb{Z}\} \quad \text{and} \quad a + b\mathbb{Z} = \{a + bx \mid x \in \mathbb{Z}\}.
\]

More explicitly, for example, \( 1 + 3\mathbb{Z} = \{\cdots, -5, -2, 1, 4, 7, \cdots \} \).

1.4. Cardinality. We denote the number of elements in a set \( X \) by \(|X|\). We call the number of elements in a set its cardinality. For example, the cardinality of the empty set is zero; it is the only set whose cardinality is zero. Two sets are said to have the same cardinality if they have the same number of elements.

The cardinality of infinite sets is a subtle matter but we can’t say more about this until we discuss the notion of injective, surjective, and bijective functions. We will do this shortly and say more about cardinality once we have those notions in hand.
1.5. Subsets and containment. We say that a set $X$ is contained in a set $Y$ if every element of $X$ is an element of $Y$. More formally, we say $X$ is a subset of $Y$ if it is contained in $Y$ and write
\[ X \subset Y \]
to denote this situation. Thus the symbol “$\subset$” is read as is a subset of or is contained in. For example,
\[ \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}. \]
If $X$ is any set, then $X \subset X$ and $\phi \subset X$.

It is quite common to prove the equality of two sets $X$ and $Y$ by proving that $X \subset Y$ and $Y \subset X$. In other words, $X = Y$ if and only if $X \subset Y$ and $Y \subset X$.

1.6. Intersection and union. Two important operations on sets are intersection and union. They bear some resemblance to addition and multiplication of numbers.

The intersection of sets $X$ and $Y$ consists of the elements that are in both $X$ and $Y$. It is denoted by $X \cap Y$. Using the examples above, we have, for example,
\[ P \cap E = \{2\} \]
because 2 is the only even prime number. You may know that there are infinitely many primes, so $P \cap O$ is an infinite set, i.e., it has infinitely many elements. Notice we wrote $P \cap E = \{2\}$ not $P \cap E = 2$. There is an important difference—the intersection of two sets is a set, and the number 2 is different from the set whose only element is 2.

The union of sets $X$ and $Y$, which we denote by $X \cup Y$, consists of the elements that are in either $X$ or $Y$. For example, \( \{1, 2, 3\} \cup \{2, 3, 4\} = \{1, 2, 3, 4\} \). Likewise,
\[ O \cup E = \mathbb{Z} \]
because every number is either even or odd. Notice that
\[ O \cap E = \phi \]
because there are no numbers that are both even and odd.

If $X$ is a subset of $Y$ it is clear that $X \cap Y = X$ and $X \cup Y = Y$.

1.7. OR versus EXCLUSIVE OR. Sometimes students are confused by the way mathematicians use the word ”or”. In everyday language we tend to use the word ”or” in a context where something is one thing or the other but can not be both. An example will help. We might say, ”my leg hurts” and the listener might respond with the question ”does your left leg or your right leg hurt” assuming that only one of your legs hurts and you will answer by stating which of your two legs hurts. However, the pedantic mathematician might answer “yes”, meaning that it is true that at least one of his legs hurts.
In mathematics when we say that either $P$ or $Q$ is true we mean that one of the following is true:

- $P$ is true but $Q$ is false;
- $Q$ is true but $P$ is false;
- both $P$ and $Q$ are true.

Thus, if someone asks “is 729 even or odd”, the pedantic mathematician will answer “yes”. Frustrating for others, but that’s the way it is.

To distinguish the mathematician’s notion of OR from the general public’s idea mathematicians introduce the notion of what is called “exclusive or”. We sometimes write XOR to mean “exclusive or”. Thus, one would answer yes to the question “is $P$ XOR $Q$ true” only if

- $P$ is true but $Q$ is false;
- $Q$ is true but $P$ is false;

I haven’t explained that very well...read about it on the web.

1.8. Properties of intersection and union.

1.8.1. Arithmetic operations. You already know the basic properties of the arithmetic operations $+$ and $\times$. Both $+$ and $\times$ are associative operations by which we mean that

$$a + (b + c) = (a + b) + c \quad \text{and} \quad a \times (b \times c) = (a \times b) \times c$$

for all numbers $a$, $b$, and $c$. The associativity properties imply that the expressions $a + b + c$ and $a \times b \times c$ are unambiguous.

Both $+$ and $\times$ are commutative operations by which we mean that

$$a + b = b + a \quad \text{and} \quad a \times b = b \times a.$$ 

Slightly more sophisticated because it involves both operations, addition and multiplication, is the distributive law

$$a \times (b + c) = a \times b + a \times c.$$ 

The number zero has two special properties, namely

$$0 \times a = 0 \quad \text{and} \quad 0 + a = a$$

for all numbers $a$. The number 1 behaves with respect to $\times$ in a way that is analogous to the way 0 behaves with respect to $+$, namely

$$1 \times a = a$$

for all numbers $a$. 
1.8.2. Set operations. There are analogous properties for the set operations \( \cup \) and \( \cap \). For any sets \( X, Y, \) and \( Z \),
\[
X \cup Y = Y \cup X \\
X \cap Y = Y \cap X \\
X \cup (Y \cup Z) = (X \cup Y) \cup Z \\
X \cap (Y \cap Z) = (X \cap Y) \cap Z \\
X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z) \\
\phi \cap X = \phi \\
\phi \cup X = X.
\]
Mathematicians like it when there are similarities like this between the arithmetic operations + and \( \times \) and the set operations \( \cup \) and \( \cap \). Of course, there are some significant differences too. For example, \( X \cap X = X \cup X = X \).

One other difference is that there are two distributive laws for \( \cap \) and \( \cup \) but only one distributive law for + and \( \times \) (+ does not distribute across \( \times \)).

All these properties are easy to check. The only ones that might require some care are the distributive laws. You should try to prove them yourself. The strategy to use is that I mentioned earlier for showing two sets are equal: show each is a subset of the other.

There is also a similarity between \( \subset \) and \( \leq \). It is a good exercise for you to write down some of the similarities.

1.9. Disjoint union. The union of two subsets \( B \) and \( C \) of a set \( A \) is denoted by \( B \cup C \). If \( B \cap C = \phi \), we say that \( B \) and \( C \) are disjoint. We sometimes write \( B \triangle C \) to denote the union of two disjoint sets, and call it the disjoint union of \( B \) and \( C \). For example, \( \mathbb{Z} \) is the disjoint union of the even and the odd numbers. More generally, if \( n \) is a positive integer, then \( \mathbb{Z} \) is the disjoint union of the \( n \) subsets,
\[
i + n\mathbb{Z} := \{a \in \mathbb{Z} \mid a \text{ leaves a remainder of } i \text{ when divided by } n\}
\]
as \( i \) runs through the numbers 0 to \( n - 1 \). Of course, you know this already but a rigorous proof requires some thought (see Theorem 11.1).

Notice that \( n\mathbb{Z} \) consists of the numbers divisible by \( n \), which is exactly \( \{nb \mid b \in \mathbb{Z}\} \); that is why we write \( n\mathbb{Z} \) for this set—it is all multiples of \( n \). And the notation \( i + n\mathbb{Z} \) is explained by the fact that \( i + n\mathbb{Z} = \{i + nb \mid b \in \mathbb{Z}\} \).

Of course, we could define \( i + n\mathbb{Z} \) for any integer \( i \) to be \( \{i + nb \mid b \in \mathbb{Z}\} \).

We will do this later, but you need to be warned that \( i + n\mathbb{Z} = j + n\mathbb{Z} \) if \( i - j \) is divisible by \( n \). For example, the set of odd numbers is \( 1 + 2\mathbb{Z} = 3 + 2\mathbb{Z} = -57 + 2\mathbb{Z}, \ldots \)

1.10. Set difference. If \( A \subseteq B \) we define
\[
B - A := \{b \in B \mid b \notin A\}.
\]
2. Functions

We also use this notation when $A$ and $B$ are subsets of a set $C$, even though $B$ need not be contained in $A$. For example, if $A = \{0, 1, 2, 3, 4\} \subset \mathbb{Z}$ and $E$ is the set of even integers, then $A - E = \{1, 3\}$.

1.11. Products. The Cartesian product of sets $X$ and $Y$ is the set

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

that is, $X \times Y$ consists of all ordered pairs $(x, y)$ in which $x$ belongs to $X$ and $y$ belongs to $Y$. The Cartesian product $Y \times X$ consists of all ordered pairs $(y, x)$ in which $x$ belongs to $X$ and $y$ belongs to $Y$. Note that $X \times Y \neq Y \times X$ unless $X = Y$.

Notice this: if $X$ and $Y$ are finite, we have the formula

$$|X \times Y| = |X| \times |Y|.$$ 

That’s why we use the word “product”; the number of elements in the product of a sets is the product of the number of elements in each set. In fact, it is interesting to pause and think that thousands of years ago, before man had much mathematics, he must have had a notion of Cartesian product: if we have 3 families and each family needs two spears we need a total of six spears. I’ve expressed that poorly, but when you think about it the notion of multiplication must surely have arisen after the notion (intuitive and not explicitly expressed) of Cartesian product.

You should also wonder why we use the word Cartesian. Any ideas?

2. Functions

2.1. A function $f : X \to Y$ from a set $X$ to a set $Y$ is a rule that assigns to each element in $x$ an element $f(x)$ in $Y$. We also call a function a map or mapping.

A deceptively important, deceptive because apparently trivial, function is the identity function $id_X : X \to X$ defined by $id_X(x) = x$ for all $x \in X$. There are many identity functions, one for each set.

2.2. Composition of functions. We can compose two functions: If $f : X \to Y$ and $g : Y \to Z$, we define $g \circ f : X \to Z$ to be the function

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

We call $g \circ f$ the composition of $g$ and $f$. Sometimes we omit $\circ$ and just write $gf$.

2.2.1. Composition of functions is associative. Composition is associative, meaning that

$$(hg)f = h(gf)$$

whenever these compositions make sense; we usually write $hgf$ for this. Do not read this sentence only once and pass on, think about why it is true. Can you prove it? Associativity of composition of functions is every bit as important as the fact that multiplication of numbers is associative. Look
at the proof of Lemma 0.2 and isolate the step in which associativity of composition is used.

Associativity is a deceptively simple property. Too often beginners use it in situations where it does not apply. One of my favorite examples of its misuse, if one can use the word “favorite” to choose among terrible sins, occurs in elementary calculus. Certainly, students in an elementary calculus course would agree that subtraction is not an associative operation because \((a - b) - c \neq (a - b) - c\) in general. Perhaps with a little more thought such students might agree that division is not an associative operation. Despite that, I have seen many students write the following

\[
\frac{d}{dx} \left( \frac{1}{x} \right) = \lim_{h \to 0} \frac{1}{x+h} - \frac{1}{x} = \frac{-h}{x(x+h)}
\]

but the last expression is ambiguous because \(-\) is not associative. For example,

\[
\frac{\frac{1}{2}}{2} \neq \frac{1}{\frac{2}{2}}.
\]

2.2.2. The identity function. Identity functions have the following important property: if \(f : X \to Y\) is any function, then

\[f \circ \text{id}_X = f = \text{id}_Y \circ f.\]

In this regard the identity function behaves like the number 1. Perhaps a better analogy is with matrices: there are many identity matrices, one of size \(n \times n\) for each \(n \in \mathbb{N}\), and the result of the product of a matrix \(A\) with the appropriate sized identity matrix is \(A\).

If \(f : X \to X\) we often write \(f^2\) rather than \(ff\) or \(f \circ f\). Likewise we write \(f^3\) for \(fff\) and so on. It is convenient to define \(f^0\) to be \(\text{id}_X\). We then have the wonderful formula

\[f^m f^n = f^{m+n}\]

for all \(m, n \in \mathbb{N}\).

2.3. Injective, surjective, and bijective, functions. We now meet some particularly important classes of functions.

Let \(f : X \to Y\). We say that \(f\) is

- \textbf{injective}, or \textbf{one-to-one} or \textbf{1-1}, if \(f(x) \neq f(x')\) whenever \(x \neq x';\)
- \textbf{surjective}, or \textbf{onto}, if each \(y \in Y\) is equal to \(f(x)\) for some \(x \in X;\)
- \textbf{bijective} if it is both injective and surjective.

An injective, surjective, or bijective function is called an \textbf{injection}, \textbf{surjection}, or \textbf{bijection} for short.

To check your understanding of these notions prove the following result.

\textbf{Lemma 0.1.} Let \(g : X \to Y\) and \(f : Y \to Z\) be functions.

\footnote{Many people think "one-to-one" is a lousy name. I agree. It would be better to say such a function is "different-to-different. After all, that is its defining property: \(f\) is injective if it sends different elements of \(X\) to different elements of \(Y\).}
(1) If \( f \circ g \) is injective so is \( g \).
(2) If \( f \circ g \) is surjective so is \( f \).
(3) \( f \circ g \) is injective if \( f \) and \( g \) are.
(4) \( f \circ g \) is surjective if \( f \) and \( g \) are.

You should also decide whether the converse of each of these statements is true or not.

2.3.1. The range of a function. The range of a function \( f : X \to Y \), which we denote by \( \mathcal{R}(f) \), is the set of all values it takes, i.e.,

\[
\mathcal{R}(f) := \{ f(x) \mid x \in X \}.
\]

I like this definition because it is short but some prefer the equivalent definition

\[
\mathcal{R}(f) := \{ y \in Y \mid y = f(x) \text{ for some } x \in X \}.
\]

2.3.2. Consider the function

\[
f : \{ \text{men} \} \to \{ \text{women} \}, \quad f(x) := \text{the mother of} \ x.
\]

Is \( f \) well-defined, injective, surjective, bijective? What is the range of \( f \),
Give reasons, in standard English prose, for your answers.

2.3.3. The sine function \( \sin : \mathbb{R} \to \mathbb{R} \) is not one-to-one because \( \sin \pi = \sin 2\pi \). It is not onto because its range is \([-1, 1] = \{ y \mid -1 \leq y \leq 1 \} \). However, if we consider \( \sin \) as a function from \( \mathbb{R} \) to \([-1, 1] \) it becomes
onto, though it is still not one-to-one. If we consider \( \sin \) as a function from \([-\frac{\pi}{2}, \frac{\pi}{2}] \) to \([-1, 1] \) it is becomes bijective. That is why we define the inverse sine function \( \sin^{-1} \) as a function from \([-1, 1] \) to \([-\frac{\pi}{2}, \frac{\pi}{2}] \).

2.3.4. The function \( f : \mathbb{R} \to \mathbb{R} \) given by \( f(x) = x^2 \) is not one-to-one because \( f(2) = f(-2) \) for example, and is not onto because its range is \( \mathbb{R}_{\geq 0} := \{ y \in \mathbb{R} \mid ly \geq 0 \} \). However, the function

\[
f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \quad f(x) = x^2,
\]

is both one-to-one and onto.

2.3.5. Exercise. Let \( X \) be a finite set and \( f : X \to X \) a function. Show that \( f \) is injective if and only if it is surjective.

2.3.6. Domain and codomain. If \( f \) is a function from \( X \) to \( Y \) we call \( X \) the domain and \( Y \) the codomain of \( f \). I don’t use these words very often myself but you can read this paragraph as part of your general cultural education. The examples in sections 2.3.3 and 2.3.4 show that for a give formula \( f(x) \) the question of whether \( f \) is injective or surjective depends in a delicate way on the choice of domain and codomain.

If we want to be very precise, we define a function as a triple \( (f, X, Y) \) consisting of a domain \( X \), a codomain \( Y \), and a subset \( \Gamma_f \) of \( X \times Y \) such that if \( x \in X \), there is a unique \( y \in Y \) such that \( (x, y) \in \Gamma_f \). We then write \( f(x) \) for the unique element in \( Y \) such that \( (x, f(x)) \) is in \( \Gamma_f \). We call \( \Gamma_f \) the graph of the function.
Thus, when we just give a formula \( f(x) \) to define a function we are not really giving a complete definition. We must also state the domain and codomain to give a complete definition of a function.

**Lemma 0.2.** The following properties of a function \( f : X \to Y \) are equivalent:

1. \( f \) is bijective;
2. for each \( y \) in \( Y \) there is a unique \( x \in X \) such that \( f(x) = y \);
3. there is a unique function \( g : Y \to X \) such that \( fg = id_Y \) and \( gf = id_X \).

**Proof.** (1) \( \Rightarrow \) (2) Let \( y \in Y \). Since \( f \) is surjective, \( y = f(x) \) for some \( x \in X \). Since \( f \) is injective there can be only one \( x \in X \) such that \( f(x) = y \).

(2) \( \Rightarrow \) (3) Define \( g : Y \to X \) by declaring that \( g(y) \) is the unique \( x \in X \) such that \( f(x) = y \). If \( y \in Y \), then \( fg(y) = y \) because \( g(y) \) is defined just so it has the property that \( f(g(y)) = y \). Thus \( fg = id_Y \). If \( x \in X \), and \( y = f(x) \), then \( g(y) \) is defined to be \( x \); i.e., \( x = g(y) = gf(x) \). Thus \( gf = id_X \). Thus the function \( g \) we have just defined has the claimed properties.

Suppose \( g' : Y \to X \) also had the property that \( fg' = id_Y \) and \( g'f = id_X \). Let \( y \in Y \). Then \( y = id_Y(y) \) so

\[
g(y) = g(id_Y(y)) = g(fg'(y)) = (gf)(g'(y)) = id_X(g'(y)) = g'(y).
\]

Thus \( g(y) = g'(y) \) for all \( y \in Y \) and we conclude that \( g = g' \). (Notice we used the associative law for composition of functions in our calculation.)

(3) \( \Rightarrow \) (1) Since \( fg = id_Y \) and \( id_Y \) is surjective, \( f \) is surjective. Since \( gf = id_X \) and \( id_X \) is injective, \( f \) is injective. \( \square \)

### 2.4. The inverse of a function, if it exists.

The function \( g \) in part (3) of Lemma 0.2 is called the **inverse** of \( f \) and is denoted by \( f^{-1} \).

It may be defined by declaring that \( f^{-1}(y) = x \) provided that \( f(x) = y \). We then have

\[
f^{-1} \circ f = id_X \quad \text{and} \quad f \circ f^{-1} = id_Y.
\]

These formulas provide the defining property of \( f^{-1} \) provided it exists.

If \( f \) is a bijective function from a set \( X \) to itself and \( n \in \mathbb{N} \), we define \( f^{-n} \) to be

\[
(f^{-1})^n = f^{-1} \circ \cdots \circ f^{-1}.
\]

We then have the wonderful formula

\[
f^m f^n = f^{m+n}
\]

for all \( m, n \in \mathbb{Z} \).
2.4.1. *An example.* The inverse of the function \( f : \mathbb{R} \to \mathbb{R}, f(x) = 2x \), is the function \( g : \mathbb{R} \to \mathbb{R}, g(x) = \frac{1}{2}x \) because
\[
fg(x) = 2g(x) = 2\left(\frac{1}{2}x\right) = x \quad \text{and} \quad gf(x) = g(2x) = \frac{1}{2}(2x) = x.
\]

2.4.2. *Warning:* A function need not have an inverse! For example, the function \( h : \mathbb{Z} \to \mathbb{Z}, h(n) = 2n \), does not have an inverse. Why?

2.4.3. *Important:* Notice I defined the inverse of a function (if it exists) not an inverse. The result in the next lemma explains why I used the and not an.

**Lemma 0.3.** A function can have at most one inverse.

**Proof.** Let \( f : X \to Y \). Suppose \( g_1 : Y \to X \) and \( g_2 : Y \to X \) are such that
\[
g_1f = g_2f = \text{id}_Y \quad \text{and} \quad g_1f = g_2f = \text{id}_X.
\]
If \( y \in Y \), then \( y = \text{id}_Y(y) = g_1(y) \) so
\[
g_2(y) = g_2(fg_1(y)) = (g_2f)g_1)(y) = (\text{id}_Xg_1)(y) = g_1(y).
\]
Since \( g_1(y) = g_2(y) \) for all \( y \in Y \), \( g_1 = g_2 \). \( \square \)

Because \( f \) can have at most one inverse we denote that inverse by \( f^{-1} \) whenever it exists.

2.4.4. *One-sided inverses.* Let \( f : X \to Y \). It is possible for there to be a function \( g : Y \to X \) such that \( fg = \text{id}_Y \) but \( gf \neq \text{id}_X \). Likewise, there might be a function \( h : Y \to X \) such that \( hf = \text{id}_X \) but \( fh \neq \text{id}_Y \). For example, let \( \mathbb{N} = \{0, 1, 2, \ldots\} \) and let \( f : \mathbb{N} \to \mathbb{N} \) be the function \( f(n) = n + 1 \). Let \( g : \mathbb{N} \to \mathbb{N} \) be the function
\[
g(n) = \begin{cases} n & \text{if } n = 0 \\ n - 1 & \text{if } n \geq 1. \end{cases}
\]
Then \( gf = \text{id}_\mathbb{N} \) but \( fg \neq \text{id}_\mathbb{N} \) because \( fg(0) = 1 \). Thus \( g \) is a left inverse to \( f \) but not a right inverse. Similarly, \( f \) is a right inverse to \( g \) but not a left inverse.

However, if \( f : X \to Y \) has a left inverse and a right inverse, then the left inverse must equal the right inverse so \( f \) has an inverse. To see this, suppose there are functions \( g : Y \to X \) and \( h : Y \to X \) such that \( fg = \text{id}_Y \) and \( hf = \text{id}_X \). Then
\[
g = \text{id}_X \circ g = (hf) \circ g = (hf)g = h(fg) = h \circ \text{id}_Y = h.
\]

2.4.5. *Warning about notation.* Do not confuse the function \( f^{-1} \) with the function \( 1/f \). In fact, we have not even defined a function called \( 1/f \) so there can’t possibly be any confusion. Still, you might imagine a situation in which it is logical to use the label \( 1/f \) for a certain function related to \( f \); indeed, there are such situations but it need not be the case that \( 1/f \) has anything to do with \( f^{-1} \).
2.4.6. Exercise. Think of a situation in which one might reasonably label a function $1/f$ yet $1/f$ need have nothing to do with $f^{-1}$.

2.5. More about cardinality. Sets $X$ and $Y$ have the same cardinality if there is a bijection $f : X \to Y$. For example the set of integers has the same cardinality as the set of even integers, even though $2\mathbb{Z}$ is a proper subset of $\mathbb{Z}$. The function $f : \mathbb{Z} \to 2\mathbb{Z}$, $f(n) := 2n$, is a bijection.

Show that the property of having the same cardinality is an equivalence relation. The equivalence classes are called cardinals. It is convenient to actually say that $|X|$ denotes the equivalence class that $X$ belongs to. We write $|X| \leq |Y|$ if there is an injective function $f : X \to Y$. If $|X| \leq |Y|$ and $|Y| \leq |X|$ is $|X| = |Y|$?

A set having the same cardinality as $\mathbb{Z}$ is said to be countable.

Two important examples, both for illustrating the idea of cardinality, and for their historical impact are these: (1) $\mathbb{Z}$ and $\mathbb{Q}$ have the same cardinality; (2) $\mathbb{Z}$ and $\mathbb{R}$ have different cardinalities. The latter example shows there are different kinds of infinity. Indeed, we say that

Show for finite sets $X$ and $Y$ that there is a bijection (see below) $f : X \to Y$ if and only if $|X| = |Y|$.

Infinite cardinals. Different cardinals.

One of the most famous arguments in mathematics is Cantor’s diagonal argument to show that $|\mathbb{Q}| = |\mathbb{Z}|$. This is famous not just because it proves an important fact but famous because no one had even thought of asking the question before Cantor. Indeed, before Cantor there was not even a mathematical framework in which one could rigorously ask the question whether $\mathbb{Z}$ and $\mathbb{Q}$ have the same number of elements.

To appreciate this you might ask yourself this: if there are injective functions $f : X \to Y$ and $g : Y \to X$, does it make sense to say that $X$ and $Y$ have the same number of elements? Is $|X| = |Y|$?

Is $|\mathbb{R}| = |\mathbb{Z}|$. If not, is there a set $S$ with $|\mathbb{Z}| < |S| < |\mathbb{R}|$?

You can read about these topics on the web.

2.5.1. Exercise. Let’s see whether you have understood some of the above. First we introduce the notation

$$\text{Map}(X,Y) := \{ \text{all functions } f : X \to Y \}.$$  

If $|X| = |X'|$ and $|Y| = |Y'|$ prove that

$$|\text{Map}(X,Y)| = |\text{Map}(X',Y')|.$$  

Prove this by choosing explicit bijections $a : X \to X'$ and $b : Y \to Y'$ and using those to define an explicit bijection

$$\Phi : \text{Map}(X,Y) = \text{Map}(X',Y').$$

Let $X$, $Y$, and $Z$, be any sets; show there is a (natural) bijection

$$\Psi : \text{Map}(X \times Y,Z) \to \text{Map}(X,\text{Map}(Y,Z)).$$
3. Writing Mathematics

3.1. In a mathematics book you will see the words

Lemma, Proposition, Theorem, Corollary.

Each of these is followed by a precise statement of a result/fact, and that is followed by a proof of the statement. Or it should be! These four words indicate, to some extent, the status of the result. A theorem is very important and a proposition important. A lemma is usually a result proved in preparation for proving a theorem or proposition. Some lemmas that have a short and simple proof acquire a particular importance because they encapsulate simple observations that are used over and over again. For example, we frequently make use of the observation that a function from a finite set to itself is injective if and only if it is surjective. A corollary is a useful consequence of a theorem, usually with a much simpler proof than the theorem. A deep theorem might have many corollaries worth stating.

A good topic for discussion when you next kick back with some other math geeks is why we use such labels. Why do we bother to organize the results is such a way? Why do we care about proofs anyway? And, what constitutes a proof? Why are some proofs so hard to understand? If you find a particular proof difficult to understand it is often a good idea to try writing it out in a different way. For example, you might try to rearrange the order of the different paragraphs or steps in it. Why should we bother understanding the proofs in this course? The flippant answer to the last question is because your grade depends on it.

This will be the first or second course in which you encounter

the axiomatic method.

Although the axiomatic method is the rock on which Euclid’s books are built, it is only over the past 200 years that mathematicians have adopted this formal and precise way of presenting their subject as the right way to present the material.

The foundation of the axiomatic method consists of

Definitions.

We make precise definitions of the objects and precise definitions of their properties.

In a group theory course the primary objects are groups and rings; there are many secondary objects too that require precise definitions: subgroups, normal subgroups, quotients, ideals, ... It is all rather frightening at first. But it is through making precise definitions that we place our subject on a solid foundation; and I mean rock-solid. There can be no debate about what a group is once the definition has been made.
In a linear algebra course the primary objects are matrices, vector spaces, 
linear transformations, bases, and notions such as dimension, invertibility, 
and so on.

It is having rock-solid definitions and rock-solid proofs that distinguishes 
mathematics from all other lines of human inquiry. There is a certainty in 
mathematics that you will not see in physics, chemistry, biology, economics, 
psychology, sociology, astrology, and so on. All sciences strive to mathe-
maticize, to place themselves on a solid footing. Man seeks certainty.

Having precise definitions, and making precise statements in Theorems 
and their proofs, can give mathematics a dry appearance. However, the 
examples and applications are what give the subject life; an analogy with 
biology can help; the examples, for example the various groups that we know, 
the groups that nature has given us, are the living creatures that we study; 
they all have their special characteristics and features; there are patterns, 
they can occur in families, there are the cyclic groups, the symmetric groups, 
the simple groups, and so on. We study these creatures. The theorems and 
so on are reports on the results of our studies: for example, every symmetric 
group $S_n$ for $n \geq 5$ contains a simple subgroup with exactly half as many 
elements as itself; if $H$ is a subgroup of a group $G$, then the number of 
elements in $H$ divides the number of elements in $G$; and so on.

3.2. English. These notes are written in a mixture of standard english 
prose and mathematical notation. So are mathematics books. That is how 
to express the ideas of mathematics: in prose and notation, woven together 
into grammatically correct sentences. You must do the same in this course. 
It is the only way to express your ideas clearly. When you write out a 
solution to a problem your goal is to convey certain ideas to the reader. 
It is your responsibility to do that clearly and unambiguously. It is not the 
reader’s responsibility to struggle to discern your meaning. You must be the 
reader’s guide and friend, making his or her life easy.

You must begin every sentence with an upper-case letter. You must 
end each sentence with a period. A sentence has a subject, a verb, and an 
object. Sentences should not be unduly long. Follow the rules of English 
grammar. These rules are introduced in third grade and by sixth grade most 
children have learned them. You are now at a university and must write at 
the university level.

I have sometimes heard the complaint that this is a mathematics course, 
not an english course. But, whatever the field of human inquiry, the only 
way to convey its ideas is through language, and that language is always 
a mixture of everyday prose and the technical terminology and notation of 
the particular field.

So, write, and write well.

3.3. Proofs. You must write a lot of proofs in this course. A proof is 
an argument designed to convince your reader that something is true, or 
false.
A proof is not a list of equations that you have used to persuade yourself that you understand why something is true. A proof is an argument to persuade someone else that something is true. You must not expect the reader to assemble your list of equations into an argument by inserting appropriate phrases and punctuation. It is your job to do that. Your equations and calculations are typically the work you do prior to assembling an argument. A good argument will use your equations but will weave them into an argument.

Think of the analogy of cooking. A proof is something like a recipe. It is not just a list of ingredients. Certainly one needs to know the ingredients, but one must also know how to combine them and in what order and in what proportions. It is no good giving a list of ingredients and then in the narrative part of the recipe mentioning only some of them. You leave the reader baffled—was I supposed to add sugar or not? Likewise, in a proof, you should say everything that is necessary and nothing that is superfluous. Like cooking, it is a difficult skill. It takes years to develop and you can spend a lifetime honing that skill. Developing that skill is an important part of this course. Practice, experience, and judgment, are required to do this well.

Your proof should use phrases like “If..., then ...”, and words like “because”, “since”, “therefore”, “so”, “since”, ”but”, and so on.

3.4. Definitions. I will ask you to state a large number of definitions in the midterm and finals exams. Definitions are as important as proofs. I can do no better than to quote Giuseppe Peano:

\begin{quote}
There is no need to prove every theorem in class, but let us at least have precise concepts and correct definitions.
Rigor does not consist in proving everything. It consists in saying what is true and not saying what is not true.
\end{quote}

Definitions are important historical landmarks. They generally emerge over an extended period of time as mathematicians come to isolate the important, essential, and fruitful concepts. The important features of a mathematical object have names. For example, important features of a group include whether it is finite, abelian, cyclic, of prime order, simple, and so on? Other important features of a group involve its subgroups, conjugacy classes, and so on. To compare two groups we use homomorphisms, their kernels and images, isomorphisms, and so on. There is no avoiding the fact that you must learn these and other definitions. There are lots of them and it looks daunting at first. As with any new vocabulary, whether English or foreign, using the new words is the way to embed them in your mind. Use them in conversation with others. Doing many exercises will help those definitions to take root in your mind.
CHAPTER 1

Integers

We begin with elementary properties of the integers, the set
\[ \mathbb{Z} = \{ \ldots, -2, -1, 0, 1, 2, \ldots \}. \]
We will not prove anything new or surprising but we will examine carefully familiar properties of the integers and see what fundamental principles make them true.

1. The natural numbers

Children experience a degree of bewilderment when they first encounter negative numbers. Before that a child’s notion of number is confined to the set of natural numbers, the set
\[ \mathbb{N} := \{ 1, 2, 3, \ldots \}. \]
Before learning about the algebraic operations $+ \text{ and } \times$ a child understands the order relation on $\mathbb{N}$—a child is aware that two is larger than one, three larger than two, and so on.

That order relation is extended to the integers: if $x$ and $y$ are integers we say that $x$ is less than $y$, and write $x < y$, if $x + n = y$ for some $n \in \mathbb{N}$. We write $x \leq y$ if $x = y$ or $x < y$.

Even before they can speak, children are intuitively aware of the well-ordering principle:

\textit{every non-empty subset of } $\mathbb{N}$ \textit{has a smallest element.}

(Just ask a child to choose among several boxes of chocolate.) Although the well-ordering principle can be proved from a more fundamental set of axioms, we will take it on faith. No doubt, you have used the well-ordering principle for many years, and believe it, so the leap of faith is small. However, just because you take something for granted does not mean that we should not look afresh at it with a critical eye.

Actually, if you think about it, young children, and even monkeys and other animals, are aware of the following fact which is equivalent to the well-ordering principle:

\textit{every non-empty subset of } $\mathbb{N}$ \textit{has a largest element.}

Just ask a child to choose from a collection of different piles of jelly beans: if you say \textit{"you can choose one pile"} the child will unerringly zero in on the pile with the most jelly beans. Likewise with monkeys and bananas.

Why is this equivalent to the well-ordering principle?
Fractions. For now we will restrict our attention to the integers, but let’s keep in the back of our mind the rational numbers, or fractions, consisting of such things as 0, $\frac{15}{31}, 7\frac{1}{4}, \ldots$ More formally, we define the rational numbers to be

$$Q = \{ \frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0 \}.$$ 

Notice that a subset of the non-negative rational numbers need not have a smallest element: for example, $\{1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots\}$ has no smallest element.

2. Division and remainders

Let $a$ and $b$ be integers. We say that $b$ divides $a$, or that $a$ is a multiple of $b$, if there is an integer $c$ such that $a = bc$. We write $b|a$ to mean “$b$ divides $a$”, and $b \not| a$ to mean “$b$ does not divide $a$.

**Basic Facts about divisibility.** You should know and be able to use the following facts. We will use them frequently with no further comment.

1. every integer divides 0;
2. 1 divides every integer;
3. if $a|b$ and $x \in \mathbb{Z}$, then $a|bx$;
4. if $a|b$ and $b|c$, then $a|c$;
5. if $a|b$ and $a|c$, then $a|(bx + cy)$ for all $x, y \in \mathbb{Z}$;
6. if $a|b$ and $b|a$, then either $a = b$ or $a = -b$.

You should be able to prove all the above. You might just read out loud items (3)–(6) to see if you can translate the symbols into fluent sentences.

Generally speaking, when we try to divide one integer by another, we end up with a remainder; for example, trying to divide 89 by 12, we find that

$$89 = 12 \times 7 + 5.$$ 

We call 89 the **dividend**, 12 the **divisor**, 7 the **quotient**, and 5 the **remainder**.

**Theorem 1.1** (The division algorithm). Let $a, b \in \mathbb{Z}$ with $b > 0$. Then there exist unique $q, r \in \mathbb{Z}$ such that $a = bq + r$ and $0 \leq r < b$.

Some comments before the proof. We are dividing $a$ by $b$ to obtain a quotient $q$ and a remainder $r$. There is no requirement that $a$ be positive, although we do assume that $b$ is. Thus, $q$ might be negative, but $r$ is not negative. The point of the theorem is that $q$ and $r$ are unique! If we did not require $r$ to lie between 0 and $b - 1$, there would be no uniqueness.

The key step in the proof is the existence of $r$ which we will prove as a consequence of the well-ordering principle.

With this preamble, let’s prove the theorem.

**Proof.** [Theorem 1.1] Let

$$a - b\mathbb{Z} := \{a - bt \mid t \in \mathbb{Z}\}.$$ 

Because $b$ is positive, if $t$ is sufficiently negative $a - bt$ is positive; thus $a - b\mathbb{Z}$ contains non-negative integers; by the well-ordering principle $a - b\mathbb{Z}$ has a smallest element; we call that $r$ and let $q$ be the integer such that $a - bq = r$. 


If \( r \geq b \), then \( r - b \geq 0 \) and \( r - b = a - b(q - 1) \), so \( r - b \) is a non-negative element of \( a - b\mathbb{Z} \) and is smaller than \( r \); that contradicts our choice of \( r \) so we conclude that \( r < b \). Hence \( r \) and \( q \) exist as claimed, and it remains to prove their uniqueness.

Suppose \( a = bq' + r' \) and \( 0 \leq r' < b \). Then \( r = r' = b(q - q') \) and \( |r - r'| < b \); but the only integer that has absolute value strictly less than \( b \) and is a multiple of \( b \) is zero, so \( r = r' = 0 \); thus \( r = r' \) and \( q = q' \), proving uniqueness of \( q \) and \( r \).

\[ \square \]

**Corollary 1.2.** If \( n \) is a positive integer, then \( \mathbb{Z} \) is the disjoint union of the sets \( r + n\mathbb{Z} \), \( 0 \leq r \leq n - 1 \).

Here is a cute application of the division algorithm.

**Example 1.3.** If \( a \) is an integer that is divisible neither by 2 nor 3, then 24 divides \( a^2 - 1 \). We use the “rabbit out of the hat” method of proof. By the division algorithm, we can write \( a = 6b + r \) with \( 0 \leq r < 5 \). If \( r \) were 0, 2, or 4, then \( a \) would be even, hence divisible by 2; but we assumed this was not the case. If \( r \) were 3, then 3 would divide \( a \) contrary to our hypothesis. The only remaining possibility is that \( r = 1 \) or \( r = 5 \).

Thus we can write \( a = 6n \pm 1 \). Therefore

\[ a^2 - 1 = (6n + 1)^2 - 1 = 36n^2 + 12n = 12n(3n \pm 1). \]

If \( n \) is even, then 24 divides \( 12n \), and therefore divides \( a^2 - 1 \). If \( n \) is odd, then \( 3n \pm 1 \) is even, so again 24 divides \( a^2 - 1 \).

\[ \square \]

### 3. The Euclidean Algorithm

**3.1. Greatest common divisor.** The notion of greatest common divisor appears repeatedly in this course. If \( a \) and \( b \) are non-zero integers, their **greatest common divisor** is the largest integer that divides both \( a \) and \( b \). It is denoted \( (a, b) \) or \( \text{gcd}(a, b) \). E.g., \( (12, -12) = 12 \).

You should be asking yourself “do \( a \) and \( b \) actually have a greatest common divisor?” They do and we give the argument in the next paragraph.

Consider the set of integers that divide both \( a \) and \( b \). This set contains 1 so is non-empty. However, it cannot contain any number bigger than \( |a| \), the absolute value of \( a \), so it must have a biggest member. (Actually, writing out the details of this argument carefully, you will see we have used the Well Ordering Principle again: if \( X \) denotes the set of positive integers which divide both \( a \) and \( b \), then \( X' = \{|a| - x \mid x \in X \} \) is a non-empty set of non-negative integers, so has a smallest element, say \( x' = |a| - x \), and it is easy to check that \( x \) must be the largest member of \( X \), so \( x \) is the gcd).

**3.2. Finding the greatest common divisor.** The previous section shows that the greatest common divisor of \( a \) and \( b \) exists, but how do we go about finding it? We find it by using the **Euclidean Algorithm** which consists of repeatedly using the division algorithm. We begin with integers \( a \) and \( b \),
\( b > 0 \), and obtain sequences of integers \( q_0, q_1, q_2, \ldots \) and \( r_0, r_1, r_2, \ldots \) such that

\[
\begin{align*}
a &= bq_0 + r_0 & \text{and} & & 0 \leq r_0 < b, \\
b &= r_0q_1 + r_1 & \text{and} & & 0 \leq r_1 < r_0, \\
r_0 &= r_1q_2 + r_2 & \text{and} & & 0 \leq r_2 < r_1, \\
\vdots \\
r_{t-2} &= r_{t-1}q_t + r_t & \text{and} & & 0 \leq r_t < r_{t-1}, \\
r_{t-1} &= r_tq_{t+1}.
\end{align*}
\]

You keep dividing the latest remainder into the previous one; the procedure stops because the remainders keep getting smaller

\( b > r_0 > r_1 > \cdots \geq 0 \).

When one hits a remainder of zero, as indicated above, then \( r_t \) is the gcd of \( a \) and \( b \). Before proving this we illustrate the idea.

**Example 1.4.** Find \((1547, 560)\).

\[
\begin{align*}
1547 &= 560 \times 2 + 427, \\
560 &= 427 \times 1 + 133, \\
427 &= 133 \times 3 + 28, \\
133 &= 28 \times 4 + 21, \\
28 &= 21 \times 1 + 7, \\
21 &= 7 \times 3.
\end{align*}
\]

So \((1547, 560) = 7\). \(\diamondsuit\)

**Proposition 1.5.** The last remainder in the Euclidean Algorithm gives the gcd.

**Proof.** We adopt the notation set up above. Write \( d = \gcd(a, b) \). Since \( d \mid a \) and \( d \mid b \), \( d \mid r_0 = a - bq_0 \). Hence \( d \mid r_1 = b - r_0q_1 \); et cetera. Eventually, we see that \( d \mid r_t \). Since \( r_t > 0 \), \( d \leq r_t \). But also, \( r_t \mid r_{t-1} \), so \( r_t \mid r_{t-2} = r_{t-1}q_t + r_t \), and \( r_t \mid r_{t-3} = r_{t-2}q_{t-1} + r_{t-2} \); et cetera. Eventually, we see that \( r_t \mid b \) and \( r_t \mid a \). Therefore, \( r_t \) is the gcd. \(\square\)

**Theorem 1.6.** If \( d = (a, b) \), then there exist \( u, v \in \mathbb{Z} \) such that

\[ d = au + bv. \]

**Proof.** Just use the sequence of equalities in the Euclidean Algorithm from the bottom up to express \( r_t \) in terms of earlier remainders:

\[
\begin{align*}
r_t &= r_{t-2} - r_{t-1}q_t \\
      &= r_{t-2} - (r_{t-3} - r_{t-2}q_{t-1})q_t \\
\vdots & \quad \vdots
\end{align*}
\]
The next example makes the procedure clear.

**Example 1.7.** Reconsider Example 1.4. We will use the calculations in that exercise to show there are integers \( u \) and \( v \) such that \( 7 = 1547u + 560v \). We get

\[
7 = 28 - 1 \times 21 \\
= 28 - (133 - 28 \times 4) = 28 \times 5 - 133 \\
= (427 - 133 \times 3) \times 5 - 133 = 427 \times 5 - 133 \times 16 \\
= 427 \times 5 - (560 - 427) \times 16 = 427 \times 21 - 560 \times 16 \\
= (1547 - 560 \times 2) \times 21 - 560 \times 16 \\
= (1547 \times 21 - 560 \times 58).
\]

**Corollary 1.8.** If \( a | bc \) and \( (a, b) = 1 \), then \( a | c \).

Mention/look ahead to the Chinese Remainder Theorem.

**Theorem 1.9 (The Fundamental Theorem of Arithmetic).** *Every integer is a product of prime numbers in a unique way.*

3.2.1. **The greatest common divisor of several elements.** It should be obvious that we can define the greatest common divisor of any set of integers as the largest integer that divides all of them. We do that. You should check that

\[
\text{gcd} \left( a, \text{gcd}(b, c) \right) = \text{gcd} \left( \text{gcd}(a, b), c \right) = \text{gcd}(a, b, c),
\]

and so on for larger sets of integers.

3.3. **The least common multiple.**
CHAPTER 2

Groups

1. Permutations

A permutation of a set $X$ is a bijective function $\sigma : X \to X$. For example, the identity function $\text{id}_X$ is a permutation of $X$. We call it the trivial permutation.

In this section we examine individual permutations and the set of all permutations on $X$.

1.1. The set of permutations of a set $X$. Let $X$ be any set.

The set of permutations of $X$ is our first example of a group. If we combine our first result about permutations, Proposition 2.1 below, with the observation that $\text{id}_X$ is a permutation, and the fact that composition of functions is associative, we have, in effect, verified that the set of permutations of $X$ is a group.

So, all results about permutations in this section are, in fact, results about permutation groups or, as they are usually called when $X$ is finite, the symmetric groups. Symmetric groups are of great importance.

**Proposition 2.1.** Let $f$ and $g$ be permutations of $X$. Then

(1) $fg$ is a permutation of $X$;
(2) $f^{-1}$ is a permutation of $X$.

**Proof.** These are facts about bijections (see section 2.3 in Chapter 0). A composition of bijections is bijective. A bijection has an inverse and that inverse is itself a bijection. \qed

1.1.1. The number of permutations of a set. There are infinitely many permutations of an infinite set. For now we will mostly be concerned with permutations of a finite set.

If $n$ is a positive integer, then $\text{factorial}$ is the number

$$n! := n(n-1)(n-2)\cdots21,$$

the product of the integers between 1 and $n$. For example,

$$1! = 1, \quad 2! = 2, \quad 3! = 6, \quad 4! = 24, \quad 5! = 120,$$
$$6! = 720, \quad 7! = 5040, \quad 8! = 40320, \quad \text{and so on.}$$

It is convenient to define

$$0! = 1.$$
Lemma 2.2. There are $n!$ permutations of a set having $n$ elements.

Proof. Let’s assume $X = \{1, 2, \ldots, n\}$. In defining a bijective function $f : X \to X$, there are $n$ choices for $f(1)$ then, once $f(1)$ has been selected, $n - 1$ choices for $f(2)$, then $n - 2$ choices for $f(3)$, and so on, giving a total of $n \cdot (n - 1) \cdots 3 \cdot 2 \cdot 1 = n!$ choices for $f$. □

1.1.2. Convention and warning. If $\sigma$ and $\tau$ are permutations, the notation $\sigma \tau$ means first do $\tau$, then do $\sigma$. After all, since $\sigma$ and $\tau$ are functions what could be more sensible than using the convention we are already familiar with: $fg$ means first perform the function $g$ then the function $f$. However, not all books adopt this convention. For example, P.M. Cohn’s three volume text Algebra uses the opposite convention. In all other respects I love these books.

1.2. Cycle notation for permutations.
1.2.1. Remarks on notation. We need good notation for permutations. Sometimes it is easy to do that: for example, the function $f(n) = n + 5$ is a permutation of the set $\mathbb{Z}$. For permutations of finite sets, the case of most interest to us, we can rarely use such simple formulas.

Before dealing with permutations let’s consider some notations for integers. The notation 10010 for the number ten thousand and ten immediately gives us a sense of its size—it’s around ten thousand. It is also apparent that 10010 is divisible by 10, hence by the primes 2 and 5. But the notation 10010 does not reveal any arithmetic properties of 10010 other than its divisibility by 2 and 5. In contrast, if we write 10010 as $2 \times 5 \times 7 \times 11 \times 13$ we immediately see its arithmetic properties.

I don’t need to say anything about Roman numerals like MDCCCLIV. The fact that this notation was discarded long ago, except for some specialized and ceremonial uses, is sufficient testament to its inadequacies.

1.2.2. Analogy with prime numbers. We will adopt a notation for permutations that is analogous to writing an integer as a product of prime numbers. The analogue of a prime number is a cyclic permutation, or cycle, for short. In analogy with the fact that every integer can be written as a product of prime numbers in a unique way we will show that every permutation can be written as a product of disjoint cycles in a unique way. We will now explain these terms, first somewhat informally.

1.2.3. Cycles. Let’s start with an example. Let $\sigma$ be the permutation of \{1, \ldots, 9\} defined by

$$\sigma(1) = 1, \sigma(2) = 5, \sigma(3) = 7, \sigma(4) = 9, \sigma(5) = 2,$$
$$\sigma(6) = 4, \sigma(7) = 6, \sigma(8) = 8, \sigma(9) = 3.$$  

The expression for $\sigma$ as a product of disjoint cycles will be written

$$\sigma = (25)(37649).$$
The meaning is this. First, $\sigma$ is the product/composition of two simpler permutations/functions, $(25)$ and $(37649)$. The permutation denoted $(25)$ is the function that sends $2$ to $5$ and $5$ to $2$ and fixes the integers $1, 3, 4, 6, 7, 8, 9$. The permutation denoted $(37649)$ is the function that sends $3$ to $7$, $7$ to $6$, $6$ to $4$, $4$ to $9$, $9$ to $3$, and fixes the integers $1, 2, 5, 8$. We call the factors $(25)$ and $(37649)$ cycles. Each factor cycles around some of the numbers in $\{1, \ldots, 9\}$. The following picture to illustrate this:

\[ (25) = \begin{pmatrix} 2 & 5 \end{pmatrix}, \quad (37649) = \begin{pmatrix} 3 & 9 & 4 \end{pmatrix}. \]

Each arrow indicates the action of $\sigma$. For example, $7 \to 6$ means $\sigma(7) = 6$. The notation $(25)(37649)$ immediately reveals the behavior of $\sigma$, just as writing $10010$ as $2 \times 5 \times 7 \times 11 \times 13$ immediately reveals its fundamental arithmetic properties. The notation $(25)(37649)$ is efficient because we omit the numbers $i$ for which $\sigma(i) = i$. There could be no shorter notation than $(257)$ for the permutation $\tau$ defined by $\tau(2) = 5$, $\tau(5) = 7$, $\tau(7) = 2$, and $\tau(i) = i$ for all other $i$.

In some books the permutation $\tau = (257)$ will be written as

\begin{equation}
(2-1) \quad \begin{pmatrix} 1 & 2 & \ldots & 9 \\
1 & 2 & \ldots & 9
\end{pmatrix}.
\end{equation}

The general principle is that we write $\tau(i)$ below $i$, i.e.,

\[ \tau = \begin{pmatrix} 1 & 2 & \ldots & 9 \\
\tau(1) & \tau(2) & \ldots & \tau(9)
\end{pmatrix}. \]

This notation is cumbersome. You must examine it carefully to uncover the essential properties of $\tau$. And it is a lot of writing compared to $(257)$!

Likewise, \[ \begin{pmatrix} 1 & 2 & \ldots & 9 \\
1 & 2 & \ldots & 9
\end{pmatrix} \]

is a cumbersome and obscure notation for the permutation $(25)(3764893)$.

1.2.4. Warning: Different notations for the same cycle. Be warned that the permutations $(37649)$ and $(64937)$ are the same!

1.2.5. Disjoint cycles. The cycles $(25)$ and $(37649)$ are said to be disjoint because none of the terms in $(25)$ appears in $(37649)$.\footnote{Recall that two subsets, $A$ and $B$ say, of a set $X$ are disjoint if $A \cap B = \emptyset$. We say three subsets $A$, $B$, and $C$, are disjoint if $A \cap B = B \cap C = A \cap C = \emptyset$. And so on.} Likewise, the cycles $(138)$, $(24)$, and $(567)$ are disjoint. On the other hand the cycles $(25)$ and $(27649)$ are not disjoint because $2$ appears in each of them.

The permutation $\theta := (123)(34)(4526)(7153)$ is a product of four cycles but they are not disjoint. Our convention for composition of functions is that $fgh$ means first apply the function $h$, then $g$, then $f$. Hence $\theta$ is the
function first apply the permutation $\delta = (7153)$, then $\gamma = (4526)$, then $\beta = (34)$, then $\alpha = (123)$. So, the effect of $\theta$ is as follows:

\[
\begin{align*}
1 & \xrightarrow{\delta} 5 \xrightarrow{\gamma} 2 \xrightarrow{\beta} 2 \xrightarrow{\alpha} 3 \\
3 & \xrightarrow{\delta} 7 \xrightarrow{\gamma} 7 \xrightarrow{\beta} 7 \xrightarrow{\alpha} 7 \\
7 & \xrightarrow{\delta} 1 \xrightarrow{\gamma} 1 \xrightarrow{\beta} 1 \xrightarrow{\alpha} 2 \\
2 & \xrightarrow{\delta} 2 \xrightarrow{\gamma} 6 \xrightarrow{\beta} 6 \xrightarrow{\alpha} 6 \\
6 & \xrightarrow{\delta} 6 \xrightarrow{\gamma} 4 \xrightarrow{\beta} 3 \xrightarrow{\alpha} 1 \\
4 & \xrightarrow{\delta} 4 \xrightarrow{\gamma} 5 \xrightarrow{\beta} 5 \xrightarrow{\alpha} 5 \\
5 & \xrightarrow{\delta} 3 \xrightarrow{\gamma} 3 \xrightarrow{\beta} 4 \xrightarrow{\alpha} 4.
\end{align*}
\]

Hence $\theta = (13726)(45)$. In particular, $\theta$ is a product of disjoint cycles. Perhaps you can already see that every permutation of a finite set is a product of disjoint cycles. Although this might be intuitively obvious to you we will give a rigorous proof, i.e., an explanation, of this fact in Theorem 2.5 below.

An important property of disjoint cycles is that they commute with one another: for example, $(13726)(45) = (45)(13726)$. It is best for you to think about this yourself and why this is rather than have someone explain it to you. As another example, observe that the six different products of the three disjoint cycles $(13)$, $(247)$, and $(56)$, are equal to each other.

1.3. The orbits of a permutation. It is useful to have a graphic, dynamic mental image of a permutation. To encourage this we say “$\sigma$ acts on $X$” and speak of “the action of $\sigma$ on $X$.” The word “action” has a dynamic feel to it.

Let $x \in X$. The orbit of $x$ under the action of $\sigma$ is defined to be the set

$$
O_x := \{\sigma^n(x) \mid n \in \mathbb{Z}\} = \{\ldots, \sigma^{-2}(x), \sigma^{-1}(x), x, \sigma(x), \sigma^2(x), \ldots\}.
$$

We call $O_x$ a $\sigma$-orbit. The size of the orbit is the number of elements in it.

The association of the word orbit with the movement of planets further encourages us to have a dynamic picture of a permutation. Permutations of $X$ move the elements of $X$ around.

**Lemma 2.3.** Let $O_x$ and $O_y$ be the $\sigma$-orbits of $x$ and $y$. Then either $O_x = O_y$ or $O_x \cap O_y = \emptyset$.

**Proof.** Suppose $O_x \cap O_y \neq \emptyset$. Then $\sigma^m(x) = \sigma^n(y)$ for some integers $m$ and $n$. It follows that $\sigma^k(x) = \sigma^{k+n-m}(y)$ and $\sigma^\ell(y) = \sigma^{\ell+m-n}(x)$ for all integers $k$ and $\ell$. Hence $O_x \subseteq O_y$ and $O_y \subseteq O_x$. Thus $O_x = O_y$. \qed

If $A_1, A_2, \ldots$ are subsets of a set $X$ such that $X = A_1 \cup A_2 \cup \cdots$ and $A_i \cap A_j = \emptyset$ whenever $i \neq j$ we say that $X$ is the disjoint union of $A_1, A_2, \ldots$ and write $X = A_1 \sqcup A_2 \sqcup \cdots$ to denote this fact.
Proposition 2.4. Let \( \sigma \) be a permutation of \( X \). Then \( X \) is the disjoint union of its \( \sigma \)-orbits.

Proof. Since \( x \) belongs to the orbit \( O_x \), \( X \) is the union of all the \( \sigma \)-orbits. By Lemma 2.3, different orbits are disjoint. Hence if \( A_1, A_2, \ldots \) are the distinct orbits, then \( X = A_1 \sqcup A_2 \sqcup \cdots \). \( \square \)

Now assume \( \sigma \) is a permutation of a finite set \( X \). Then every \( \sigma \)-orbit is finite and there are only finitely many \( \sigma \)-orbits. Thus \( X = A_1 \sqcup A_2 \sqcup \cdots \sqcup A_k \) where \( A_1, A_2, \ldots, A_k \) are the different \( \sigma \)-orbits. Suppose

\[
A_1 = \{a_{11}, a_{12}, \ldots, a_{1r_1}\}
\]

and

\[
\sigma(a_{11}) = a_{12}, \quad \sigma(a_{12}) = a_{13}, \quad \ldots, \quad \sigma(a_{1r_1-1}) = a_{1r_1}, \quad \sigma(a_{1r_1}) = a_{11}.
\]

Then the action of the cycle \( (a_{11} a_{12} \ldots a_{1r_1}) \) on \( A_1 \) is the same as the action of \( \sigma \). Likewise, if

\[
A_2 = \{a_{21}, a_{22}, \ldots, a_{2r_2}\}
\]

and

\[
\sigma(a_{21}) = a_{22}, \quad \sigma(a_{22}) = a_{23}, \quad \ldots, \quad \sigma(a_{2r_2-1}) = a_{2r_2}, \quad \sigma(a_{2r_2}) = a_{21}
\]

then the action of the cycle \( (a_{21} a_{22} \ldots a_{2r_2}) \) on \( A_2 \) is the same as the action of \( \sigma \). And so on. It follows that

\[
\sigma = (a_{11} a_{12} \ldots a_{1r_1})(a_{21} a_{22} \ldots a_{2r_2}) \cdots (a_{k1} a_{k2} \ldots a_{kr_k}).
\]

Since the different orbits are disjoint we have proved the next result.

Theorem 2.5. Every permutation can be written as a product of disjoint cycles in a unique way.

Strictly speaking, we only proved this in the case when \( X \) is finite but the same idea and a little more notation will prove the same result for any set.

1.4. Cycles. You probably noticed that I have not given a precise definition of a cycle yet, woeful sinner that I am.

Suppose \( \sigma \) is a permutation of a set \( X \). Suppose too that \( \sigma \neq \text{id}_X \). We call \( \sigma \) a cycle if it has at exactly one orbit of size \( > 1 \). If that orbit has size \( r \) we call \( \sigma \) an \( r \)-cycle if the size of that orbit is \( r \).

We also declare that \( \text{id}_X \) is a cycle, the unique 1-cycle. We sometimes call it the trivial cycle.

Two non-trivial cycles \( \sigma \) and \( \tau \) are disjoint if their non-trivial orbits are disjoint.

The inverse of a cycle is a cycle. For example, \((3 5 2 6 1)^{-1} = (1 6 2 5 3)\). A 2-cycle is called a transposition.

Proposition 2.6. Every permutation is a product of transpositions.
Proof. Every cycle is a product of transpositions because, for example, 
\((12\ldots m - 1 m) = (1 m)(1 m - 1)\cdots(14)(13)(12)\). But every permutation 
is a product of cycles, so the result follows. □

There is no uniqueness to the factorization of a permutation as a product 
of transpositions. For example,

\[(1 2 3 4) = (1 2)(2 3)(3 4) = (1 4)(13)(12)\]

1.5. The symmetric groups. Let \(n\) be a positive integer. The set of 
all permutations of the set \(\{1, 2, \ldots, n\}\) is called the \(n\)th symmetric group, or 
just the symmetric group if \(n\) is clear from the context. It is often denoted 
by \(S_n\) or \(\Sigma_n\).

The number of elements in \(S_n\) is \(n!\).

Thus \(S_1\) has a single element, the identity permutation \(\text{id}_{\{1\}}\).

The second symmetric group \(S_2\) has two elements, \(\text{id}_{\{1,2\}}\) and \(1\).

The third symmetric group, \(S_3\), has six (=3!) elements, namely

\[1, \ (12), \ (23), \ (13), \ (123), \ (321)\]

The multiplication table for \(S_3\) is:

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
\cdot & 1 & (12) & (13) & (23) & (123) & (321) \\
\hline
1 & 1 & (12) & (13) & (23) & (123) & (321) \\
(12) & (12) & 1 & (321) & (123) & (23) & (13) \\
(13) & (13) & (123) & 1 & (321) & (12) & (23) \\
(23) & (23) & (321) & (123) & 1 & (13) & (12) \\
(123) & (123) & (13) & (23) & (12) & (321) & 1 \\
(321) & (321) & (23) & (12) & (13) & 1 & (123) \\
\hline
\end{array}
\]

The entry in the row labelled \(12\) and the column labelled \(123\) is the 
product \((12)(123) = (23)\). The entry in the row labelled \(123\) and the 
column labelled \(12\) is the product \((123)(12) = (13)\). And so on.

2. Groups and Examples

2.1. Definition and first observations. In the previous section we 
encountered a group, the symmetric group \(S_n\). Keep that example in mind 
when reading the next definition—why does \(S_n\) satisfy axioms \((1)-(3)\) in the 
definition?

Definition 2.7. A group \(G\) is a non-empty set \(G\) together with a map 
\(G \times G \rightarrow G\) that we usually denote by \((x, y) \mapsto xy\), satisfying the following properties:

(1) \((xy)z = x(yz)\) for all \(x, y, z \in G\);
(2) there is an element \(e \in G\) called the identity with the property that

\[ex = xe = x \quad \text{for all} \ x \in G;\]
(3) for each \(x \in G\) there is an element \(x^{-1}\) in \(G\) such that \(xx^{-1} = x^{-1}x = e\). We say the group is \textit{abelian} if \(xy = yx\) for all \(x, y \in G\).\footnote{Named after Niels Henrik Abel, 1802-29.}

We call the map \(G \times G \to G, (x, y) \mapsto xy\), the \textit{group operation} when we want to be a little vague but in particular situations we often call it multiplication, or addition, or \textit{composition}.

2.1.1. \textit{The symmetric group is a group.} Let’s check that \(S_n\), the set of permutations of the set \(\{1, 2, \ldots, n\}\), endowed with the binary operation "composition of functions" satisfies the axioms of Definition 2.7. A composition of bijections is a bijection so "composition of functions" is indeed a map \(S_n \times S_n \to S_n\). We also note that \(S_n\) is non-empty because it contains the identity function \(id : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}\). Composition of functions is associative so the associative axiom is satisfied. There is an identity element, namely the identity map. Finally, the inverse of a bijection is a bijection so every element in \(S_n\) has an inverse. Thus \(S_n\) is a group.

It is OK if you skip immediately to the examples in the next two sections and return to the next result after looking at some of those. The next lemma is an important, albeit elementary, result and you might find it easier to appreciate its proof after you have some examples in mind.

\textsc{Lemma 2.8.} Let \(G\) be a group. Then

(1) \(G\) has exactly one identity element;
(2) each \(x \in G\) has a unique inverse;
(3) you can cancel in a group: if \(x, y, z \in G\) and \(xy = xz\), then \(y = z\).

\textit{Likewise, if} \(yx = zx\), \textit{then} \(y = z\).

\textbf{Proof.} (1) If \(e\) and \(e'\) both satisfy condition (2) in Definition 2.7 they are equal: we have \(e = ee' = e'\), the first "\(=\)" because \(e'\) satisfies (2) and the second "\(=\)" because \(e\) satisfies (2).

(2) if \(x \in G\) and \(xy = yx = e\) and \(xz = zx = e\), then \(y = ye = y(xz) = (yx)z = ez = z\). The proof used condition (1), the associativity of multiplication.

(3) If \(xy = xz\), then

\[ y = ey = (x^{-1}x)y = x^{-1}(xy) = x^{-1}(xz) = (x^{-1}x)z = ez = z. \]

A similar string of equalities with \(x^{-1}\) on the right instead of the left proves the other cancellation rule. \hfill \Box

2.2. \textit{Some infinite groups.} You already know lots of groups! Let’s now mention some of them. Although most will be familiar look at each one in light of the axioms for a group and check that all three of the axioms are satisfied.
2.2.1. The group of integers. The first group you meet as a child is the group of integers \( \mathbb{Z} \) with the group operation being addition. We denote the result of the group operation by \( x + y \) rather than \( xy \). To check that \( \mathbb{Z} \) really is a group first observe that it is non-empty and that its addition is associative, \( x + (y + z) = (x + y) + z \). The identity is zero because \( 0 + x = x + 0 = x \) and the inverse of \( x \) is \(-x\).

Note that \( \mathbb{Z} \) is not a group under multiplication: although multiplication is associative, \( x(yz) = (xy)z \), and there is an identity, namely 1 (\( 1 \cdot x = x \cdot 1 = x \)), inverses do not exist—e.g., neither 2 nor 0 has an inverse.

2.2.2. The group of non-zero real numbers under multiplication. The non-zero real numbers with their ordinary multiplication form a group, an abelian group because \( xy = yx \). The identity element is 1, and the inverse of \( x \) is \( 1/x \). There are two common notations for the set of non-zero real numbers, \( \mathbb{R} - \{0\} \) and \( \mathbb{R}^\times \), and we often write \( (\mathbb{R}^\times, \cdot) \) or \( (\mathbb{R} - \{0\}, \cdot) \) for it—the first position denotes the set \( G \) and the second position denotes the operation. Following that convention, the previous example is \( (\mathbb{Z}, +) \) and the next example is \( (\mathbb{R}, \cdot) \).

2.2.3. The group of positive real numbers under multiplication. The set of positive real numbers with their ordinary multiplication form a group. We denote this group by \( (\mathbb{R}_{>0}, \times) \).

2.2.4. The group of real numbers under addition. The real numbers under addition form a group that we denote by \( (\mathbb{R}, +) \). The identity in \( (\mathbb{R}, +) \) is 0 and the inverse of \( x \) is \(-x\).

2.2.5. Rational numbers. Similar examples with the rational numbers in place of the reals are \( (\mathbb{Q} - \{0\}, \cdot) \), \( (\mathbb{Q}_{>0}, \times) \), and \( (\mathbb{Q}, +) \).

2.2.6. The general linear groups. The set of invertible \( n \times n \) matrices with real entries is denoted by \( \text{GL}(n, \mathbb{R}) \). It is a group under matrix multiplication: the product of two \( n \times n \) invertible matrices is an invertible matrix of size \( n \times n \); the identity matrix is the identity; the inverse of \( A \) is exactly the matrix you already know as its inverse; multiplication of matrices is associative. We call \( \text{GL}(n, \mathbb{R}) \) the general linear group (of size \( n \) over \( \mathbb{R} \)). If \( n \geq 2 \), \( \text{GL}(n, \mathbb{R}) \) is not abelian.

There are other general linear groups such that \( \text{GL}(n, \mathbb{Z}) \), the group of \( n \times n \) invertible matrices whose entries, and the entries of its inverse, are integers.

2.2.7. The special linear groups. We denote by \( \text{SL}(n, \mathbb{R}) \) the set of \( n \times n \) matrices having determinant 1. The identity matrix belongs to \( \text{SL}(n, \mathbb{R}) \). Since \( \det(AB) = (\det A)(\det B) \), a product of two matrices in \( \text{SL}(n, \mathbb{R}) \) is in \( \text{SL}(n, \mathbb{R}) \). Since \( \det(A^{-1}) = (\det A)^{-1} \), the inverse of a matrix in \( \text{SL}(n, \mathbb{R}) \) is in \( \text{SL}(n, \mathbb{R}) \). We call \( \text{SL}(n, \mathbb{R}) \) the special linear group.

2.3. Some finite groups. You have already met one of the most important classes of finite groups, the symmetric groups \( S_n \). Symmetric groups
are quite complicated. Indeed, they have been intensely studied for well over a century now and although an enormous amount is known about them there is still much we don’t know about them. You could devote a lifetime to their study. Many people have done just that. But let’s start here with some simple examples, some of which will be familiar.

2.3.1. The trivial group. The simplest group of all consists of only one element, \( G := \{e\} \) with \( ee = e \). We call it the trivial group.

2.3.2. The group \( \{\pm 1\} \). A familiar group with two elements is the set \( \mu_2 := \{1, -1\} \) with the usual multiplication.

2.3.3. The multiplication table of a group. At primary school you probably wrote out some multiplication tables.

The order of a group is the number of its elements. We write \(|G|\) for the number of elements in \( G \). A group is said to be finite if it has only finitely many elements.

For a finite group it is conceivable that we could write out its entire multiplication table. We do this for some small groups below. We adopt the following convention: the rows and columns are labeled by the elements of the group; we use the same order for the labeling of the rows and columns; the entry in row \( a \) and column \( b \) is the product \( ab \); the entry in row \( b \) and column \( a \) is \( ba \).

2.3.4. Patterns in the multiplication table. You should be aware of some “patterns” in the multiplication table.

If \( G \) is abelian, there is a symmetry of the table about the diagonal line of slope \(-1\). That symmetry expresses the fact that \( ab = ba \).

The identity must appear exactly once in each row and column because given \( x \) there is an element \( y \) such that \( xy = e = yx \). The identity \( e \) appears on the diagonal exactly when \( x^2 = e \), i.e., when an element is its own inverse.

In fact, every element of the group must appear in each row and column; for example, the entries in the row labelled by \( x \) are \( \{xg \mid g \in G\} \), and given \( y \in G \) there is a \( g \) such that \( xg = y \), namely \( g = x^{-1}y \), so \( y \) appears in the row labelled \( x \).

2.3.5. Positive and negative numbers. Here is a group whose elements are themselves sets. Let \( P \) and \( N \) denote the sets of positive and negative real numbers respectively. We define the group operation, which we call multiplication, by declaring that

\[
\begin{array}{c|cc}
\times & P & N \\
\hline
P & P & N \\
N & N & P
\end{array}
\]
The group operation says that

\[
\begin{align*}
\text{positive} \times \text{positive} &= \text{positive} & \text{negative} \times \text{negative} &= \text{positive} \\
\text{positive} \times \text{negative} &= \text{negative} & \text{negative} \times \text{positive} &= \text{negative}.
\end{align*}
\]

2.3.6. The parity group \( \mathbb{P} \). The elements in this group are themselves sets. Let \( E \) denote the set of even integers and \( O \) the set of odd integers. Young children know that

\[
\text{odd} + \text{odd} = \text{even}, \quad \text{even} + \text{odd} = \text{odd} + \text{even} = \text{even}.
\]

In other words, they know that the addition table for the parity group \( \mathbb{P} := \{E, O\} \) is

<table>
<thead>
<tr>
<th>+</th>
<th>( E )</th>
<th>( O )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E )</td>
<td>( E )</td>
<td>( O )</td>
</tr>
<tr>
<td>( O )</td>
<td>( O )</td>
<td>( E )</td>
</tr>
</tbody>
</table>

The identity in \( \mathbb{P} \) is \( E \) and \( O \) is its own inverse.

2.3.7. The group \( \mathbb{F} \). This group has two elements, labeled 0 and 1, and its addition table is defined to be

<table>
<thead>
<tr>
<th>+</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

2.3.8. Remark. The last four examples are essentially the same. This idea will be made precise when we introduce the notion of isomorphism. We will then be able to say that the last four groups are isomorphic to one another where the word isomorphism has a precise meaning.

2.3.9. Exercise: a group with three elements. There is a group with three elements. There is essentially only one group with three elements. When you begin to write out its multiplication table the axioms force your hand. It is a very good exercise to see why this is. For example, write 1 for the identity element and let \( x \) be some other element in your putative three-element group. Now \( xx \) is either 1, \( x \), or the third element in it. Consider all three possibilities—perhaps you can exclude one of them, or even two of them. Perhaps you should consider \( xxx \).

2.3.10. The 4th roots of unity. Let \( i \) denote a square root of \(-1\). Then the set of complex numbers

\[ \mu_4 := \{1, -1, i, -i\} \]

is a group under multiplication.
2. GROUPS AND EXAMPLES

\[ \begin{array}{cccc}
1 & i & -1 & -i \\
1 & 1 & i & -1 \\
i & i & 1 & -1 \\
-1 & -1 & -i & 1 \\
-1 & -i & 1 & -1 \\
\end{array} \]

2.3.11. *Exclusive OR.* Fix a set \( X \). It might help if you think of an explicit \( X \) such as the set of integers. Let \( G \) be the set of all finite subsets of \( X \) including the empty set. If \( A \) and \( B \) are in \( G \), i.e., subsets of \( X \), we define

\[ A \oplus B := \{ \text{elements that are in either } A \text{ or } B \text{ but not both} \}. \]

The identity element is the empty set \( \phi \). Every element is its own inverse because \( A \oplus A = \phi \). You can draw a Venn diagram involving three sets in order to convince yourself that the associative law \( A \oplus (B \oplus C) = (A \oplus B) \oplus C \) holds.

The “multiplication table” for the subsets of the set \( X = \{3\} \) is

\[
\begin{array}{c|ccc}
\oplus & \phi & \{3\} \\
\hline
\phi & \phi & \{3\} \\
\{3\} & \{3\} & \phi \\
\end{array}
\]

The “multiplication table” for the subsets of the set \( X = \{3, 8\} \) is

\[
\begin{array}{c|ccccc}
\oplus & \phi & \{3\} & \{8\} & \{3, 8\} \\
\hline
\phi & \phi & \{3\} & \{8\} & \{3, 8\} \\
\{3\} & \{3\} & \phi & \{3, 8\} & \{8\} \\
\{8\} & \{8\} & \{3, 8\} & \phi & \{3\} \\
\{3, 8\} & \{3, 8\} & \{8\} & \{3\} & \phi \\
\end{array}
\]

2.3.12. *Another group with four elements, \( \mathbb{F}^2 \).* The set

\[ \mathbb{F}^2 := \{00, 01, 10, 11\} \]

can be made into a group with addition table

\[
\begin{array}{c|cccc}
+ & 00 & 01 & 10 & 11 \\
00 & 00 & 01 & 10 & 11 \\
01 & 01 & 00 & 11 & 10 \\
10 & 10 & 11 & 00 & 01 \\
11 & 11 & 10 & 01 & 00 \\
\end{array}
\]
2.3.13. The group generated by the reflections in the $x$- and $y$-axes. The set consisting of the matrices

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad ST = TS = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

is a group under matrix multiplication. Note a product of two elements in $G$ is still in $G$. Each element is its own inverse. We know that multiplication of any matrices is associative so the associative law holds in $G$. Each of these matrices represents a linear transformation in the plane: $S$ is the reflection in the $x$-axis, $T$ is the reflection in the $y$-axis, and $ST$ is rotation by 180 degrees. This group is “essentially the same” as the previous example as you can see by comparing the multiplication/addition tables:

<table>
<thead>
<tr>
<th></th>
<th>$I$</th>
<th>$S$</th>
<th>$T$</th>
<th>$ST$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I$</td>
<td>$I$</td>
<td>$S$</td>
<td>$T$</td>
<td>$ST$</td>
</tr>
<tr>
<td>$S$</td>
<td>$S$</td>
<td>$I$</td>
<td>$ST$</td>
<td>$T$</td>
</tr>
<tr>
<td>$T$</td>
<td>$T$</td>
<td>$ST$</td>
<td>$I$</td>
<td>$S$</td>
</tr>
<tr>
<td>$ST$</td>
<td>$ST$</td>
<td>$T$</td>
<td>$S$</td>
<td>$I$</td>
</tr>
</tbody>
</table>

2.4. Multiplicative versus additive notation. In the above examples we used various notations for the group operations. The only hard and fast rule is that we never use the symbol + for the group operation in a non-abelian group. That would be too confusing because in all our prior experience + has been a commutative operation whether adding numbers, or matrices, or functions, etc. Thus if we speak about a group $G$ with no real knowledge of its nature we will always write $xy$ for the result of the group operation on the elements $x$ and $y$. We then say that we are using multiplicative notation for the group operation.

For example, we used multiplicative notation when we defined a group. And, when we define isomorphism below we will again use multiplicative notation.

2.4.1. Exponent Notation. If $x$ is an element of a group $G$ and $n$ a positive integer we write $x^n$ for the product of $x$ with itself $n$ times. For example, $x^2 = xx$, $x^3 = xxx$, and so on. The associative law ensures that $x^3$, for example, is unambiguous (because $x(xx) = (xx)x$). We also declare that $x^0 = 1$, the identity element in $G$. If $n$ is a positive integer, we write $x^n$ for $(x^{-1})^n$.

It is a good exercise to check that $x^m x^n = x^{m+n}$ for all integers $m$ and $n$, and that $(x^m)^n = x^{mn}$.

For a familiar group like $(\mathbb{Z}, +)$ it would be very confusing to use any symbol other than + for the addition in $\mathbb{Z}$. However, the exponent notation above is not sensible to use in a group where we use + for the group operation. It makes more sense to write $nx$ for $x + x + \cdots + x$, the $n$-fold sum of $x$ with itself. Now the “exponent rules” become $(n + m)x = nx + mx$. 

2.4.2. Notation for the identity. In the definition of a group we wrote \( e \) for its identity element. It is common to use the symbol 1 to denote the identity in a group where we are using multiplicative notation. It is also common to write 0 for the identity element in a group where we are using + for the group operation. However, when we do this you must not confuse 1 with the real number 1. We are simply overworking the symbol 1 by using it to denote two different things. Of course, you already have some experience in doing that with matrices. The \( m \times n \) matrices form a group under addition and we denote the identity in this group, the \( m \times n \) zero matrix, by 0. Thus, in matrix algebra the symbol 0 is very overworked. For example, you might even write \( 0 \times 0 = 0 \) where the first 0 is the \( 2 \times 3 \) zero matrix, the second 0 is the \( 3 \times 4 \) zero matrix, and the 0 on the right is the \( 2 \times 4 \) zero matrix.

2.5. Complex roots of unity. For each integer \( n \geq 1 \) we define

\[
\mu_n := \{ z \in \mathbb{C} \mid z^n = 1 \}.
\]

This is a group under multiplication: it contains 1 which serves as the identity in \( \mu_n \); if \( z \) is an \( n \text{th} \) root of unity, so is \( z^{-1} \); if \( z^n = w^n = 1 \), then \( (zw)^n = 1 \); and, of course, the multiplication is associative. We call \( \mu_n \) the group of \( n \text{th} \) roots of unity. We can explicitly list its elements

\[
\mu_n := \{e^{2\pi im/n} \mid 0 \leq m \leq n - 1 \}.
\]

Notice that \(|\mu_n| = n\).

Perhaps you have previously used a different notation for the \( n \text{th} \) roots of unity using the fact that

\[
e^{i\theta} = \cos \theta + i \sin \theta.
\]

2.6. The quaternion group. The quaternion group is a non-abelian group with eight elements. So far the smallest non-abelian group we have met is the symmetric group \( S_3 \) which has 6 elements. It turns out that \( S_3 \) is the only non-abelian group with < 8 elements.\(^3\)

The quaternion group is therefore the second smallest non-abelian group. Actually, there are two non-abelian groups with eight elements—the other one is the dihedral group \( D_4 \), the symmetry group of the square. We shall see later that there are no non-abelian groups with nine elements, exactly one non-abelian group with 10 elements, no non-abelian groups with eleven elements, and three non-abelian groups with twelve elements. Perhaps you already suspect a pattern—the more divisors an integer has the more non-abelian groups there are of that size. That is roughly true.

\(^3\)I should really say “the only non-abelian group up to isomorphism” but we haven’t met isomorphisms yet so I can’t say that. But even professional mathematicians often omit “up to isomorphism”, that being understood. So, \( S_3 \) really is the “only” non-abelian group with < 8 elements.
The elements of the quaternion group are usually denoted \( \pm 1, \pm i, \pm j, \) and \( \pm k \). We will denote the group by the letter \( Q \). Thus
\[
Q = \{ \pm 1, \pm i, \pm j, \pm k \}
\]
and the multiplication is given by
\[
(2.2) \quad i^2 = j^2 = k^2 = ijk = -1,
\]
and multiplication by \(-1\) is as you would expect. Oh, 1 is the identity. We can work out the complete multiplication table from (2.2) by simply requiring that the multiplication be associative. For example, from \( i^2 = -1 \) we see that \( i^{-1} = -i \). Similarly, \( j^{-1} = -j \) and \( k^{-1} = -k \). Now multiply both sides of the equality \( ijk = -1 \) by \( k^{-1} \) on the right to get \( ij = ijkk^{-1} = (-1)k^{-1} = (-1)(-k) = k \).

Eventually we get
\[
ij = k, \quad jk = i, \quad ki = j, \quad ji = -k, \quad kj = -i, \quad ik = -j.
\]

Notice that the subset \( \{ \pm 1, \pm i \} \) is the the group of 4th roots of unity that appeared first in section 2.3.10. So too are the subsets \( \{ \pm 1, \pm j \} \) and \( \{ \pm 1, \pm k \} \), except the notation differs from that in section.

2.6.1. The other non-abelian group with eight elements. Paul: \( \mathbb{F}^2 \times \mathbb{Z}_2 \)

\[ ... \]

3. The groups \( \mathbb{Z}/d \)

For each positive integer \( d \) there is a particularly important abelian group having \( d \) elements that is denoted by \( \mathbb{Z}/d \), or \( \mathbb{Z}_d \).

3.1. A convenient notation: addition of subsets of \( \mathbb{Z} \). If \( A \) and \( B \) are subsets of \( \mathbb{Z} \) we introduce the notation
\[
A + B := \{ a + b \mid a \in A, b \in B \}.
\]

We call it \emph{addition} but it does \emph{not} make the set of all subsets of \( \mathbb{Z} \) a group (why not?). Nevertheless, this addition is commutative, i.e., \( A + B = B + A \), associative, i.e., \( A + (B + C) = (A + B) + C \), and has an identity, \( \{0\} \).

If \( A \) consists of a single element, say \( A = \{a\} \), we write \( a + B \) rather than \( A + B \).

3.1.1. Warning. This notation for addition of subsets should not be confused with the notation for set difference. If \( A \) and \( B \) are subsets of a set \( X \) we write \( A - B \) for \( \{ a \in A \mid a \notin B \} \). Sorry about this, but we have to deal with the language as it is spoken. Before getting upset about this talk to a non-native English speaker about the words \emph{thought}, \emph{through}, \emph{rough}, and \emph{thorough}. Ugh! And ouch!
3.2. Definition of $\mathbb{Z}/d$. Let’s start with the case $d = 5$. We pronounce $\mathbb{Z}/5$ as zee mod five. Its elements are the following five subsets of $\mathbb{Z}$:

$$5\mathbb{Z} := \{5n \mid n \in \mathbb{Z}\},$$

$$1 + 5\mathbb{Z} := \{1 + 5n \mid n \in \mathbb{Z}\},$$

$$2 + 5\mathbb{Z} := \{2 + 5n \mid n \in \mathbb{Z}\},$$

$$3 + 5\mathbb{Z} := \{3 + 5n \mid n \in \mathbb{Z}\},$$

$$4 + 5\mathbb{Z} := \{4 + 5n \mid n \in \mathbb{Z}\}.$$

For brevity we will write $r’ = r + 5\mathbb{Z}$ for all integers $r$. Thus

$$\mathbb{Z}/5 := \{0’, 1’, 2’, 3’, 4’\}.$$

Note that $2’ = 2 + 5\mathbb{Z} = 7 + 5\mathbb{Z} = 7’$, and so on. Notice too, that if $0 \leq r \leq 4$, then $r’$ consists of the integers that leave a remainder of $r$ when divided by 5. To make $\mathbb{Z}/5$ a group we define an addition on it by declaring that

$$r’ + s’ = (r + s)’.$$

The addition is well-defined (why?) and the addition table is

<table>
<thead>
<tr>
<th></th>
<th>0’</th>
<th>1’</th>
<th>2’</th>
<th>3’</th>
<th>4’</th>
</tr>
</thead>
<tbody>
<tr>
<td>0’</td>
<td>0’</td>
<td>1’</td>
<td>2’</td>
<td>3’</td>
<td>4’</td>
</tr>
<tr>
<td>1’</td>
<td>1’</td>
<td>2’</td>
<td>3’</td>
<td>4’</td>
<td>0’</td>
</tr>
<tr>
<td>2’</td>
<td>2’</td>
<td>3’</td>
<td>4’</td>
<td>0’</td>
<td>1’</td>
</tr>
<tr>
<td>3’</td>
<td>3’</td>
<td>4’</td>
<td>0’</td>
<td>1’</td>
<td>2’</td>
</tr>
<tr>
<td>4’</td>
<td>4’</td>
<td>0’</td>
<td>1’</td>
<td>2’</td>
<td>3’</td>
</tr>
</tbody>
</table>

To verify that this addition rule makes $\mathbb{Z}/5$ a group you must check associativity, the existence of an identity, and existence of inverses. Associativity follows from the associative law $A + (B + C) = (A + B) + C$ that was mentioned above for subsets of $\mathbb{Z}$. It is clear that $0’$ is an identity in $\mathbb{Z}/5$. Inverses exist because $r’ + (5 - r)’ = 0’$. Hence $(\mathbb{Z}/5, +)$ is a group. It is an abelian group because $r’ + s’ = s’ + r’$, i.e., because the addition of subsets of $\mathbb{Z}$ is commutative.

In the preceding example no special properties of the number 5 were used. For every positive integer $d$ we define

$$\mathbb{Z}/d := \{r + d\mathbb{Z} \mid r \in \mathbb{Z}\}$$

and make $\mathbb{Z}/d$ a group by defining

$$(r + d\mathbb{Z}) + (s + d\mathbb{Z}) := (r + s) + d\mathbb{Z}.$$

**Proposition 2.9.** For every integer $d \geq 1$ there is an abelian group with $d$ elements, namely $\mathbb{Z}/d$.

**Proof.** Use the same arguments as those we used above to show that the addition on $\mathbb{Z}/5$ made it a group. \qed
3.2.1. Warning: It is very important to notice that \( r_1 + d\mathbb{Z} = r_2 + d\mathbb{Z} \) if and only if \( r_1 - r_2 \) is a multiple of \( d \). Thus each element of \( \mathbb{Z}/d \) has infinitely many different names. It is for this reason that we must take care to check the addition in \( \mathbb{Z}/d \) is well-defined. For example, if \( r_1 + d\mathbb{Z} = r_2 + d\mathbb{Z} \) the addition in \( \mathbb{Z}/d \) must be such that \( (r_1 + d\mathbb{Z}) + (s + d\mathbb{Z}) = (r_2 + d\mathbb{Z}) + (s + d\mathbb{Z}) \).

4. Isomorphisms

4.1. A preparatory result. Our next result is an important result about group homomorphisms and their kernels but we don’t want to introduce that terminology yet so we will state the result without any fancy language. In any case, we will use parts of it frequently once we begin our discussion of isomorphisms.

We will use the notation \( e_G \) for the identity element in a group \( G \) when we need to distinguish it from the identity element in a group \( H \).

**Lemma 2.10.** Let \( f : G \rightarrow H \) be a function between two groups such that
\[ f(xy) = f(x)f(y) \]
for all \( x, y \in G \). Define \( K := \{ g \in G \mid f(g) = e_H \} \). Then

1. \( e_G \in K \);
2. \( f(x^{-1}) = f(x)^{-1} \) for all \( x \in G \);
3. the multiplication in \( G \) makes \( K \) a group;
4. \( f \) is injective if and only if \( K = \{ e_G \} \).

**Proof.** (1) The hypothesis on \( f \) implies that
\[ e_H f(e_G) = f(e_G) = f(e_G e_G) = f(e_G)f(e_G) . \]
However, we can cancel in \( H \) so \( e_H = f(e_G) \).

(2) Let \( x \in G \). Then \( f(x)f(x^{-1}) = f(xx^{-1}) = f(e_G) = e_H \). Similarly,
\[ f(x^{-1})f(x) = f(x^{-1}x) = f(e_G). \]
Thus \( f(x^{-1}) = f(x)^{-1} \).

(3) The following five observations show that \( K \) is a group.

(a) By (1), \( e_G \in K \), so \( K \) is not empty.

(b) If \( x, y \in K \), then \( f(xy) = f(x)f(y) = e_H e_H = e_H \) so \( xy \in K \).

Thus, when restricted to elements of \( K \), the multiplication in \( G \) gives a composition \( K \times K \rightarrow K \).

(c) That composition is associative because it is associative for all elements in \( G \).

(d) If \( x \in K \), then \( x e_G = x = e_G x \) so \( K \) contains an identity element, namely \( e_G \).

(e) If \( g \in K \), then \( f(g^{-1}) = f(y)^{-1} = e_H^{-1} = e_H \) so \( g^{-1} \in K \). Hence every element in \( K \) has an inverse in \( K \).

(4) Certainly, if \( f \) is injective \( K = \{ e_G \} \). To prove the converse suppose that \( K = \{ e_G \} \). If \( f(x) = f(y) \), then
\[ f(xy^{-1}) = f(x)f(y^{-1}) = f(y)f(y^{-1}) = f(yy^{-1}) = f(e_G) = e_H \]
so \( xy^{-1} \in K = \{ e_G \} \), i.e., \( x = y \). Hence \( f \) is injective.
4.2. **When are two groups the same?** The notion of isomorphism formalizes the notion of sameness for groups. Two groups are “essentially the same” if and only if they are isomorphic (see below).

The multiplication tables of the groups $\mu_2 = \{1, -1\}$, $P = \{\{E, O\}, +\}$, and $F = \{\{0, 1\}, +\}$, are the “same” in the sense that the only difference is one of labeling. For example, if we replace every 1 in the multiplication table of $\mu_2$ by $E$ and every $-1$ in the multiplication table of $\mu_2$ by $O$ we get the addition table for $P$. A similar statement is true with $F$ in place of $P$, 0 in place of $E$, and 1 in place of $O$.

The three sets $\mu_2$, $P$, and $F$, are different because their elements are different. Thus if someone asks “are any two of these groups the same” we are bound to reply “no, their elements are different.” However, if we focus on the group operations and compare their multiplication tables they are essentially the “same”. To see this we write their multiplication tables side-by-side and compare them.

It is the algebraic structure of a group that is important rather than the individual nature of its elements. Group theory is not about how we label the elements of a group. It is about the behavior of the group operation, the algebraic features of the group.

**Definition 2.11.** Let $G$ and $H$ be groups. A bijection $f : G \to H$ is called an isomorphism if

$$f(xy) = f(x)f(y)$$

for all $x, y \in G$. In this case we say $G$ and $H$ are isomorphic groups and use the notation $G \cong H$ to denote this fact.

Every group is isomorphic to itself: the identity map $\text{id}_G : G \to G$ is an isomorphism.

**Proposition 2.12.** Let $G$ and $H$ be groups and suppose $f : G \to H$ is an isomorphism. Then

1. $f(e_G) = e_H$;
2. $f(x^{-1}) = f(x)^{-1}$ for all $x \in G$;
3. $f^{-1} : H \to G$ is an isomorphism;
4. If $g : H \to K$ is an isomorphism, then $gf$ is an isomorphism.

**Proof.** Parts (1) and (2) were already proved in Lemma 2.10.

(3) Let $a, b \in H$. Then there are elements $x, y \in G$ such that $f(x) = a$ and $f(y) = b$. By definition of $f^{-1}$, $f^{-1}(a) = x$ and $f^{-1}(b) = y$. Because $f$ is an isomorphism, $f(xy) = f(x)f(y) = ab$ so

$$f^{-1}(ab) = xy = f^{-1}(a)f^{-1}(b)$$

which shows that $f^{-1}$ is an isomorphism.

(4) Since $f$ and $g$ are bijective so is $gf$. If $x, y \in G$, then

$$gf(xy) = g(f(x)f(y)) = gf(x)gf(y)$$

so $gf$ is an isomorphism. □
The point is this. A group as a set with additional structure. That additional structure is the product rule which satisfies the conditions in the definition of a group. The requirement that \( f(xy) = f(x)f(y) \) is saying, in effect, that the bijection \( f \) is “compatible” with the additional structure, namely the products.

4.2.1. Remark. We used multiplicative notation when we defined “isomorphism” and when we stated and proved Proposition 2.12. However, often we will have isomorphisms between groups \( G \) and \( H \) in which the group laws are written additively. In that case the condition \( f(xy) = f(x)f(y) \) must be replaced by \( f(x + y) = f(x) + f(y) \). Here’s a simple example of that.

4.2.2. An example: \( \mathbb{Z} \cong \mathbb{Z}_d \). Let \( d \) be a non-zero integer. We write \( \mathbb{Z}_d \) for the set of all multiples of \( d \); i.e., \( \mathbb{Z}_d = \{ nd \mid n \in \mathbb{Z} \} \). Then \( (\mathbb{Z}_d, +) \) is a group: it is non-empty; a sum of two multiples of \( d \) is a multiple of \( d \) so ordinary addition is a binary operation \( \mathbb{Z}_d \times \mathbb{Z}_d \to \mathbb{Z}_d \); addition is certainly associative; 0 is a multiple of \( d \) so is in \( \mathbb{Z}_d \) and is the identity in \( \mathbb{Z}_d \); and every element in \( \mathbb{Z}_d \) has an inverse in \( \mathbb{Z}_d \), namely its negative. Now we know \( (\mathbb{Z}_d, +) \) is a group, I claim that the function

\[
    f : \mathbb{Z} \to \mathbb{Z}_d, \quad f(n) := nd,
\]

is an isomorphism. Certainly it is bijective; its inverse being the function that sends an element \( x \in \mathbb{Z}_d \) to \( x/d \). Furthermore,

\[
    f(m + n) = (m + n)d = md + nd = f(m) + f(n)
\]

so \( f \) is an isomorphism, and we may therefore write \( \mathbb{Z} \cong \mathbb{Z}_d \).

There was nothing special about \( d \) in what we just did. If \( r \) is any non-zero real number that set of all \emph{integer} multiples of \( r \) is a group under addition and is isomorphic to \( \mathbb{Z} \). Fill in the details if this is not immediately obvious,

4.2.3. Remark. Sometimes we have an isomorphism between groups where the group operation is written additively in one and multiplicatively in the other. In that case one has to change the condition \( f(xy) = f(x)f(y) \) in the appropriate way. We will now give a simple example of this phenomenon. This example can be considered a warm-up for the much more important example that appears in section 4.4 below.

4.2.4. Doubly infinite geometric progressions. Let \( a \in \mathbb{R} - \{-1, 0, 1\} \). Then

\[
    \mathbf{A} := \{ a^n \mid n \in \mathbb{Z} \}
\]

is a doubly infinite geometric progression: it consists of the numbers

\[
    \ldots a^{-2}, a^{-1}, 1, a, a^2, \ldots.
\]

The set \( \mathbf{A} \) is a group under multiplication. Of course, 1 is the identity, the inverse of \( a^n \) is \( a^{-n} \), and multiplication is associative. I claim that the function

\[
    f : \mathbb{Z} \to \mathbf{A}, \quad f(n) := a^n,
\]
is an isomorphism of groups. Because $a$ is not $-1$, 0, or $-1$, all the powers of $a$ are different. Hence $f$ is injective. It is obviously surjective and therefore bijective. Since

$$f(m + n) = a^{m+n} = a^m a^n = f(m)f(n)$$

for all $m, n \in \mathbb{Z}$, $f$ is an isomorphism. Notice that $f^{-1}$, which is also an isomorphism, satisfies $f^{-1}(xy) = f^{-1}(x) + f^{-1}(y)$.

4.2.5. *Groups with four elements.* We have seen several groups having four elements. Are they all isomorphic to one another or not? For example, the group $\mathbb{F}_2$ in section 2.3.12 is isomorphic to the group in section 2.3.13 with an isomorphism $f : \mathbb{F}_2 \rightarrow \{I, S, T, ST\}$ being given by the function

$$f(00) := I, \quad f(01) := S, \quad f(10) := T, \quad f(11) := ST.$$

Certainly $f$ is bijective but a little more care is required to see that

$$f(x + y) = f(x)f(y)$$

for all $x, y \in \mathbb{F}_2$. Although the group operation in $\mathbb{F}_2$ is denoted by $+$ and the operation in the other group is written multiplicatively this has no bearing on the question of isomorphism. Both groups are abelian but it is natural to use $+$ for the group operation in $\mathbb{F}_2$ and $\times$ for the group operation in $\{I, S, T, ST\}$.

The above $f$ is not the only isomorphism from $\mathbb{F}_2$ to $\{I, S, T, ST\}$. The function $g : \mathbb{F}_2 \rightarrow \{I, S, T, ST\}$ given by

$$g(00) := I, \quad g(01) := T, \quad g(10) := ST, \quad g(11) := S.$$

There are two more isomorphisms from $\mathbb{F}_2$ to $\{I, S, T, ST\}$. Can you find them?

Is $\mathbb{F}_2$ isomorphic to $\mu_4$? The answer is *no*. The best way to see this is to look for some algebraic feature that one of the groups has but the other does not have. For example, every element $x$ in $\mathbb{F}_2$ has the property that $x + x$ is equal to the identity, but in $\mu_2$ we have $i^2 \neq 1$. (The fact that the group operations are written differently has no relevance.) Alternatively, there is an element $\xi$ in $\mu_4$ such that every element in $\mu_4$ is a power of $\xi$, i.e., $\mu_4 = \{\xi, \xi^2, \xi^3, \xi^4\}$. But there is no element $x$ in $\mathbb{F}_2$ such that $\{x, x + x, x + x + x, x + x + x + x\}$ is equal to $\mathbb{F}_2$.

We must turn these observations into a rigorous proof, but that is easy—the harder part is to make the observations in the last paragraph.

We use a proof by contradiction to prove that $\mathbb{F}_2 \not\cong \mu_4$. Suppose to the contrary that $f : \mathbb{F}_2 \rightarrow \mu_4$ is an isomorphism. Then for all $x \in \mathbb{F}_2$ we would have $f(x + x) = f(00) = 1$ (remember that an isomorphism sends the identity to the identity) and therefore $f(x)f(x) = 1$. But $f$ is surjective so we would have $\xi^2 = 1$ for all $\xi \in \mu_4$ and that is not the case. Hence no such $f$ can exist.

Thus we have two different, meaning *non-isomorphic*, groups of order 4. Are there any others? No. If $G$ is a group with 4 elements it is isomorphic to $\mathbb{F}_2$ or $\mu_4$. It is a good exercise to try proving that.
4.3. Groups having a prime number of elements. We have just seen there are (at least) two different groups with four elements. In sharp contrast, if \( p \) is a prime number all groups with \( p \) elements are isomorphic to one another.

**Proposition 2.13.** Let \( p \) be a positive prime number. Then there is only one group with \( p \) elements, i.e., if \( G \) and \( H \) are groups having \( p \) elements, then \( G \cong H \).

**Proof.** Let \( G \) be any group with \( p \) elements. Let \( e \) denote the identity element in \( G \) and fix an element \( x \) in \( G \) that is not \( e \).

Let \( \mu_p \) be the group of complex \( p^\text{th} \) roots of unity. Write \( \xi = e^{2\pi i/p} \). Then \( \mu_p = \{1, \xi, \xi^2, \ldots, \xi^{p-1}\} \). Define \( f : \mu_p \to G \) by

\[
f(\xi^k) = x^k.
\]

Then \( f(\xi^k \xi^\ell) = f(\xi^{k+\ell}) = x^{k+\ell} = x^k x^\ell = f(\xi^k) f(\xi^\ell) \).

Let \( K = \{\xi^n \mid f(\xi^n) = e\} \). Then \( K \) is a subgroup of \( \mu_p \) by Lemma 2.10. Suppose \( K \) contains some element other than 1, say \( \xi^n \) with \( 1 \leq n \leq p - 1 \). Since \( \gcd(n, p) = 1 \), there are integers \( a \) and \( b \) such that \( an + bp = 1 \). Because \( K \) is a group it contains

\[
(\xi^n)^a = \xi^{na} = \xi^{1-bp} = \xi.
\]

But that is absurd because \( f(\xi) = x \neq e \). We conclude that \( K \) must equal \{1\}. It now follows from Lemma 2.10 that \( f \) is injective. But \( G \) and \( \mu_p \) have the same number of elements so \( f \) is bijective and therefore an isomorphism.

We have shown that \( G \cong \mu_p \). If \( H \) is another group with \( p \) elements then \( H \cong \mu_p \). Hence \( H \cong G \).

The group of \( p^\text{th} \) complex roots of unity is not the only group with \( p \) elements that we have encountered. The group \( \mathbb{Z}/p \) also has \( p \) elements. Hence \( \mathbb{Z}/p \cong \mu_p \) and we could restate Proposition 2.13 in the following way.

**Corollary 2.14.** Let \( p \) be a positive prime. If \( G \) is a group with \( p \) elements, then

\[
G \cong \mathbb{Z}/p.
\]

4.4. Logarithms. The discovery that the groups \((\mathbb{R}_{>0}, \times)\) and \((\mathbb{R}, +)\) are isomorphic had a large impact on science. There are many isomorphisms between. Fix a real number \( b > 1 \). The functions

\[
f : \mathbb{R} \to \mathbb{R}_{>0} \quad f(x) := b^x \\
g : \mathbb{R}_{>0} \to \mathbb{R} \quad g(x) = \log_b x
\]

are mutually inverse isomorphisms.

The historical significance of this isomorphism lies in the fact that it simplified the task of calculation. For example, if one has to multiply two positive real numbers \( x \) and \( y \), one uses the fact that

\[
xy = fg(xy) = f(g(x) + g(y)) = f(\log_b x + \log_b y)
\]
Large tables giving the values of \( f(x) \) and \( g(x) \) were published so one could simply look up the values of \( g(x) \) and \( g(y) \), add them, then look up the value of \( f(g(x) + g(y)) \). Similarly, if one wanted to compute \( x^r \), one used the fact that

\[
x^r = f g(x^r) = f(\log(x^r)) = f(r \log x).
\]

The discovery of the method of logarithms, i.e., the discovery of the isomorphisms \( f \) and \( g \), is usually attributed to the Scotsman John Napier, the 8th Laird of Merchiston (1550-1617). The method was first revealed to the public with the publication of Napier’s book *Mirifici Logarithmorum Canonis Descriptio* in 1614. It was written in Latin, the language of science at the time, and contained 57 pages of explanation and 90 pages of tables of logarithms. Napier did not use a base as we now understand it, but his logarithms were, up to a scaling factor, effectively to base \( 1/e \).

Henry Briggs (1561-1630), Professor of Mathematics at Gresham College, London, from 1596 to 1619, and from 1619 Savilian professor of geometry at Oxford, visited Scotland in 1615 and 1617 seeking Napier’s permission to publish a table of common logarithms, i.e., logarithms to the base 10. The first installment of Briggs’ table of common logarithms, containing a brief account of logarithms and a long table of the logarithms of all integers below 1000 to 8 decimal places, was published in 1617 under the title *Logarithmorum Chilias Prima*.

In 1624, Briggs published *Arithmetica Logarithmica*, containing the logarithms of all integers from 1 to 20,000 and from 90,000 to 100,000 to fourteen places of decimals, together with an introduction in which the theory and use of logarithms are fully developed. The interval from 20,000 to 90,000 was filled by Adriaan Vlacq, a Dutch mathematician, but in his table, which appeared in 1628, the logarithms were given to only ten places of decimals. Vlacq’s table was later found to contain 603 errors, but considering that the table was the result of original hand-calculation, and contained more than 2,100,000 digits, the number of errors is remarkably small. An edition of Vlacq’s work, containing many corrections, was issued at Leipzig in 1794 under the title *Thesaurus Logarithmorum Completus* by Jurij Vega.

The great astronomer, Johannes Kepler (1571-1630) was an enthusiastic supporter of Napier’s work and in 1624 published a clear explanation of how they worked. Napier had not done that and many had been reluctant to use Napier’s logarithms prior to Kepler’s explanation. At that time the main task of astronomers was to produce tables of astronomical data, in large part as an aid to astrology. The production of these tables involved huge amounts of computation so Kepler’s enthusiasm is understandable.

Napier formed the word logarithm to mean a number that indicates a ratio: the Greek word logos meant proportion, and arithmos meant number. Napier chose that because the difference of two logarithms determines the ratio of the numbers they represent, so that an arithmetic series of logarithms corresponds to a geometric series of numbers. The term antilogarithm was
introduced in the late 17th century and persisted in collections of tables until they fell into disuse around the 1970s.

Jost Bürgi (1552-1632), a Swiss clockmaker and maker of astronomical instruments, invented logarithms independently of John Napier. There is evidence that Bürgi invented the method of logarithms as early as 1588, six years before Napier began work on the same idea. By delaying the publication of his work to 1620, and even then publishing only after repeated requests from Johannes Kepler, Bürgi lost his claim for priority.\footnote{In 1615, Kepler’s mother, Katharina Kepler, was accused of witchcraft by a prostitute. European witch hunting was at its peak during Kepler’s career, and was supported by all levels of society, including secular officials and intellectuals in universities. Kepler spent several years making legal appeals and hiding his mother from legal authorities seeking to torture her into confessing to witchcraft. Examining an accused witch ad torturam was a standard court procedure during this era. In 1620, under court order, Kepler’s mother was kidnapped in the middle of the night from her daughter’s home and taken to prison. Kepler spent the next year appealing to the Duke of Württemberg to prevent his imprisoned mother from being examined ad torturam. However, on September 28, 1621 Frau Kepler was taken from her prison cell into the torture room, shown the instruments of torture and ordered to confess. She replied “Do with me what you want. Even if you were to pull one vein after another out of my body, I would have nothing to admit,” and said the Lord’s Prayer. She was taken back to prison and freed on October 4 upon order of the duke, who ruled that her refusal to confess under threat of torture proved her innocence. He also ordered her accusers to pay the cost of her trial and imprisonment. After having spent most of the last seven years under the legal threat of imminent torture, Katharina Kepler died on April 13, still being threatened with violence from those who insisted she was a witch.}

4.5. The circle group. Let $U(1)$ denote the set of complex numbers of length 1. If $z$ and $w$ are complex numbers, then $|zw| = |z||w|$ so the usual multiplication of complex numbers is an associative binary operation $U(1) \times U(1) \rightarrow U(1)$. The number 1 has length one so belongs to $U(1)$ and is its identity element. If $z$ is a non-zero complex number, then $|z^{-1}| = |z|^{-1}$ so every element in $U(1)$ has an inverse in $U(1)$. Hence $U(1)$ is a group.

We call $U(1)$ the circle group because if you think of the complex numbers as points in the euclidean plane $\mathbb{R}^2$ the complex numbers of length one form a circle of radius one centered at 0.

4.6. Rotations in the plane. Given an angle $\theta$ we write $T_\theta$ for the linear transformation of the plane $\mathbb{R}^2$ that rotates a point by an angle of $\theta$ radians in the counterclockwise direction about the origin. Notice that $T_\theta = T_{\theta + 2\pi n}$ for all $n \in \mathbb{Z}$, so each rotation can be labelled in infinitely many different ways.

It is obvious that $T_\theta T_\psi = T_{\theta + \psi}$ and that rotation in the counterclockwise direction by an angle of $-\theta$ is the same as rotation in the clockwise direction by an angle of $\theta$ so $T_\theta$ has an inverse, namely $T_{-\theta}$. It follows at once that the set of all rotations is a group. We call it the rotation group in $\mathbb{R}^2$ and denote it by $SO(2)$. Actually, in keeping with its importance, it goes by a grander name, the special orthogonal group. As the notation suggests
there are special orthogonal groups $SO(n)$ for all integers $n \geq 0$ but for now we confine our attention to $SO(2)$. The identity element is $T_0$, the rotation through zero degrees. In fact, $T_{2n\pi} = \text{id}_{\mathbb{R}^2}$ for all $n \in \mathbb{Z}$, and $T_{(2n+1)\pi} = -\text{id}_{\mathbb{R}^2}$ for all $n \in \mathbb{Z}$.

Let’s use column vectors $\mathbf{z} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ to label the points in the plane $\mathbb{R}^2$ and write $A_\theta$ for the unique $2 \times 2$ matrix such that $T_\theta(\mathbf{z}) = A_\theta \mathbf{z}$ for all $\mathbf{z} \in \mathbb{R}^2$. If we write our matrix with respect to the ordered basis

$$
\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}
$$

the first column of $A_\theta$ is given by $T_\theta(\mathbf{e}_1)$ and the second column of $A_\theta$ is given by $T_\theta(\mathbf{e}_2)$. Elementary trigonometry (draw the diagrams and check!) then gives

$$
(2.3) \quad A_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.
$$

There are many interesting aspects to this:

- if you have forgotten your formulas for $\sin(\theta + \psi)$ and $\cos(\theta + \psi)$ you can recover them by using the fact that $A_\theta A_\psi = A_{\theta + \psi}$ and computing the product on the left;
- the determinant of $A_\theta$ is 1 because $\cos^2 \theta + \sin^2 \theta = 1$;
- $A_\theta^{-1} = A_{-\theta}$ and using the formula for the inverse of a $2 \times 2$ matrix in you can recover the formulae for $\sin(-\theta)$ and $\cos(-\theta)$ if you have forgotten them.

4.7. The isomorphism between $U(1)$ and $SO(2)$. Define $f : SO(2) \to U(1)$ by

$$
f(T_\theta) := e^{i\theta}.
$$

This is a well-defined function: although $T_\theta = T_{\theta + 2n\pi}$, $f(T_\theta) = f(T_{\theta + 2n\pi})$ because $e^{2n\pi i} = 1$. Furthermore, $f$ is bijective: it is surjective because if $z = a + ib \in U(1) - \{ \pm 1 \}$, then $z = e^{i\theta}$ where

$$
\theta = \tan^{-1} \left( \frac{b}{a} \right)
$$

and $i = e^{i\pi/2}$ and $-i = e^{-i\pi/2}$; it is injective because if $f(T_\theta) = f(T_\psi)$, then $e^{i\theta} = e^{i\psi}$ which implies $e^{i(\theta - \psi)} = 1$ and $\theta - \psi = 2n\pi$ for some integer $n$, whence $T_\theta = T_\psi$.

Finally, since

$$
f(T_\theta)f(T_\psi) = e^{i\theta}e^{i\psi} = e^{i(\theta + \psi)} = f(T_{\theta + \psi}) = f(T_\theta T_\psi)
$$

we conclude that $f$ is an isomorphism.
Suppose $G$ and $H$ are isomorphic groups. Then

1. $|G| = |H|$;
2. $G$ is abelian if and only if $H$ is abelian;
3. 

4.8.1. What does isomorphism mean? Two groups are isomorphic if and only if all their “essential” group-theoretic, or algebraic, features are the same. The word essential is vague, but here is an example of a difference that is not essential: let $\mu_2 = \{1, -1\}$ and let $P = \{E, O\}$ be the parity group. The groups $\mu_2$ and $P$ are different because we use different labels for their elements and different symbols for the group operation ($\cdot$ for $\mu_2$ and addition $+$ for $P$). However, in terms of their group-theoretic properties they are the same: they each have two elements and one other element that is its own inverse.

There are some obvious questions you might ask. For example, are any two groups with the same numbers of elements isomorphic? No; for example, $\mathbb{Z}_6$ and $S_3$ have six elements but are not isomorphic because one is abelian and the other is not. We have already seen that $\mu_4$ and $\mathbb{F}^2$ both have order 4 but are not isomorphic because every element in $\mathbb{F}^2$ is its own inverse but $\mu_4$ does not have that property.

4.8.2. Joke. An exam question describes two groups and then asks “Are these two groups isomorphic?” The student answers “The first one is but the second one isn’t.” :)

5. Subgroups

5.1. Definition. Let $G$ be a group. A subset $H \subset G$ is a subgroup of $G$ if the multiplication on $G$ makes $H$ a group.

Note that $G$ itself is a subgroup of $G$ and so is $\{e\}$. These are the boring subgroups.

To show that a subset $H$ of $G$ is a subgroup we must show $H$ satisfies axioms (1), (2), and (3), of Definition 2.7. Even before checking those we must check condition (0), that $xy$ is in $H$ whenever $x$ and $y$ are. If $H$ passes that test we say $H$ is closed under multiplication.

Oh, we must check $H$ is non-empty; let’s assume it is.

Condition (1), associativity of multiplication, will be satisfied by elements $x, y, z \in H$ because it is already satisfied for all $x, y, z \in G$. So we do not need to check condition (1).

Condition (2) says $H$ must have an identity. Could $H$ have an identity that is different from $e$, the identity of $G$? If $e' \in H$ is such that $e'x = x$ for every one element $x \in H$, then $e'x = ex$; but we can cancel in $G$ so $e' = e$. So, the identity in $H$ has to be the identity in $G$. Thus $H$ must contain $e$.

To see if condition (3) holds for $H$, the uniqueness of inverses tells us that $x^{-1}$ must belong to $H$ whenever $x$ does. All this proves the next result.
5. SUBGROUPS

Proposition 2.15. A subset $H$ of $G$ is a subgroup if and only

1. $H$ contains the identity of $G$, and
2. $xy$ belongs to $H$ whenever $x$ and $y$ do, and
3. $x^{-1}$ belongs to $H$ whenever $x$ does.

A shorter characterization of subgroups is given by the next result.

Proposition 2.16. Let $H$ be a non-empty subset of a group $G$. Then $H$ is a subgroup if and only $xy^{-1}$ belongs to $H$ whenever $x$ and $y$ do.

Proof. ($\Rightarrow$) This is trivial.

($\Leftarrow$) By hypothesis, $H$ is non-empty so contains an element, $h$ say. It also contains $hh^{-1}$, the identity. Thus, if $h \in H$, so is $eh^{-1} = h^{-1}$, so $H$ contains the inverse of every element in it. Finally, if $u, v \in H$, then $v^{-1} \in H$ so $H$ contains $u(v^{-1})^{-1} = uv$. Hence $H$ is a subgroup.

5.2. Examples of subgroups. The group of integers $(\mathbb{Z}, +)$ is a subgroup of the group of rational numbers $(\mathbb{Q}, +)$ which is a subgroup of the group of real numbers $(\mathbb{R}, +)$ which is a subgroup of the complex numbers $(\mathbb{C}, +)$.

The multiplicative group of non-zero rational numbers $(\mathbb{Q} - \{0\}, \times)$ is a subgroup of the group of non-zero real numbers $(\mathbb{R} - \{0\}, \times)$ which is a subgroup of the non-zero complex numbers $(\mathbb{C} - \{0\}, \times)$.

The group of positive real numbers $(\mathbb{R}_{>0}, \times)$ is a subgroup of $(\mathbb{R}-\{0\}, \times)$. The set of negative real numbers is not a subgroup of $(\mathbb{R} - \{0\}, \times)$ (why?).

Let $d$ be an integer. Because the difference of two multiples of $d$ is a multiple of $d$ Proposition 2.16 tells us that $d\mathbb{Z} := \{dn \mid n \in \mathbb{Z}\}$ is a subgroup of $\mathbb{Z}$.

Proposition 2.17. The subgroups of $\mathbb{Z}$ are the subsets $d\mathbb{Z}$, $d \geq 0$.

Proof. Let $H$ be a subgroup of $\mathbb{Z}$. Since $\{0\} = 0\mathbb{Z}$ we may assume $H \neq \{0\}$.

If $h$ is in $H$ so is $-h$. Hence there is a smallest positive integer in $H$.

Let $d$ be that integer. Because $H$ is a subgroup, every multiple of $d$ belongs to $H$. Thus $d\mathbb{Z} \subset H$.

Now let $h$ be any element in $H$. Then $h = qa + r$ for some integers $q$ and $r$ with $0 \leq r < d$. Since $r = h - qa$, $r$ belongs to $H$. By choice of $d$ it follows that $r = 0$, whence $h \in d\mathbb{Z}$. Thus $d\mathbb{Z} = H$. □

5.2.1. Exercise: If $H$ is a subgroup of $G$ and $f : G \rightarrow G'$ an isomorphism show that $f(H)$ is a subgroup of $G'$ and that the restriction of $f$ to $H$ is an isomorphism from $H$ to $f(H)$.

Proposition 2.18. if $H$ and $K$ are subgroups of $G$ so is $H \cap K$.

Proof. Certainly $H \cap K$ is non-empty because both $H$ and $K$ contain the identity element. Suppose that $x$ and $y$ belong to $H \cap K$. Then $xy^{-1} \in H$ because $H$ is a subgroup and $xy^{-1} \in K$ because $K$ is a subgroup. Thus
2. GROUPS

$xy^{-1} \in H \cap K$ and it now follows from the previous result that $H \cap K$ is a subgroup of $G$. □

The next result explains, in part, the importance of the symmetric groups—they are ubiquitous.

**Theorem 2.19 (Cayley’s Theorem).** If $G$ is a group with $n$ elements, then $G$ is isomorphic to a subgroup of the symmetric group $S_n$.

**Proof.** The idea of the proof is simple: if $x \in G$, then the function $\lambda_x : G \to G$ defined by $\lambda_x(y) := xy$ is a permutation of $G$. Write $P$ for the group of permutations of $G$. Then $P$ is a group isomorphic to $S_n$ so it is enough to show that $G$ is isomorphic to a subgroup of $P$.

Define $f : G \to P$ by $f(x) := \lambda_x$. First we check that $\lambda_x$ really is in $P$: if $y \neq y'$, then $xy \neq xy'$ so $\lambda_x(y) \neq \lambda_x(y')$ which implies that $\lambda_x$ is injective; if $g \in G$, then $g = \lambda_x(x^{-1}g)$ so $\lambda_x$ is surjective; thus $\lambda_x$ is a permutation of $G$. It is clear that $f(x)f(y) = f(xy)$ is the function that sends $g$ to $\lambda_x \lambda_y(g) = xyg = \lambda_{xy}(g)$ so $f(x)f(y) = f(xy)$.

Certainly $f(G)$ is a subgroup of $P$ because $f(1) = \text{id}_G$; $f(x)f(y) = f(xy)$; and $f(x)^{-1} = f(x^{-1})$.

It is clear that $f$ is injective because if $x \neq y$, $f(x)(1) = x$ whereas $f(y)(1) = y$ so $f(x) \neq f(y)$. Thus, considered as map from $G$ to $f(G)$, which is a subgroup of $P$, $f$ is bijective and hence an isomorphism. □

### 5.3. Finite subgroups of $\text{U}(1)$ and $\text{SO}(2)$

The circle group $\text{U}(1)$ consists of the complex numbers of length 1. If $n$ is a positive integer, then the group of $n^{\text{th}}$ roots of unity, $\mu_n$, is a subgroup of $\text{U}(1)$. Furthermore, $\mu_m$ is a subgroup of $\mu_n$ whenever $m$ divides $n$.

It is easy to check, and you should do so, that an isomorphism $f : G \to H$ sends a subgroup $K$ of $G$ to the subgroup $f(K)$ subgroup of $H$. Furthermore, $f(K) \cong K$.

Because $\text{U}(1) \cong \text{SO}(2)$ the subgroups $\mu_n$ of $\text{U}(1)$ correspond to subgroups of $\text{SO}(2)$. More explicitly, the isomorphism $f : \text{SO}(2) \to \text{U}(1)$ given by

$$f \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = e^{i\theta} = \cos \theta + i\sin \theta$$

sends the rotation $T_{2\pi r/n}$ to the $n^{\text{th}}$ root of unity

$$e^{2\pi ir/n} = \cos \left(\frac{2\pi ir}{n}\right) + i \sin \left(\frac{2\pi ir}{n}\right).$$

If you have a good understanding of the complex numbers as points in the complex plane and realize that multiplying a complex number $z$ by $\cos \theta + i\sin \theta$ just rotates $z$ by an angle $\theta$ in the counter-clockwise direction all this will be “obvious”.
5. SUBGROUPS

5.4. The center of $G$. The center of $G$ is the subgroup

$$Z(G) := \{z \in G \mid zg = gz \text{ for all } g \in G\}.$$ 

In German, the word for center is *zentrum*. That’s why we use the notation $Z(G)$. Germany was a hotbed of group theory in the late 1800s.

The center of $G$ is a subgroup of $G$, a particularly important one.

In fact, it is a special case of the following more general result which we will prove after introducing a generalization of the center. The centralizer of a subset $S$ of $G$ is defined to be

$$C_G(S) := \{z \in G \mid zg = gz \text{ for all } g \in S\}.$$ 

In other words, $C_G(S)$ consists of the elements in $G$ that commute with all elements in $S$. Hence $Z(G) = C_G(G)$.

**Lemma 2.20.** If $S$ is a subset of $G$, then $C_G(S)$ is a subgroup of $G$.

**Proof.** Certainly the identity of $G$ commutes with all elements in $S$—it commutes with all elements in $G$! If $x$ and $y$ commute with all elements in $S$ so do $x^{-1}$ and $xy$. To see this let $g \in S$. The cancellation law implies that $xg = gx$ and only if $x^{-1}(xg)x^{-1} = x^{-1}(gx)x^{-1}$, i.e., if and only if $gx^{-1} = x^{-1}y$, so $x^{-1} \in C_G(S)$. Furthermore, $xyg = xgy = gxy$ so $xy \in C_G(S)$. □

5.5. The subgroup generated by a subset of $G$. If $S$ is a subset of $G$ we introduce the notation

$$\langle S \rangle := \{\text{the smallest subgroup of } G \text{ containing } S\}$$

and call $\langle S \rangle$ the subgroup of $G$ generated by $S$. For the definition to make sense we must check there is a smallest such subgroup: if $H$ and $K$ are subgroups containing $S$, then $H \cap K$ is a subgroup containing $S$, so one sees that

$$\langle S \rangle = \text{the intersection of all the subgroups of } G \text{ that contain } S.$$ 

Notice that Proposition ??2.6 says that the symmetric group $S_n$ is generated by transpositions. However, one can be efficient and generate it with just $n - 1$ transpositions. Show that $S_n = \{(1\,2), (2\,3), \ldots, (n-1\,n)\}$.

**Lemma 2.21.** If $x \in G$, then $\langle x \rangle = \{x^n \mid n \in \mathbb{Z}\}$.

**Proof.** Certainly $\{x^n \mid n \in \mathbb{Z}\}$ is a group: it contains the identity, $x^0$, and contains inverses, $(x^n)^{-1} = x^{-n}$, and is closed under products, $x^m x^n = x^{m+n}$. Hence $\{x^n \mid n \in \mathbb{Z}\}$ is a subgroup of $G$ that contains $x$.

Let $H$ be a subgroup of $G$ containing $x$. Then $H$ contains $x^n$ for all $n > 0$ because it is closed under multiplication; $H$ contains inverses so contains $x^{-1}$, and products of $x^{-1}$ with itself, so contains $x^{-n}$ for all $n > 0$; $H$ contains the identity so contains $x^0$ too. Thus, $H$ contains $\{x^n \mid n \in \mathbb{Z}\}$. Hence $\{x^n \mid n \in \mathbb{Z}\}$ is the smallest subgroup of $G$ containing $x$. □
5.5.1. Exercise. Let $a$ and $b$ be elements in a group $G$. Suppose that $ab = ba$. Show that $\langle a, b \rangle$ is abelian and list all the elements in it. I don’t mind if the same element appears more than once on your list.

5.5.2. Exercise. If $a_1, \ldots, a_n \in \mathbb{Z}$ show that $\langle a_1, \ldots, a_n \rangle = d\mathbb{Z}$ where $d$ is the greatest common divisor of $a_1, \ldots, a_n$.

5.5.3. A non-example. Let $H$ be the subset of the group $G = (\mathbb{Q} - \{0\}, \cdot)$ consisting of the negative integers and 1. This fails to be a subgroup because the product of two elements in $H$ need not belong to $H$; but it satisfies all the other axioms.

5.5.4. We say more about subgroups generated by subsets of $G$ in section ??.

5.6. Two subgroups of $G$. Suppose $H$ and $K$ are subgroups of $G$. We have seen that $H \cap K$ is also a subgroup of $G$. It is the largest subgroup of $G$ contained in both $H$ and $K$.

5.6.1. The abelian case. There is an analogy with least common multiples. If $a$ and $b$ are non-zero integers, then $za \cap zb = zc$ where $c$ is the least common multiple of $a$ and $b$. For example, $20\mathbb{Z} \cap 12\mathbb{Z} = 60\mathbb{Z}$.

It is natural to ask if we can do something similar with greatest common divisors. We can. The smallest subgroup that contains $20\mathbb{Z}$ and $12\mathbb{Z}$ is $4\mathbb{Z}$ and $4 = \gcd\{12, 20\}$. If $A$ and $B$ are subsets of the integers we sometimes write

$$A + B := \{a + b \mid a \in A \text{ and } b \in B\}.$$  

For example, $20\mathbb{Z} + 12\mathbb{Z} = 4\mathbb{Z}$. If $A$ and $B$ are subgroups of $\mathbb{Z}$, then $A + B$ is a subgroup of $\mathbb{Z}$ and is the smallest subgroup that contains both $A$ and $B$.

If $a$ and $b$ are non-zero integers, then $za + zb = zd$ where $d = \gcd\{a, b\}$ and $zd$ is the smallest subgroup of $\mathbb{Z}$ that contains both $za$ and $zb$.

We extend this notation in an obvious way.

**Proposition 2.22.** Let $H$ and $K$ be subgroups of an abelian group $(G, +)$. Define

$$H + K := \{x + y \mid x \in H \text{ and } y \in K\}.$$  

Then $H + K$ is a subgroup of $G$ and is the smallest subgroup that contains both $H$ and $K$.

**Proof.**

In other words, $H + K$ is the subgroup of $G$ generated by $H \cup K$. We also write $H + K = \langle H, K \rangle$.

5.6.2. The non-abelian case. If $G$ is not abelian a little more care is required. First we will write $G$ multiplicatively and we define

$$AB := \{ab \mid a \in A \text{ and } b \in B\}$$  

whenever $A$ and $B$ are subsets of $G$. 


6. Cosets and Lagrange’s Theorem

6.1. Let \( H \) be a subgroup of \( G \). For each element \( x \in G \), we define
\[
xH := \{ xh \mid h \in H \}.
\]
We call the sets \( xH \) the **right cosets** of \( H \) in \( G \).

There is a similar notion of left cosets.

Because \( e \in H \), \( x \in xH \), so \( G \) is the union of the right cosets.

**Lemma 2.24.** If \( H \) is a subgroup of \( G \), then \( G \) is the disjoint union of the distinct right cosets \( xH \).

**Proof.** If \( xH \cap yH \neq \emptyset \), then \( xa = yb \) for some \( a, b \in H \). If \( c \in H \), then \( xc = xaa^{-1}c = yba^{-1}c \) which is in \( yH \) because \( a, b, c \in H \). Hence \( xH \subseteq yH \).

Similarly, \( yH \subseteq xH \); thus \( xH = yH \).

Lemma 2.24 says that every element of \( G \) belongs to one, and only one, coset of \( H \). Thus \( xH \) is the *unique* right coset of \( H \) that contains \( x \). We often refer to \( xH \) by calling it the **right coset containing** \( x \).

**Lemma 2.25.** If \( H \) is a subgroup of \( G \), then \( |xH| = |H| \).

**Proof.** There is a map \( H \rightarrow xH \), \( h \mapsto xh \); this map is surjective and is injective because if \( xh = xh' \), then \( h = h' \); hence \( xH \) has the same number of elements as \( H \).

**Proposition 2.26.** (Lagrange’s Theorem) Let \( H \) be a subgroup of a finite group \( G \). Then \(|H| \) divides \(|G|\), and \(|G|/|H|\) is the number of right cosets of \( H \) in \( G \).

**Proof.** Since \( G \) is a disjoint union of sets \( xH \), each of which has the same number of elements as \( H \),
\[
|G| = n|H|
\]
where \( n \) is the number of right cosets of \( H \) in \( G \).

Of course, \(|G|/|H|\) is also the number of left cosets of \( H \) in \( G \).
6.1.1. The index of a subgroup. Let $H$ be a subgroup of a group $G$. The index of $H$ in $G$ is

$$[G : H] := \text{the number of cosets of } H \text{ in } G.$$ 

The proof of Lagrange’s theorem tells us the following.

**Corollary 2.27.** If $H$ be a subgroup of a group $G$, then

$$|G| = |H| \cdot [G : H].$$

**Proof.** This is because $G$ is the disjoint union of the cosets of $H$, each of which has $|H|$ elements, and there are $[G : H]$ different cosets. $\square$

6.1.2. Warning: Although the number of left cosets of $H$ is the same as the number of right cosets of $H$, in general a left coset is not a right coset. For example, if $H$ is the subgroup $\{1, (12)\}$ of $S_3$, then

$$(13)H = \{(13), (13)(12)\} = \{(13), (123)\}$$

but

$$H(13) = \{(13), (12)(13)\} = \{(13), (132)\} \neq (13)H.$$ 

When $G$ is abelian left and right cosets coincide: if $H$ is a subgroup of $G$ and $x \in G$, then

$$xH = \{xh \mid h \in H\} = \{hx \mid h \in H\} = Hx.$$ 

6.2. Notation. We extend the notation $xH$ for a coset of a subgroup by writing

$$xS := \{xs \mid s \in S\}$$

for any subset $S \subset G$. Likewise,

$$Sx := \{sx \mid s \in S\}.$$ 

More generally, if $T$ is another subset of $G$ we write

$$ST := \{st \mid s \in S, t \in T\}.$$ 

This notation behaves in a nice way. For example, $x(yS) = (xy)S$ and $(xS)y = xSy$, so we simply write $xyS$ and $xSy$ for these subsets. Similarly, $x(ST) = (xS)T$, and so on.

6.3. Cosets when the group operation is denoted by $\cdot$. Suppose $(G, \cdot)$ is an abelian group. If $H$ is a subgroup of $G$ and $x \in G$, then the coset of $H$ that contains $x$ is

$$x + H = \{x + h \mid h \in H\}.$$ 

We already met an important example of such cosets when we defined the groups $\mathbb{Z}/d$. By definition, the elements of $\mathbb{Z}/d$ are the cosets $r + d\mathbb{Z}$, $r \in \mathbb{Z}$. 
6.4. Parallel lines as cosets. Consider the plane $\mathbb{R}^2$. It is an abelian group under addition:

$$(a, b) + (c, d) = (a + c, b + d).$$

Fix a number $m \in \mathbb{R} \cup \{\infty\}$ and let $L$ be the line through the origin with slope $m$. Then $L$ is a subgroup of $\mathbb{R}^2$. (Check this!) If $p = (a, b)$, the set

$$p + L := \{(a, b) + (c, d) \mid (c, d) \in L\}$$

is a coset of $L$ and every coset of $L$ is of the form $p + L$ for some $p$. The points in $p + L$ are exactly the points on the line through $p$ having slope $m$. Thus, the cosets of $L$ are exactly the lines of slope $m$. If you form a mental picture of this situation you see clearly that $\mathbb{R}^2$ is the disjoint union of the lines of slope $m$: each point in $\mathbb{R}^2$ lies on a unique line of slope $m$. In other words, $\mathbb{R}^2$ is the disjoint union of the cosets of $L$.

You should have a similar mental image of the right cosets $Hx$ of a subgroup $H$ of a group $G$: think of the $Hx$s as being “parallel” to one another: $H$ plays the role of the line of slope $m$ passing through the origin—it is the unique coset that contains the identity of $G$. The key idea in the proof of Lagrange’s Theorem, that all cosets have the same number of elements, might now seem more inevitable and natural—after all, $Hx$ is just the result of “sliding $H$ by a distance of $x$”.

7. Cyclic groups

7.1. Definition. A group $G$ is cyclic if it is generated by one element, i.e., if $G = \langle x \rangle$ for some $x$. We then call $x$ a generator of $G$ and say that $x$ generates $G$.

Equivalently, $G$ is cyclic if it equals $\{x^n \mid n \in \mathbb{Z}\}$ for some $x \in G$. Note we do not insist that $x^n$ and $x^m$ are different if $m$ and $n$ are different.

The group of $n$th roots of unity,

$$\mu_n := \{z \in \mathbb{C} \mid z^n = 1\},$$

is cyclic. For example, $e^{2\pi i/n}$ is a generator for $\mu_n$. More generally, if $r$ is any integer relatively prime to $n$, $e^{2\pi i/n}$ is a generator for $\mu_n$.

The group of integers $(\mathbb{Z}, +)$ is cyclic. Clearly, $\mathbb{Z} = \langle 1 \rangle$ and $\mathbb{Z} = \langle -1 \rangle$.

A cyclic group is abelian because $x^m x^n = x^{m+n} = x^{n+m} = x^n x^m$.

Every group contains a cyclic subgroup, namely $\langle x \rangle$ for any $x \in G$.

7.2. The order of an element. The order of an element $x \in G$ is the smallest positive integer $n$ such that $x^n = 1$. If there is no such $n$ we say that $x$ has infinite order.

Proposition 2.28. The order of $x$ is equal to the number of elements in $\langle x \rangle$.

Proof. Suppose first that $x$ has infinite order. If $\langle x \rangle$ was finite there would be integers $m \neq n$ such that $x^m = x^n$. But then $x^{m-n} = x^{n-m} = 1$ so $x$
would have finite order. This contradiction shows the result is true when \( x \) has infinite order.

Now suppose \( x \) has finite order, \( n \) say. We will show that
\[
\langle x \rangle = \{1, x, x^2, \ldots, x^{n-1}\}
\]
and that these \( n \) elements are distinct from one another. If \( m \) is any integer,
then \( m = nq + r \) for some \( q, r \in \mathbb{Z} \) such that \( 0 \leq r \leq n - 1 \). It follows that
\[
x^m = x^{na+r} = x^{na}x^r = (x^n)^a x^r = x^r.
\]
Hence \( \langle x \rangle = \{1, x, x^2, \ldots, x^{n-1}\} \).

If \( x^i = x^j \) for some integers \( 0 \leq i < j \leq n - 1 \), then \( x^{i-j} = x^{j-i} = 1 \),
so \( x^{i-j} = 1 \); but \( 0 < |i - j| < n \), contradicting the hypothesis that \( x \) has
order \( n \). \( \square \)

**Proposition 2.29.** Let \( x \) be an element in a finite group \( G \). Then
\begin{enumerate}
\item \( x \) has finite order;
\item the order of \( x \) divides \( |G| \);
\item \( x^{[G]} = e \).
\end{enumerate}

**Proof.** (1) Since \( \langle x \rangle \) is finite the order of \( x \) is finite.

(2) The order of \( x \) is the number of elements in \( \langle x \rangle \), and Lagrange’s
Theorem tells us that this number divides \( |G| \).

(3) This follows at once from (2) because \( x^{rs} = (x^r)^s \). \( \square \)

It is important to distinguish between the order of an element and the
order of a group. There are infinite groups in which every element has finite
order. For example, the group of all subsets of \( \mathbb{Z} \) with group operation
\[
A \oplus B = \{x \mid x \in A \cup B \text{ but } x \notin A \cap B\}
\]
is infinite but every element in it has order two (except the identity which
has order one).

7.2.1. An infinite group all of whose elements have finite order. Let \( G \)
be the subgroup of \( (\mathbb{Q},+) \) consisting of those fractions \( a/b \) such that \( b \) is
a power of \( 2 \); that is
\[
G := \left\{ \frac{a}{b} \mid a, b \in \mathbb{Z}, b = 2^d \text{ for some } d \geq 0 \right\}.
\]
Notice that \( \mathbb{Z} \) is a subgroup of \( G \). The quotient \( G/\mathbb{Z} \) is an infinite
group, and every element in it has finite order because if \( g \in G \), then \( g = a/2^d \)
for some \( d \) and, if we set \( n = 2^d \), \( ng = g + \cdots + g \in \mathbb{Z} \). Hence, in \( G/\mathbb{Z} \),
\( n(g + \mathbb{Z}) = ng + \mathbb{Z} = \mathbb{Z} \), the identity element to \( G/\mathbb{Z} \).

**Proposition 2.30.** Let \( G \) be a cyclic group and \( d \) a positive integer. Then
\begin{enumerate}
\item \( G \) is isomorphic to \( \mathbb{Z} \) if it is infinite;
\item \( G \) is isomorphic to \( \mathbb{Z}/d \) if it has \( d \) elements.
\end{enumerate}
7. CYCLIC GROUPS

Proof. By hypothesis, \( G = \langle x \rangle = \{ x^n \mid n \in \mathbb{Z} \} \). Define \( f : \mathbb{Z} \to G \) by
\[
f(n) := x^n.
\]
Notice that \( f(m + n) = x^{m+n} = x^m x^n = f(m)f(n) \). If \( f \) is bijective it is an isomorphism so in that case \( G \cong \mathbb{Z} \).

Now assume \( f \) is not bijective. In that case \( x^n = x^m \) for some \( n \neq m \). If follows that \( x^{n-m} = x^{m-n} = 1 \). Let \( d \) be the smallest positive integer such that \( x^d = 1 \). Define
\[
g : \mathbb{Z}/d \to G \quad \text{by} \quad f(r + d\mathbb{Z}) := x^r.
\]
Then \( g \) is surjective, and therefore bijective because \( |G| = d = |\mathbb{Z}/d| \). Also
\[
f((r + d\mathbb{Z}) + (s + d\mathbb{Z})) = f((r + s) + d\mathbb{Z})
\]
\[
= x^{r+s}
\]
\[
= x^r x^s
\]
\[
= f(r + d\mathbb{Z}) f(s + d\mathbb{Z})
\]
for all \( r \) and \( s \). Hence \( f \) is an isomorphism; i.e., \( G \cong \mathbb{Z}/d \). \( \square \)

The previous result says the only isomorphism; i.e., \( G \cong \mathbb{Z}/d \), \( d > 0 \). Up to isomorphism!

Corollary 2.31. Let \( n \) be a positive integer. Then there is an isomorphism
\[
\mu_n \cong \mathbb{Z}/n
\]
between the group of complex \( n \)th roots of unity and \( \mathbb{Z}/n \).

Proof. Since \( \mu_n \) is cyclic, generated by \( e^{2\pi i/n} \) for example, the corollary follows from Proposition 2.30. \( \square \)

The next result says that if \( p \) is a prime number there is only one group having \( p \) elements, namely \( \mathbb{Z}/p \).\(^5\)

Theorem 2.32. If \( p \) is a positive prime number, there is, up to isomorphism, a unique group with \( p \) elements, namely \( \mathbb{Z}/p \). In other words, if \( G \) is a group having \( p \) elements then \( G \cong \mathbb{Z}/p \).

Proof. Let \( p \) be a positive prime and \( G \) a group with \( p \) elements.

We fix an an element \( x \) in \( G \) that is not the identity. Then the order of \( x \) is \( \geq 2 \) and divides \( |G| \) by Proposition 2.29, so must be \( p \). By Proposition 2.28, the order of \( x \) is the number of elements in \( \langle x \rangle \). Thus
\[
G = \{ 1, x, \ldots, x^{p-1} \}.
\]
In particular, \( G \) is a cyclic group with \( p \) elements so \( G \) is isomorphic to \( \mathbb{Z}/p \) by Proposition 2.30. \( \square \)

\(^5\)When I say “only” I should add the qualifier “up to isomorphism”, but I’m not doing that because it would make the sentence a little clunky and I want to encourage you to add that qualifier internally.
8. The product of two groups

8.1. The definition. Let $G$ and $H$ be groups. We make their cartesian product

$$G \times H = \{(g, h) \mid g \in G, h \in H\}$$

into a group by declaring the product to be

$$(a, x) \cdot (b, y) := (ab, xy)$$

for $a, b \in G$ and $x, y \in H$. We call this the product of the groups $G$ and $H$.

8.1.1. Exercises.

1. Show (2-4) makes $G \times H$ a group.
2. Show that $G \times H \cong H \times G$.
3. If $H_1$ and $H_2$ are isomorphic groups show that $G \times H_1$ is isomorphic to $G \times H_2$.
4. Let $G_1$, $G_2$, $H_1$, and $H_2$ be groups. Suppose that $G_1 \cong G_2$ and $H_1 \cong H_2$. Show that $G_1 \times H_1 \cong G_2 \times H_2$.

8.1.2. Remark. If $G$ and $H$ are abelian we often call their direct product the direct sum and denote it by $G \oplus H$.

8.1.3. An important isomorphism. When you encountered complex numbers did you learn that every non-zero complex number can be written in a unique way as

$$r e^{i \theta}$$

where $r$ is a positive real number and $\theta \in [0, 2\pi)$. I hope so. We will review that in a moment but if you already know that fact are you aware that it is really saying that there is an isomorphism

$$f : (\mathbb{R}_{>0}, \cdot) \times U(1) \to (\mathbb{C} - \{0\}, \cdot)$$

given by the formula $f(r, e^{i \theta}) := re^{i \theta}$.

8.2. An example. Let $F = \{\{0, 1\}, +\}$ be the group with addition $0 + 0 = 1 + 1 = 0$ and $0 + 1 = 1 + 0 = 1$. Then $F \times F$ has four elements, $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$. It is simpler to write these as $00$, $01$, $10$, and $11$. The group operation on $G \times G$ given by (??) is

<table>
<thead>
<tr>
<th></th>
<th>00</th>
<th>01</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>00</td>
<td>01</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>01</td>
<td>01</td>
<td>00</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>11</td>
<td>00</td>
<td>01</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>10</td>
<td>01</td>
<td>00</td>
</tr>
</tbody>
</table>

This is the group in example 2.3.12. Now you know why we labeled it $F^2$: it is isomorphic to $F \times F$. Since $F \cong \mathbb{Z}/2$, $F^2 \cong (\mathbb{Z}/2) \times (\mathbb{Z}/2)$.

Suppose that $X$ is a set and $X = Y \cup Z$, i.e., $X = Y \cup Z$ and $X \cap Z = \phi$. Then $X! \times Z!$ is a subgroup of $X!$.

More to say!
8. The Product of Two Groups

8.2.1. Products of several groups. Let \( G, H, \) and \( K \) be groups. It is easy to show that
\[
(G \times H) \times K \cong G \times (H \times K)
\]
so we just write \( G \times H \times K \) for this group. This idea extends to arbitrary finite collections of groups. If \( G_1, \ldots, G_n \) are groups the elements of \( G_1 \times \cdots \times G_n \) are ordered \( n \)-tuples \((x_1, \ldots, x_n)\) with \( x_i \in G_i \) and this set is made into a group by declaring that
\[
(x_1, \ldots, x_n)(y_1, \ldots, y_n) = (x_1y_1, \ldots, x_ny_n).
\]

8.3. The Chinese Remainder Theorem. We will now prove the Chinese Remainder Theorem, a result that is an important step towards the classification of all finite abelian groups (up to isomorphism, of course!).

Theorem 2.33 (The Chinese Remainder Theorem). Suppose \( d \) and \( e \) are positive integers. If \( \gcd(d, e) = 1 \), then there is an isomorphism
\[
f : \mathbb{Z}/de \rightarrow (\mathbb{Z}/d) \times (\mathbb{Z}/e)
\]
given by \( f(r + de \mathbb{Z}) := (r + d \mathbb{Z}, r + e \mathbb{Z}) \).

Proof. First we will show \( f \) is well defined. If \( r + de \mathbb{Z} = s + de \mathbb{Z} \), then \( de \) divides \( r - s \) so \( r - s \in d \mathbb{Z} \) and \( r - s \in e \mathbb{Z} \). It follows that \( r + d \mathbb{Z} = s + d \mathbb{Z} \) and \( r + e \mathbb{Z} = s + e \mathbb{Z} \). Hence \( f \) is well-defined.

We will now show \( f \) is injective. It will follow from this that \( f \) is bijective because the sets \( \mathbb{Z}/de \) and \( (\mathbb{Z}/d) \times (\mathbb{Z}/e) \) have the same number of elements. By Lemma 2.10(2), to show \( f \) is injective we need only show that if \( f(r + de \mathbb{Z}) = (d \mathbb{Z}, e \mathbb{Z}) \), then \( r \in de \mathbb{Z} \).

Suppose \( f(r + de \mathbb{Z}) = (d \mathbb{Z}, e \mathbb{Z}) \). Then \( r \in d \mathbb{Z} \) and \( r \in e \mathbb{Z} \); say \( r = du = ev \). Because \( \gcd(d, e) = 1 \) there are integers \( a \) and \( b \) such that \( ad + be = 1 \). It follows that
\[
u = adu + be u = av + bu = e(au + bu).
\]
Hence \( r = du = de(au + bu) \in ed \mathbb{Z} \). Thus \( r + ed \mathbb{Z} = ed \mathbb{Z} \). Thus, \( f^{-1}((d \mathbb{Z}, e \mathbb{Z})) = \{de \mathbb{Z}\} \) and it follows from Lemma 2.10(2) that \( f \) is injective and therefore bijective.

Finally, the calculation
\[
f((r + de \mathbb{Z}) + (s + de \mathbb{Z})) = f((r + s) + de \mathbb{Z})
\]
\[
= (r + s + d \mathbb{Z}, r + s + e \mathbb{Z})
\]
\[
= (r + d \mathbb{Z}, r + e \mathbb{Z}) + (s + d \mathbb{Z}, s + e \mathbb{Z})
\]
\[
= f(r + de \mathbb{Z}) + f(s + de \mathbb{Z})
\]
shows \( f \) is an isomorphism. \( \Box \)

Thus, for example,
\[
\mathbb{Z}/60 \cong \mathbb{Z}/5 \times \mathbb{Z}/12 \cong \mathbb{Z}/5 \times \mathbb{Z}/4 \times \mathbb{Z}/3.
\]
In contrast, we have already seen that
\[ \mathbb{Z}/4 \nless \neq \mathbb{Z}/2 \times \mathbb{Z}/2. \]

8.3.1. History. The Chinese Remainder Theorem appeared in ancient times in the following form: if \( m_1, \ldots, m_n \) are pairwise relatively prime integers, and \( a_1, \ldots, a_n \) are any integers, then there is an integer \( d \) such that \( d \equiv a_i (\text{mod } m_i) \) for all \( i \). This statement appears in the manuscript Mathematical Treatise in Nine Sections written by Chin Chiu Shao in 1247 (search on the web if you want to know more).

8.4. How to recognize a product. If we have a group \( G \) and can then show it is a product of two smaller groups we have made progress: generally speaking smaller groups are easier to understand than large ones so we can understand \( G \) by understanding the two factors in its expression as a product. The next result gives a criterion for recognizing when a group can be written as a product of two of its subgroups.

Theorem 2.34. Let \( H \) and \( K \) be subgroups of a group \( G \). Suppose that \( hk = kh \) for all \( h \in H \) and \( k \in K \). If \( HK = G \) and \( H \cap K = \{e\} \), then
\[ G \cong H \times K. \]

The map \( f : H \times K \to G \) given by \( f(h,k) := hk \) is an isomorphism.

Proof. Since \( HK = G \), \( f \) is surjective. Furthermore, if \( x = (h,k) \) and \( y = (h',k') \), then
\[
\begin{align*}
f(xy) &= f((h,k) \cdot (h',k')) \\
       &= f(hh',kk') \\
       &= hh'kk' \\
       &= hh'k'k \\
       &= f(h,k)f(h',k') \\
       &= f(x)f(y).
\end{align*}
\]
We will now use Lemma 2.10 to show that \( f \) is injective. If \( f(h,k) = e \), then \( hk = e \) so \( h = k^{-1} \) and therefore \( h \in H \cap K \). But \( H \cap K = \{e\} \) so \( h = e \) and \( k = e \). Hence \( f^{-1}(e) = (e,e) \) and Lemma 2.10 implies \( f \) is injective. \( \square \)

In the situation of the Theorem 2.34 we usually write
\[ G = H \times K. \]
This equality means, among other things, that every element in \( G \) can be written in a unique way as a product \( hk \), i.e., given \( g \in G \), there is a unique \( h \) in \( H \) and a unique \( k \) in \( K \) such that \( g = hk \).

8.4.1. The abelian case. If \( H \) and \( K \) are subgroups of an abelian group \((G,+)\) such that \( H + K = G \) and \( H \cap K = \{0\} \), then Theorem 2.34 tells us that \( G = H \times K \). We will make use of this in the proof of Theorem 2.37 in the next section.
9. Finite abelian groups

In the section we classify all finite abelian groups. Our treatment follows J.S. Milne’s notes at http://www.jmilne.org/math/CourseNotes/GT.pdf.

First, because all the groups in this section are abelian we will use the symbol + for the group operation in each.

In the next section we introduce the idea of a basis for a finite abelian group. The analogy with the idea of a basis for a vector space should help you. First observe that \( \mathbb{R}^n \) is an abelian group under the usual addition of vectors. Although the only element of finite order in \( \mathbb{R}^n \) is the zero vector, i.e., the identity don’t let that worry you. Recall that a basis for \( \mathbb{R}^n \) is a subset \( \{ y_i \mid i \in I \} \) that spans it and is linearly independent. The spanning condition means that every element in \( \mathbb{R}^n \) can be written as a sum \( \sum_i \lambda_i y_i \) for some \( \lambda_i \)s in \( \mathbb{R} \), and the linear independence means that if \( \sum_i \lambda_i y_i = 0 \), then all \( \lambda_i \) are zero. Of course a basis for \( \mathbb{R}^n \) must consist of exactly \( n \) elements and leads to an isomorphism

\[
\mathbb{R}^n = \mathbb{R}^1 \times \cdots \times \mathbb{R}^n.
\]

9.1. The notion of a basis for a finite abelian group. A **basis** for an abelian group \( G \) is a subset \( \{ x_1, \ldots, x_n \} \) such that

\[
G = \langle x_1 \rangle \times \langle x_2 \rangle \times \cdots \times \langle x_n \rangle.
\]

**Lemma 2.35.** A set \( \{ x_1, \ldots, x_n \} \) is a basis for \( G \) if and only if

1. every element in \( G \) can be written as \( a_1 x_1 + \cdots + a_n x_n \) for some \( a_1, \ldots, a_n \in \mathbb{Z} \) and
2. \( a_1 x_1 + \cdots + a_n x_n = 0 \) implies \( a_1 x_1 = a_2 x_2 = \cdots = a_n x_n = 0 \).

**Proof.** Exercise. \( \Box \)

**Lemma 2.36 (Milne).** Suppose \( G \) is generated by \( \{ x_1, \ldots, x_n \} \). Suppose \( a_1, \ldots, a_n \) are integers whose greatest common divisor is 1. Then \( G \) is generated by a set \( \{ y_1, \ldots, y_n \} \) where \( y_1 = a_1 x_1 + \cdots + a_n x_n \).

**Proof.** If some \( a_i < 0 \) we replace \( a_i \) by \(-a_i\) and \( x_i \) by \(-x_i\). This allows us to assume that all \( a_i \)s are positive. We argue by induction on \( s = a_1 + \cdots + a_n \). If \( s = 1 \), the hypotheses imply \( G \) is generated by \( x_1 \) and \( y_1 = x_1 \) so the result is obviously true.

Now suppose \( s > 1 \). Then at least two \( a_i \) are positive, say \( a_1 \geq a_2 > 0 \). It is clear that

\[
G = \langle x_1, \ldots, x_n \rangle = \langle x_1, x_2 + x_1, x_3, \ldots, x_n \rangle
\]

and \( \text{gcd}\{ a_1 - a_2, a_2, a_3, \ldots, a_n \} = 1 \). But the sum of these numbers is less than \( s \) so the induction hypothesis implies that \( G \) is generated by a set \( \{ y_1, \ldots, y_n \} \) where

\[
\begin{align*}
y_1 &= (a_1 - a_2)x_1 + a_2(x_2 + x_1) + a_3x_3 + \cdots + a_n x_n \\
&= a_1 x_1 + \cdots + a_n x_n.
\end{align*}
\]
This completes the proof.

**Theorem 2.37.** Every finite abelian group has a basis. In other words, if \( G \) is a finite abelian group, there are positive integers \( n_1, \ldots, n_k \) such that

\[
G \cong (\mathbb{Z}/n_1) \times \cdots \times (\mathbb{Z}/n_k).
\]

**Proof.** [Milne] Let \( G \) be a finite abelian group. We will argue by induction on the number of generators of \( G \). If \( G \) has a single generator it is cyclic and therefore has a basis. We therefore assume that \( G \) needs \( n > 1 \) generators. We pick a generating set \( \{x_1, \ldots, x_n\} \) with \( x_1 \) having minimal order. We will show that

\[
\langle x_1 \rangle \cap \langle x_2, \ldots, x_n \rangle = 0
\]

which, by the remarks in section 8.4.1, implies that \( G = \langle x_1 \rangle \times \langle x_2, \ldots, x_n \rangle \), so proving the theorem.

Suppose the intersection is not zero. Then there are integers \( a_1, \ldots, a_n \)

\[
a_1 x_1 + \cdots + a_n x_n = 0
\]

and \( a_1 x_1 \neq 0 \). Of course, we can assume that \( a_1 \) is strictly smaller than

the order of \( x_1 \). Let \( d = \gcd(a_1, \ldots, a_n) \) and define \( b_i := a_i/d_i \). Then

\[
\gcd\{b_1, \ldots, b_n\} = 1
\]

so the lemma implies that \( G = \langle y_1, y_2, \ldots, y_n \rangle \) with

\[
y_1 = b_1 x_1 + \cdots + b_n x_n.
\]

But

\[
d y_1 = a_1 x_1 + \cdots + a_n x_n = 0
\]

and \( d \) divides \( a_1 \) so is \( \leq a_1 < \text{order}(x_1) \). This contradicts the choice of \( x_1 \).

The integers \( n_1, \ldots, n_k \) are not uniquely determined by \( G \). For example, the Chinese Remainder Theorem tells us that

\[
\mathbb{Z}/mn \cong (\mathbb{Z}/m) \times (\mathbb{Z}/n)
\]

if \( \gcd(m, n) = 1 \). However, using the Chinese Remainder Theorem and factoring the \( n_i \)s and then using the Chinese remainder theorem to multiply the factors together in appropriate ways we can show there are prime numbers \( p_1, \ldots, p_m \), possibly with repetitions, and integers \( r_1, \ldots, r_m \) such that

\[
G \cong \frac{\mathbb{Z}}{p_1^{r_1}} \times \cdots \times \frac{\mathbb{Z}}{p_m^{r_m}}
\]

and the primes and the \( r_i \)s are uniquely determined by \( G \). We can also use the

Chinese Chinese Remainder Theorem to deduce that

\[
G \cong \frac{\mathbb{Z}}{a_1} \times \cdots \times \frac{\mathbb{Z}}{a_{\ell}}
\]

where \( a_1 \) divides \( a_2 \), \( a_2 \) divides \( a_3 \), and so on and so on. The \( a_i \)s are uniquely
determined by \( G \).

For example, if

\[
G = \frac{\mathbb{Z}}{18} \times \frac{\mathbb{Z}}{40} \times \frac{\mathbb{Z}}{8} \times \frac{\mathbb{Z}}{48}
\]

then
\[ G \cong \frac{\mathbb{Z}}{2} \times \frac{\mathbb{Z}}{8} \times \frac{\mathbb{Z}}{8} \times \frac{\mathbb{Z}}{16} \times \frac{\mathbb{Z}}{3} \times \frac{\mathbb{Z}}{9} \times \frac{\mathbb{Z}}{5} \]

and
\[ G \cong \frac{\mathbb{Z}}{2} \times \frac{\mathbb{Z}}{8} \times \frac{\mathbb{Z}}{24} \times \frac{\mathbb{Z}}{720}. \]

There are lots of other ways of writing this group. For example, using the Chinese Remainder Theorem we could remove the factor of 5 from 720 and then replace 24 by \(24 \times 5 = 120\) to write
\[ G \cong \frac{\mathbb{Z}}{2} \times \frac{\mathbb{Z}}{8} \times \frac{\mathbb{Z}}{120} \times \frac{\mathbb{Z}}{144}. \]

Endless fun.

10. Putting together some of the ideas in this chapter

10.1. Some subgroups of \(S_n\). The next result should be “obvious”. Indeed, writing out a proof of it seems to make it less obvious.

**Proposition 2.38.** If \(m \leq n\), then \(S_n\) contains a subgroup isomorphic to \(S_m\).

**Proof.** Since \(m \leq n\) we can write \(\{1, 2, \ldots, n\}\) as the disjoint union of two sets \(X\) and \(Y\) such that \(X\) has \(n - m\) elements and \(Y\) has \(m\) elements. Let
\[ H := \{\sigma \in S_n \mid \sigma(x) = x \text{ for all } x \in X\}. \]

Then \(H\) is a subgroup of \(S_n\): certainly the identity permutation is in \(H\); if \(\sigma(x) = x\) for all \(x \in X\), then \(\sigma^{-1}(x) = x\) for all \(x \in X\) so \(\sigma^{-1}\) belongs to \(H\) whenever \(\sigma\) does; if \(\sigma\) and \(\tau\) fix every element in \(X\) so does \(\sigma \tau\).

We can characterize \(H\) as
\[ \{1, \ldots, n\} = X \cup \{y_1, \ldots, y_m\}. \]

Define \(f : S_m \to S_n\) by \(f(\sigma)(x) = x\) if \(x \in X\) and \(f(\sigma)(y_i) := y_{\sigma(i)}\) for \(i = 1, \ldots, m.\). If \(\sigma, \tau \in S_m\), then
\[ f(\sigma \tau)(y_i) = y_{\tau \sigma(i)} = f(\sigma)(y_{\tau(i)}) = f(\sigma)(f(\tau)(y_i)) = (f(\sigma) \circ f(\tau))(y_i) \]
and, if \(x \in X\), \(f(\sigma \tau)(x) = x = f(\sigma)(f(\tau)(x)) = (f(\sigma) \circ f(\tau))(x)\) so
\[ f(\sigma \tau) = f(\sigma) \circ f(\tau). \]

\[ \square \]

**Proposition 2.39.** If \(p + q \leq n\), then \(S_n\) contains a subgroup isomorphic to \(S_p \times S_q\).

10.2.
10.2.1. Example: generators of the symmetric group $S_3$. Given a group $G$, it is often useful to know that $G$ is generated by some particular subset of elements.

For example, $S_3$ is generated by $(12)$ and $(123)$. A laborious way to verify this claim is to take repeated products of these elements and eventually see that all elements of $S_3$ can be so obtained. A better way is this. Because $(12)$ has order two any subgroup of $S_3$ that contains $(12)$ must have even order. Because $(123)$ has order three the order of any subgroup of $S_3$ that contains $(123)$ must be divisible by 3. It follows that the order of any subgroup of $S_3$ that contains both $(12)$ and $(123)$ must be divisible by both 2 and 3, hence divisible by 6. But the only subgroup of $S_3$ whose order is divisible by 6 is $S_3$ itself. Thus $S_3$ is the smallest subgroup of $S_3$ that contains both $(12)$ and $(123)$.

A slightly different argument shows that $S_3$ is generated by $(12)$ and $(23)$. Let $H = \langle (12), (23) \rangle$. Since $H$ contains 1, $(12)$, and $(23)$, $|H| \geq 3$. But $H$ must have even order because it contains the order-two element $(12)$ and, by Lagrange’s Theorem, the order of $H$ must divide the order of $S_3$; the only possibility is that $|H| = 6$, whence $\langle (12), (23) \rangle = S_3$.

10.2.2. An example. The subgroup of $S_7$ generated by $(25)$ and $(125)$ is isomorphic to $S_3$. To see this first notice that the subgroup of $S_7$ consisting of the permutations that fix the elements in $\{3, 4, 6, 7\}$ is the same as the set of permutations that only move elements in $\{1, 2, 5\}$. But the latter subgroup is obviously isomorphic to $S_3$ and contains $(25)$ and $(125)$. It now follows from the argument in section 10.2.1 that $\langle (25), (125) \rangle \cong S_3$.

10.2.3. Another example. The subgroup of $S_7$ generated by $(25)$ and $(631)$ is isomorphic to $(\mathbb{Z}/2) \times (\mathbb{Z}/3)$.
Homomorphisms between groups

1. Definition

Let $G$ and $H$ be groups. A function $f : G \rightarrow H$ is a group homomorphism if

$$f(xy) = f(x)f(y) \quad \text{for all } x, y \in G.$$ 

Isomorphisms are exactly the bijective group homomorphisms.

Here are some simple examples. The determinant function for $n \times n$ matrices has the property that

$$\det(AB) = \det A \times \det B.$$ 

This is expressing the fact that

$$\det : \text{GL}(n, \mathbb{R}) \rightarrow (\mathbb{R} - \{0\}, \times)$$

is a homomorphism from the group of invertible $n \times n$ matrices to the multiplicative group of non-zero real numbers.

The trace function for square matrices has the property that $\text{Tr}(A+B) = \text{Tr}(A) + \text{Tr}(B)$ so $\text{Tr}$ is a homomorphism from the additive group of $n \times n$ matrices to $(\mathbb{R}, +)$.

The function $f : (\mathbb{R}, +) \rightarrow (\mathbb{C} - \{0\}, \times)$ defined by

$$f(x) := e^{2\pi ix}$$

is a group homomorphism simply because the law of exponents $a^x a^y = a^{x+y}$ holds for all real numbers $x$ and $y$ and all complex numbers $a$.

The function $f : \mathbb{Z} \rightarrow \mathbb{Z}/n$ defined by $f(r) := r + n\mathbb{Z}$ is a group homomorphism because the addition in $\mathbb{Z}/n$ was defined so that $f(r) + f(s) = f(r+s)$.

1.1. The kernel of a homomorphism. The kernel of a homomorphism $f : G \rightarrow H$ is defined to be

$$\ker f := \{ x \in G \mid f(x) = e \};$$

i.e., $\ker f$ is all the elements in $G$ that $f$ maps to the identity in $H$. The next result, which is of great importance, shows, among other things, that $\ker f$ is a subgroup of $G$. Whenever a homomorphism appears one immediately asks oneself, what is its kernel.

**Theorem 3.1.** Let $f : G \rightarrow H$ be a group homomorphism and write $K$ for $\ker f$. Then

1. $f(e) = e;
3. HOMOMORPHISMS BETWEEN GROUPS

(2) \( f(x) = f(x^{-1}) \) for all \( x \in G \).
(3) \( K \) is a subgroup of \( G \);
(4) \( gK = Kg \) for all \( g \in G \);
(5) if \( f(g) = h \), then \( \{ x \in G \mid f(x) = h \} = Kg \).

**Proof.** (1) If we set \( x = y = e \) in the equation \( f(xy) = f(x)f(y) \) we get \( f(e) = f(e)f(e) \) so by cancelling in \( H \), we see that \( f(e) \) is the identity element of \( H \).

(2) If \( x \in G \), then

\[
e_H = f(eG) = f(xx^{-1}) = f(x)f(x^{-1})
\]

so \( f(x^{-1}) = f(x)^{-1} \).

(3) By (1), \( K \) contains the identity. By (2), \( K \) contains \( x^{-1} \) whenever it contains \( x \). Finally, if \( x, y \in K \), then \( f(xy) = f(x)f(y) = ee = e \) so \( xy \in K \).

Hence \( K \) is a subgroup.

(5) Suppose \( f(g) = h \). Let \( x \in G \). Then

\[
f(x) = h \iff f(x) = f(g) \iff f(x)f(g)^{-1} = e \iff f(xg^{-1}) = e \iff xg^{-1} \in K \iff x \in Kg.
\]

Hence \( \{ x \in G \mid f(x) = h \} = Kg \). Similarly,

\[
f(x) = h \iff f(x) = f(g) \iff f(g)^{-1}f(x) = e \iff f(g^{-1}x) = e \iff g^{-1}x \in K \iff x \in gK.
\]

Hence \( \{ x \in G \mid f(x) = h \} = gK \). Hence \( Kg = gK \), so proving (4). \( \square \)

2. Normal subgroups

When we first introduced the notion of right and left cosets of a subgroup we emphasized the fact that a left coset need not be a right coset. More succinctly,

Let \( H \) be a subgroup of \( G \). If \( x \in G \), then \( xH \) is usually not the same as \( Hx \). For example, if \( G = S_3 \), \( H = \{1, (12)\} \), and \( x = (13) \), then

\[
xH = \{(13), (123)\}
\]

whereas

\[
Hx = \{(13), (12)\}.
\]

There are of course circumstances when \( xH \) is equal to \( Hx \); for example, if \( x \in H \), then \( xH = H = Hx \); if \( G \) is abelian, then \( xH = Hx \) for all subgroups \( H \) and all \( x \in G \).

Notice that \( xH = Hx \) if and only if \( xHx^{-1} = H \).

Subgroups \( H \) having the property that \( xH = Hx \) for all \( x \in G \) are particularly important; they are called **normal** subgroups.

Thus every subgroup of an abelian group is normal. The subgroup \( H = \{1, (12)\} \) of \( S_3 \) is not normal. The subgroups \( \{e\} \) and \( G \) are normal subgroups of \( G \). The subgroup \( \{1, (123), (132)\} \) of \( S_3 \) is normal. You can check this by hand or use Lemma 3.2 below.

The following simple result is very useful in practice.
3. The Quotient Group Construction

Lemma 3.2. Let $H$ be a subgroup of $G$ such that $[G : H] = 2$, i.e., $H$ has two cosets in $G$. Then $H$ is a normal subgroup of $G$.

Proof. Let $a \in G - H$. Then $aH \neq H$ so $aH \cap H = \emptyset$. Therefore $aH$ and $H$ are the left cosets of $H$. But $G$ is the disjoint union of the distinct left cosets of $H$ so $G = aH \sqcup H$ and we deduce that $aH = G - H$. Similarly, $Ha = G - H$. In particular, $aH = Ha$, so $H$ is normal. \hfill \Box

3. The quotient group construction

Suppose that $N$ is a normal subgroup of a group $G$. We write

$$G/N := \text{the cosets of } N \text{ in } G = \{xN \mid x \in G\} = \{Nx \mid x \in G\}.$$ 

Thus, elements of $G/N$ are subsets of $G$.

Definition 3.3. Let $N$ be a normal subgroup of a group $G$. We make $G/N$ a group by defining

$$(xN)(yN) := (xy)N$$

for all $x, y \in G$. We call $G/N$ the quotient group of $G$ by $N$. \hfill \Diamond

Proposition 3.4. This product does make $G/N$ a group.

Proof. First, the multiplication is well-defined (a particular coset can be labelled in several ways) so we must check the multiplication is unambiguously defined: if $xN = aN$ and $yN = bN$, then $xyN$ is the unique coset containing $xy$, but $x \in aN$ and $y \in bN$ so $x = am$ and $y = bn$ with $m, n \in N$, so $xy = ambn$; but $Nb = bN$ so $mb = bm'$ for some $m' \in N$, so $xy = abm = abn' \in abN$, whence $xyN = abN$.

Now check the multiplication is associative:

$$((xN)(yN))(zN) = (xyN)(zN) = (xy)zN = x(yz)N = (xN)((yN)(zN)).$$

There is an identity element, namely $N = eN$ because

$$(xN)(eN) = xeN = xN = exN = (eN)(xN).$$

And inverses exist because

$$(xN)(x^{-1}N) = (xx^{-1})N = eN = (x^{-1}x)N = (x^{-1}N)(xN),$$

so $x^{-1}N$ is the inverse of $xN$. \hfill \Box

If $\alpha : H \to K$ and $\beta : G \to H$ are group homomorphisms so is their composition $\alpha \beta : G \to K$. Check this.

If $G$ and $H$ are any groups we have the trivial homomorphism $\psi : G \to H$ defined by $\psi(x) = e$ for all $x \in G$.

If $G$ is any group and $x \in G$, there is a homomorphism $\psi : \mathbb{Z} \to G$ defined by

$$\psi(n) = x^n$$

for all $n \in \mathbb{Z}$. 
Let $G$ be the set of all matrices of the form
\[
\begin{pmatrix}
1 & r \\
0 & 1
\end{pmatrix}.
\]
Check that $G$ is a group under multiplication and construct an isomorphism
\[
\Phi : \mathbb{Z} \rightarrow G.
\]
If $N$ is a normal subgroup of a group $G$, then map $\psi : G \rightarrow G/N$ defined by $\psi(x) = xN$ is a group homomorphism—check this, and see that it is a homomorphism precisely because of the way multiplication is defined in $G/N$.

We say that two groups $G$ and $H$ are isomorphic if there is some isomorphism $\psi : G \rightarrow H$. Show that isomorphism is an equivalence relation.

**Example 3.5.** We write $\mu$ for the subgroup of $\mathbb{C}^\times$ consisting of the roots of unity. The map $\psi : \mathbb{Q}/\mathbb{Z} \rightarrow \mu$ defined by
\[
\psi([r]) := e^{2\pi ir}
\]
is an isomorphism, so
\[
\mu \cong \mathbb{Q}/\mathbb{Z}.
\]
This is rather trivial isomorphism is surprisingly important.

**Definition 3.6.** The kernel of a group homomorphism $\psi : G \rightarrow H$ is
\[
\ker \psi := \{ x \in G \mid \psi(x) = e \}.
\]
The image of $\psi$ is
\[
\text{im } \psi := \{ h \in H \mid h = \psi(x) \text{ for some } x \in G \}.
\]

**Theorem 3.7.** Let $\psi : G \rightarrow H$ be a group homomorphism. Then
(1) $\ker \psi$ is a normal subgroup of $G$;
(2) $\text{im } \psi$ is a subgroup of $H$;
(3) $\text{im } \psi \cong G/\ker \psi$.

**Proof.** Write $K := \ker \psi$. This is a subgroup: it contains $e$; if $a, b \in K$, then $\psi(ab) = \psi(a)\psi(b) = ee = e$, so $ab \in K$ too; if $a \in K$, then $\psi(a^{-1}) = \psi(a)^{-1} = e^{-1} = e$, so $a^{-1} \in K$ too.

To see that $K$ is normal, we must show that $xKx^{-1} \subset K$ for all $x \in G$; if $a \in K$, then $\psi(xax^{-1}) = \psi(x)\psi(a)\psi(x)^{-1} = \psi(x)e\psi(x)^{-1} = e$, so $xax^{-1} \in K$.

Define $\theta : G/K \rightarrow H$ by $\theta(xK) = \psi(x)$. First we check that $\theta$ is well-defined: if $xK = yK$, then

**Remark.** Let $\psi : G \rightarrow H$ be a group homomorphism and $K$ its kernel. If $h \in H$, the fiber over $h$ is $\psi^{-1}(h) := \{ g \in G \mid \psi(g) = h \}$. Two elements $x, y \in G$ belong to the same fiber if and only if $\psi(x) = \psi(y)$ if and only if $xK = yK$. (Prove this.)

$U(1) \cong \mathbb{R}/2\pi \mathbb{Z}$. Picture a rope being coiled up.
3. THE QUOTIENT GROUP CONSTRUCTION

3.0.1. Three formulas and an analogy. Mathematicians love analogies. They illustrate the unity of the universe.

If $U$ and $V$ are subspaces of some larger vector space, $\mathbb{R}^n$ for example, then
\[
\dim(U + V) = \dim U + \dim V - \dim(U \cap V)
\]
If $A$ and $B$ are subsets of some common set, $X$ say, then
\[
|A \cup B| = |A| + |B| - |A \cap B|.
\]
If $A$ and $B$ are events in some probability space, then
\[
P[A \cup B] = P[A] + P[B] - P[A \cap B].
\]

Theorem 3.8 below is analogous to these three formulas. To see the analogy, first rewrite each one as
\[
\dim(U + V) - \dim U = \dim V - \dim(U \cap V)
\]
\[
|A \cup B| - |A| = |B| - |A \cap B|
\]
\[
P[A \cup B] - P[A] = P[B] - P[A \cap B].
\]
Second, observe that $U + V$ is the smallest subspace containing both $U$ and $V$; $A \cup B$ is the smallest set containing both $A$ and $B$; $U \cap V$ is the smallest subspace contained in both $U$ and $V$; $A \cap B$ is the smallest set contained in both $A$ and $B$.

Now consider subgroups $H$ and $N$ of a larger group $G$ and suppose that $N$ is normal in $G$. We make that additional hypothesis on $N$ to ensure that $HN$, which is $\{xa \mid x \in H, a \in N\}$, is a subgroup of $G$. Of course, $HN$ contains both $H$ and $N$, and any subgroup of $G$ that contains both $H$ and $N$ must contain the product $xa$ whenever $x$ is in $H$ and $a$ is in $N$, so contains $HN$. Thus $HN$ is the smallest subgroup of $G$ that contains both $H$ and $N$. (You should check that $HN = NH$.) The largest subgroup of $G$ that is contained in both $H$ and $N$ is, of course, $H \cap N$. The isomorphism in the next theorem is analogous to the three formulas above, even more so if one rewrites the first of those formulas as
\[
\dim\left(\frac{U + V}{U}\right) = \dim\left(\frac{V}{U \cap V}\right).
\]
Theorem 3.8 (Second Isomorphism Theorem). Let $H$ and $N$ be subgroups of a group $G$ and suppose that $N$ is a normal subgroup of $G$. Then

$$
\frac{HN}{N} \cong \frac{H}{H \cap N}
$$

Remark. Before proving the theorem we need to check that its conclusion makes sense. We have already remarked that $HN$ is a subgroup of $G$. It contains $N$ and because $xN = Nx$ for all $x \in G$, $N$ is a normal subgroup of $HN$ and we can therefore construct the quotient group $HN/N$. On the other side of the isomorphism, we must check that $H \cap N$ is a normal subgroup of $H$ to ensure that the quotient $H/H \cap N$ exists. All that does work out, so we will leave it for later to check and get on with the proof of Theorem 3.8.

Proof of Theorem 3.8. Define $f : H \to HN/N$ by $f(x) = xN$.

Lemma 3.9. Let $H$ and $N$ be subgroups of a group $G$. Suppose that $N$ is normal subgroup. Then

1. $HN$ is a subgroup of $G$;
2. $H \cap N$ is a normal subgroup of $H$.

Proof.

Example 3.10. If $G$ and $H$ are finite groups and $(|G|, |H|) = 1$, then there is only one homomorphism $\psi : G \to H$, the trivial map, $\psi(G) = \{e\}$. To see this, observe that $\text{im } \psi$ is a subgroup of $H$ so, by Lagrange’s Theorem, its order divides $|H|$. On the other hand, $\text{im } \psi \cong G/\ker \psi$ so $|\text{im } \psi|$ divides $|G|$ also (see Corollary 2.27). Hence $|\text{im } \psi| = 1$, and we conclude that $\psi(g) = e$ for all $g \in G$.

Example 3.11. We observed in Example 6.10 that the symmetric group $S_{n-1}$ is a subgroup of $S_n$. Suppose that $m \leq n$. Then there are homomorphisms $\psi : S_m \to S_n$. Fix any subset $D \subset \{1, 2, \ldots, n\}$ having $m$ elements, and fix a labelling $D = \{d_1, \ldots, d_m\}$ for the elements of $D$. Define $\psi : S_m \to S_n$ by

$$
\psi(\sigma)(i) = i \text{ if } i \notin D
$$

$$
\psi(\sigma)(d_j) = d_{\sigma(j)} \text{ otherwise.}
$$

You should check that $\psi$ is an injective homomorphism. The image of $\psi$ is the subgroup of $S_n$ consisting of those permutations $\beta$ such that $\beta(i) = i$ for all $i \notin D$. In particular, this subgroup of $S_n$ is isomorphic to $S_m$. Hence $S_n$ contains many copies of $S_m$.

3.1. A short quiz.

1. Write out the definition of a group. Use sentences.
2. Prove that the identity of a group is unique.
3. Prove that an element in a group has a unique inverse.
4. Prove that you can cancel in a group, i.e., if $ab = ac$, then $b = c$. 

(5) Consider the group \((\mathbb{Z}, +)\) of the integers with their usual addition. Is the subset \(57\mathbb{Z}\) consisting of all multiples of 57 a subgroup? Explain.

(6) Consider the set of integers \(\mathbb{Z}\) with the “multiplication”

\[ a \ast b = ab + a + b. \]

Does this make \(\mathbb{Z}\) a group? Explain why.

(7) Consider the set \(G = \mathbb{Q} - \{-1\}\), all fractions except \(-1\), with the “multiplication”

\[ a \ast b = ab + a + b. \]

Does this make \(G\) a group? Explain why.

4. ???

The next result provides a group-theoretic explanation for the importance of exponential functions.

**Lemma 3.12.** Let \(f : (\mathbb{R}, +) \rightarrow (\mathbb{R} - \{0\}, \cdot)\) be a differentiable function. Then \(f\) is a group homomorphism if and only if there is a \(c \in \mathbb{R}\) such that \(f(t) = e^{ct}\).

**Proof.** Certainly, the function \(t \mapsto e^{ct}\) is a group homomorphism. Conversely, if \(f : \mathbb{R} \rightarrow \mathbb{R}^{\times}\) is a differentiable group homomorphism, then

\[ f'(t) = \lim_{h \to 0} \frac{f(t + h) - f(t)}{h} = \lim_{h \to 0} \frac{f(t)}{h} (f(h) - 1) = f'(0) f(t). \]

Hence \(f(t) = re^{ct}\) for some \(r \in \mathbb{R} - \{0\}\) and some \(c\). But \(f(0) = 1\) so \(r = 1\).

There is a similar result for differentiable maps \((\mathbb{R}, +) \rightarrow \text{GL}_n(\mathbb{R})\). First we need the notion of the matrix exponential. If \(A\) is an \(n \times n\) matrix with real entries, we define

\[ e^A := 1 + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots. \]

Of course, we need to see that this sum converges. To see this, let \(M\) be a positive real number that is strictly greater than the absolute value of all entries in \(A\). Then every entry in \(A^i\) has absolute value less than \(M^i\) so, because

\[ \lim_{r \to \infty} \sum_{i=0}^{r} M^i/i! \]

exists, the sum \(\sum_{i=0}^{\infty} A^i/i!\) converges.

**Lemma 3.13.** Let \(f : (\mathbb{R}, +) \rightarrow \text{GL}_n(\mathbb{R})\) be a differentiable function (i.e., each coordinate function is differentiable). Then \(f\) is a group homomorphism if and only if there is an \(n \times n\) matrix \(A\) such that \(f(t) = e^{tA}\).
3. HOMOMORPHISMS BETWEEN GROUPS

**Example 3.14.** A particularly important group homomorphism occurs when

\[ A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \]

In this case

\[ f : \mathbb{R} \to \text{GL}(2, \mathbb{R}), \quad f(t) := \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \]

The proof of Lemma 3.13 shows that \( f \) is a group homomorphism, but one may also prove this as a consequence of the addition formulas for \( \cos(\alpha + \beta) \) and \( \sin(\alpha + \beta) \).

The kernel of \( f \) is \( 2\pi \mathbb{Z} \). If we view \( f(t) \) as the linear map \( \mathbb{R}^2 \to \mathbb{R}^2 \) given by left multiplication, it is rotation by an angle of \( t \) in the counter-clockwise direction.

That \( \det f(t) = 1 \) for all \( t \) is equivalent to the fact that \( \sin^2 t + \cos^2 t = 1 \).

Other remarks. Is every \( 2 \times 2 \) matrix of determinant 1 in the image of \( f \)? Action of \( \mathbb{R} \) via \( f \) on \( \mathbb{R}^2 \), and on the circle of radius \( r \), gives isomorphism \( \mathbb{R}/2\pi \mathbb{Z} \cong S^1 \).

**Pedantry.** One should distinguish between \( \text{GL}(n, \mathbb{R}) \), the group of invertible \( n \times n \) matrices, and \( \text{GL}(\mathbb{R}^n) \), the group of invertible linear transformations \( T : \mathbb{R}^n \to \mathbb{R}^n \). These groups are isomorphic but giving a particular isomorphism requires choosing a particular basis for \( \mathbb{R}^n \).

Should have a section on linear groups

**Proposition 3.15.** Let \( N \) be a normal subgroup of the group \( G \). There is a bijection

\[ \{\text{subgroups } H \text{ of } G \text{ containing } N\} \leftrightarrow \{\text{subgroups } G/N\} \]

\[ H \leftrightarrow H/N. \]

**Proof.**

**Corollary 3.16.** The subgroups of \( \mathbb{Z}/d \) are exactly the \( n\mathbb{Z}/d\mathbb{Z} \) for \( n|d \). Furthermore, if \( d = mn \), then \( n\mathbb{Z}/d\mathbb{Z} \cong \mathbb{Z}/m \).

**Proof.** Combine Propositions 3.15 and 2.17.

**Corollary 3.17.** Every subgroup of a cyclic group is cyclic.

**Proposition 3.18.** Suppose that \( H \) and \( K \) are normal subgroups of \( G \) such that \( G = HK \) and \( H \cap K = \{1\} \). Then \( G \cong H \times K \).
5. Conjugacy

Proof. First we show that the elements in \( H \) commute with elements in \( K \). To see this, let \( x \in H \) and \( y \in K \). Then \( xyx^{-1} \in K \) because \( K \) is normal, so \( xyx^{-1}y^{-1} \in K \). Because \( H \) is normal, \( yx^{-1}y^{-1} \in H \), so \( xyx^{-1}y^{-1} \in H \). But \( H \cap K = \{1\} \) so \( xyx^{-1}y^{-1} = 1 \); i.e., \( xy = yx \).

Define \( \phi : H \times K \to G \) by \( \phi(x, y) = xy \). This is surjective because \( HK = G \). It is a homomorphism because elements in \( H \) commute with elements in \( K \). It is injective because \( H \cap K = \{1\} \). Hence \( \phi \) is an isomorphism. \( \square \)

5. Conjugacy

If \( x \) and \( g \) are elements of \( G \), \( gxg^{-1} \) is called a conjugate of \( x \). It has the same order as \( x \). In fact it behaves like \( x \) in many ways (we will make this more precise later). A key observation when dealing with conjugacy-related issues is that \( (gxg^{-1})(gyg^{-1}) = g(xy)g^{-1} \).

The set of all elements of \( G \) that are conjugate to \( x \) is called a conjugacy class, or the conjugacy class of \( x \). It is denoted by \( C(x) \). Thus

\[
C(x) := \{gxg^{-1} \mid g \in G\}.
\]

You should show that two conjugacy classes are either the same or disjoint. It follows from this that \( G \) is the disjoint union of its conjugacy classes.

We also talk about conjugates of subgroups. If \( H \) is a subgroup and \( g \in G \), we write

\[
gHg^{-1} := \{ghg^{-1} \mid h \in H\}.
\]

This is again a subgroup of \( G \) (check). We call it a conjugate of \( H \).

5.1. Conjugacy classes in the symmetric group \( S_n \). Partitions.

A partition of a positive integer \( n \) is a collection of positive integers \( n_1, \ldots, n_k \) such that \( n_1 + \cdots + n_k = n \). The order of the integers is not important. It is often convenient to denote a partition by writing, for example, \( (1^32^35) \) to denote the partition \( 1, 1, 1, 2, 3, 3, 5 \) of 16. Each element of \( S_n \) determines a partition of \( n \) by taking the size of its orbits.

Lemma 3.19. Two elements of \( S_n \) are conjugate if and only if they determine the same partition of \( n \); that is, if and only if they have orbits of the same size.

Proof. Suppose that \( \sigma \) and \( \tau \) yield the same partition of \( n \). Then, we can write \( \{1, \ldots, n\} \) as a disjoint union in two ways, say

\[
\{1, \ldots, n\} = A_1 \sqcup \cdots \sqcup A_r = B_1 \sqcup \cdots \sqcup B_r,
\]

where \( |A_i| = |B_i| \) for all \( i \), and the elements of each \( A_i \) (resp., each \( B_i \)) consist of a single \( \sigma \)-orbit (resp., \( \tau \)-orbit). Fix elements \( a_i \in A_i \) and \( b_i \in B_i \) for all \( i \). It is obvious that there is an element \( \eta \in S_n \) such that \( \eta(A_i) = B_i \) for all \( i \), and even more precisely \( \eta(\sigma^j(a_i)) = \tau^j(b_i) \) for all \( i \) and \( j \). In particular, \( \eta(a_i) = b_i \), so \( \tau = \eta\sigma\eta^{-1} \).

The converse is obvious. \( \square \)
The bijection between conjugacy classes and partitions is fundamental to the analysis of the symmetric group.

**Example 3.20.** The conjugacy classes in $S_5$ are as follows:

<table>
<thead>
<tr>
<th>Partition</th>
<th>Element in the conjugacy class</th>
<th>Size of conjugacy class</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>(12345)</td>
<td>24</td>
</tr>
<tr>
<td>1, 4</td>
<td>(1234)</td>
<td>30</td>
</tr>
<tr>
<td>2, 3</td>
<td>(12)(345)</td>
<td>20</td>
</tr>
<tr>
<td>1, 1, 3</td>
<td>(123)</td>
<td>20</td>
</tr>
<tr>
<td>1, 2, 2</td>
<td>(12)(34)</td>
<td>15</td>
</tr>
<tr>
<td>1, 1, 1, 2</td>
<td>(12)</td>
<td>10</td>
</tr>
<tr>
<td>1, 1, 1, 1, 1</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>


We will use this list to find the normal subgroups of $S_5$ in Proposition 6.7.

**Some Exercises.**

1. If $H$ and $K$ are subgroups of a group $G$ show that $H \cap K$ is.
2. Give an example to show that the union of two subgroups need not be a subgroup. In particular, give an example inside $(\mathbb{Q} - \{0\}, \cdot)$.
3. There are two groups of order 4. Write out the multiplication table for each of these groups. List all the subgroups of each group (you do not need to list the trivial subgroup $\{e\}$ or the group itself).
4. Show there are groups of orders 2, 3, and 5, by explicitly constructing their multiplication tables.
5. Why does a group whose order is a prime number have no subgroups other than itself and the trivial subgroup?
6. Consider the groups $\mu_n$ of $n^{th}$ roots of unity. When is $\mu_n$ a subgroup of $\mu_m$? Describe all the subgroups of $\mu_n$.
7. Let $x$ and $y$ be elements of the group $G$. Prove that $(x^{-1})^{-1} = x$.
   Prove that $(xy)^{-1} = y^{-1}x^{-1}$.
8. Let $x$ be an element of the group $G$. The order of $x$ is the smallest positive integer $n$ such that $x^n = e$. If there is no such $n$ we say that $x$ has infinite order. Show that the order of $x$ is equal to the order of $gxg^{-1}$ for any $g \in G$.
9. Let $G$ be a group with 729 elements. Does $G$ have an element of order 5? Why?
10. If $x$ has order $n$ show that $\{e, x, x^2, \ldots, x^{n-1}\}$ is a subgroup of $G$ consisting of $n$ distinct elements.
11. Suppose that $x^2 = e$ for all $x \in G$. Show that $G$ is abelian.
12. Suppose that $x \in G$ has order $n > 0$. If $m$ is relatively prime to $n$ (i.e., $\gcd(m, n) = 1$, show that $\langle x^m \rangle = \langle x \rangle$.
13. Suppose that $x \in G$ has order $n > 0$. If $m \in \mathbb{Z}$, what is the order of $x^m$?
14. Suppose that $x \in G$ has order $n > 0$. Show that $x^i = x^j$ if and only if $n$ divides $i - j$. 
(15) Suppose that $x$ and $y$ are elements in $G$ of finite order. If $xy = yx$
find a formula for the order of $xy$ in terms of the orders of $x$ and $y$. 
CHAPTER 4

Actions of groups on sets

Groups typically occur in nature as certain kinds of permutations of a set with structure, and the group consists of permutations that preserve the structure. This is roughly what we mean when we speak of a symmetry group.

We will develop this point of view in this chapter.

1. Actions

There are two equivalent ways of making precise what we mean by an action of a group on a set.

Let $G$ be a group and $X$ a set. An action of $G$ on $X$ is a map $\alpha : G \times X \to X$, that we usually denote by $(g, x) \mapsto gx$ or $g.x$, such that

1. $e.x = x$ for all $x \in X$, and
2. $g.(h.x) = (gh).x$ for all $g, h \in G$ and $x \in X$.

Equivalently, an action of a group $G$ on a set $X$ is a group homomorphism $G \to \text{Aut} X$. If $g \in G$ and $x \in X$ we write $g.x$ for the image of $x$ under the action of $g$. Thus $1.x = x$, and $g.(h.x) = (gh).x$ for all $x \in X$, and all $g, h \in G$.

Notice that $g^{-1}.(gx) = (g^{-1}g)x = ex = x$; the action of $g^{-1}$ is the inverse of that of $g$; i.e., $g^{-1}$ undoes what $g$ does.

Each element $g \in G$ gives a permutation of $X$, i.e., a bijective (= 1-1 and onto) map $X \to X$, $x \mapsto gx$; we call this the action of $g$; the action of $g$ is 1-1 because if $gx = gy$, then $x = ex = g^{-1}gx = g^{-1}gy = ey = y$; and the action of $g$ is onto because if $y \in G$, $y = g.g^{-1}y$.

Example 4.1. 1. A group $G$ acts on itself by left multiplication, $g.x = gx$.

If $|G| = n$, this action gives a homomorphism $G \to S_n$.

Notice that the action of $G$ on itself by right multiplication is not always an action according to our definition because then $g.(h.x) = (hg).x$. It is an action if $G$ is commutative.

However, if we define $g.x = xg^{-1}$, this is an action of $G$ on itself.

2. A group $G$ acts on itself by conjugation: $g.x = gxg^{-1}$. This action defines a group homomorphism $G \to \text{Aut} G$, the automorphism group of $G$. The kernel of this is

$$\{g \in G \mid gx = xg \text{ for all } x \in G\} = Z(G),$$
the center of $G$. The image of this homomorphism is called the group of inner automorphisms of $G$, and is denoted by $\text{Inn}(G)$.

3. If $H$ is a subgroup of $G$, then $G$ acts on the set of right cosets $X = \{xH \mid x \in G\}$ by

$$g \cdot xH := gxH.$$ 

And $G$ acts on the set of left cosets $Y = \{Hy \mid y \in G\}$ by

$$Hy \cdot g := Hyg^{-1}.$$ 

4. A group $G$ acts on the set of its subgroups by conjugation, $g \cdot H = ghg^{-1} = \{gh^{-1}h \mid h \in H\}$. Subgroups $H$ and $H'$ are said to be conjugate if $H' = ghg^{-1}$ for some $g \in G$. Cosets of $H$ by $g \cdot xH = (gx)H$. \hfill $\Box$

2. The dihedral groups

2.1. Regular polygons. A polygon is regular if all its sides have the same length. For example, equilateral triangles, and squares are regular polygons. The Pentagon building in Washington D.C. is a regular pentagon. The honeycomb lattices created by bees are built up from regular hexagons. We also consider two degenerate cases. Even though its area is zero and we do not usually call it a polygon for the purposes of this chapter we will consider a line of finite length to be a regular polygon with two sides, the same side twice. We will also permit a single point to be included as a polygon. It has one side, and that side has length zero.

Often we will picture a regular pentagon as being situated in the Euclidean plane $\mathbb{R}^2$ with its center located at the origin $(0,0)$.

If a polygon has $n$ sides we will call it an $n$-gon.

2.2. Definition of the dihedral groups. Let $V$ be the set of vertices of a regular $n$-gon $P$. Consider all rigid motions of $P$ that send $V$ to $V$. Think of $P$ as a piece of wood that you may pick up, rotate, turn over, and put down again so that it is in its original position, i.e., the vertices are placed on top of the original vertices. The set of all such motions is called the dihedral group $D_n$.

This is a group. The product in $D_n$ is composition of rigid motions: if we denote two rigid motions by $g$ and $h$, we define their product $gh$ to be the rigid motion first do $h$, then do $g$. The identity element is the rigid motion do nothing. The inverse of a rigid motion $g$ is undo $g$.

**Lemma 4.2.** If $n \geq 3$, then $|D_n| = 2n$.

**Proof.** Label two adjacent vertices $A$ and $B$ with $B$ in the clockwise direction from $A$. The position of the $P$ after a rigid motion is completely determined by the new positions of $A$ and $B$. It follows that there is a bijection between

$$\{\text{elements of } D_n\} \longleftrightarrow \{\text{positions of } A \text{ and } B\}.$$ 

There are $n$ possible positions for $A$ and then two possible positions for $B$ so $D_n$ has $2n$ elements. \hfill $\Box$
LEMMA 4.3. $D_n$ is isomorphic to a subgroup of $S_n$.

Proof. We will define an injective homomorphism $\theta : D_n \rightarrow S_n$.

Label the vertices $1, 2, \ldots, n$ in a clockwise sequence. Performing a rigid motion, say $g \in D_n$, moves each vertex to another vertex; if the vertices $1, 2, 3, \ldots$ are moved to the vertices labelled $a, b, c, \ldots$ we define

$$\theta(g) = \begin{pmatrix} 1 & 2 & 3 & \ldots \\ a & b & c & \ldots \end{pmatrix}.$$ 

It is virtually a tautology to verify that $\theta$ is an injective group homomorphism. \hfill \Box

2.3. Small dihedral groups. Let’s look at the case $n = 1$. Then $P$ consists of a single vertex (and no sides at all). It is clear that the only rigid motion is the trivial motion, do nothing. So $D_1 = \{e\}$, the trivial group.

When $n = 2$, $P$ is a single interval, with two endpoints. Apart from the trivial rigid motion, there is only one other rigid motion, simply flip the interval around. Thus $D_2$ is the group with two elements.

When $n = 3$, $P$ is an equilateral triangle. It should be clear to you that given any permutation $[a, b, c]$ of $[1, 2, 3]$, there is a rigid motion of the equilateral triangle that sends the vertex labelled 1 to that labelled $a$, and simultaneously sends 2 to $b$, and 3 to $c$. Thus, every permutation of $[1, 2, 3]$ can be realized by a rigid motion. In other words, the map $\theta : D_3 \rightarrow S_3$ is surjective (onto) as well as injective; hence we have an isomorphism $D_3 \cong S_3$.

Labelling the elements of $D_n$. If clockwise rotation by $2\pi/n$ radians is denoted by $\tau$, and the flip about some fixed axis is denoted by $\sigma$, then $D_n$ is generated by $\sigma$ and $\tau$. Now $\tau^n = \sigma^2 = 1$, and $\sigma \tau \sigma^{-1} = \tau^{-1}$.

THEOREM 4.4. If $n \geq 3$, then $|D_n| = 2n$ and

$$D_n = \{\tau^i, \sigma \tau^i \mid 0 \leq i \leq n - 1\}.$$ 

Proof. We have already seen that $|D_n| = 2n$, so we must simply show that the $2n$ elements listed are all different from one another.

Label the vertices $1, 2, \ldots, n$ in a clockwise sequence; an element $g \in D_n$ is completely determined by the position of 1 and 2 after the motion. We can assume that the flip $\sigma$ sends $n$ to 1, $n - 1$ to 2, $n - 2$ to 3 and, more generally, $j$ to $n - j + 1$.

Now $\tau^i$ sends $n$ to $i$ and 1 to $i + 1$ (modulo $n$), whereas $\sigma \tau^i$ sends $n$ to $n - i + 1$ and 1 to $n - i$. It follows that these elements are different from one another when $n \geq 3$.

When $n = 2$, $\sigma = \tau$. When $n = 1$, $\sigma = \tau = \text{id}$. \hfill \Box

Perhaps a better way to prove this theorem is to observe that because $|D_n| = 2n$ and $\tau$ has order $n$, the subgroup $\langle \tau \rangle$ has index 2 in $D_n$ and is therefore normal by Lemma 3.2. This tells us that $\sigma \tau \sigma^{-1}$ is a power of $\tau$ and, because $\sigma \langle \tau \rangle = \langle \tau \rangle \sigma$,

$$D_n = \{1, \tau, \tau^2, \ldots, \tau^{n-1}\} \cup \{\sigma, \tau \sigma, \tau^2 \sigma, \ldots, \tau^{n-1} \sigma\}.$$
2.3.1. Generators and relations.

**Proposition 4.5.** Let \( p \) be an odd prime. If \( G \) is a non-abelian group of order \( 2p \), then \( G \cong D_p \).

**Proof.** If all non-identity elements of a group have order 2 that group is abelian because \( 1 = (xy)^2 = xyy \) implies \( xy = y^{-1}x^{-1} = yx \). Hence \( G \) has an element \( x \neq 1 \) whose order is not 2. If \( x \) had order 2p, then \( G \) would equal \( \langle x \rangle \) and would therefore be abelian. Thus, by Lagrange’s Theorem, the order of \( x \) is \( p \). Hence the subgroup \( \langle x \rangle \) is normal.

Pick \( a \notin \langle x \rangle \). Since \( G = \langle a, x \rangle \) is not abelian, \( xa \neq ax \). Thus \( axa^{-1} = x^i \) for some integer \( i \) with \( 2 \leq i \leq p - 1 \). Since \( G/\langle x \rangle \cong \mathbb{Z}_2 \), \( a^2 \in \langle x \rangle \) so

\[
x = a^2xa^{-2} = a(axa^{-1})a^{-1} = ax'a^{-1} = (axa^{-1})^i = (x^i)^i = x^{i^2}
\]

whence \( x^{i^2-1} = 1 \). Hence \( p|i^2 - 1 \). But \( i^2 - 1 = (i - 1)(i + 1) \) and \( p \nmid i - 1 \) so \( p|i + 1 \). Hence \( i = p - 1 \) and \( axa^{-1} = x^{p-1} = x^{-1} \).

Since \( a^2 \notin \langle x \rangle \), if \( a^2 \) were \( \neq 1 \), then \( a^2 \) would have order \( p \) and \( a \) would therefore have order \( 2p \) and \( G \) would equal \( \langle a \rangle \) contradicting the fact that \( G \) is not abelian. Hence \( a^2 = 1 \) and \( G \cong D_p \).

\( \Box \)

3. Orbits and stabilizers

**Definition 4.6.** Let \( G \) be a group acting on a set \( X \). The orbit of \( x \in X \) is \( G.x = \{ g.x \mid g \in G \} \). The stabilizer of \( x \) is \( \text{Stab}_G(x) = \{ g \in G \mid g.x = x \} \).

We also use the notation \( G_x \) for \( \text{Stab}_G(x) \).

Notice that the stabilizer of \( x \) is a subgroup of \( G \).

**Lemma 4.7.** Let a group \( G \) act on a set \( X \). Then \( X \) is the disjoint union of its orbits.

**Proof.** Since each \( x \) in \( X \) belongs to some orbit, namely \( G.x \) because \( x = e.x \), \( X \) is certainly the union of its orbits. To show that \( X \) is the disjoint union of its orbits we must show that if two orbits intersect then they are, in fact, the same. Suppose that \( z \in Gx \cap Gy \). Then \( z = gx = hy \) for some \( g, h \in G \). If \( b \in G \), then

\[
bx = bg^{-1}z = bg^{-1}hy \in Gy;
\]

thus \( Gx \subseteq Gy \); similarly, \( Gy \subseteq Gx \); thus \( Gx = Gy \). \( \Box \)

This provides an equivalence relation on \( X \),

\[
x \sim y \iff y \in G.x \iff G.x = G.y.
\]

**Proposition 4.8.** Let \( G \) be a finite group acting on a set \( X \). If \( x \in X \), then

\[
(1) \quad |G| = |G.x| \times |G_x|;
(2) \quad |G.x| = |G : G_x|;
(3) \quad |G.x| \text{ divides } |G|.
\]
Proof. Let \( g, h \in G \). Then
\[
g \cdot x = h \cdot x \Leftrightarrow g^{-1}h \cdot x = x \Leftrightarrow g^{-1}h \in G_x \Leftrightarrow gG_x = hG_x.
\]
Therefore \( |G \cdot x| \) is equal to the number of cosets of \( G_x \) in \( G \); that number is
\[
|G : G_x| = |G| / |G_x|.
\]

The following is a trivial consequence, but its triviality belies its significance.

**Corollary 4.9 (The Orbit Formula).** Let \( G \) be a finite group acting on a finite set \( X \). Let \( X_1, \ldots, X_n \) be the distinct \( G \)-orbits in \( X \), and for each \( i \) choose \( x_i \in X_i \). Then
\[
|X| = \sum_{i=1}^{n} |X_i| = \sum_{i=1}^{n} |G : \text{Stab}_G(x_i)|.
\]

Notice that if \( G \) acts on \( X \) and \( x \) and \( y \) belong to the same orbit, then their stabilizers are conjugate: if \( y = g \cdot x \), then \( G_y = gG_x g^{-1} \). Conversely, if two subgroups of \( G \) are conjugate to one another, then one is a stabilizer if and only if the other is.

**Proposition 4.10 (Burnside’s formula).** \(^1\) Let \( G \) be a finite group acting on a finite set \( X \). For each \( g \in G \), define
\[
X^g := \{ x \in X \mid g \cdot x = x \}.
\]
Then
\[
\sum_{x \in X} |G_x| = \sum_{g \in G} |X^g|
\]
and
\[
|G| \cdot \text{(the number of orbits)} = \sum_{x \in X} |G_x|.
\]

Proof. Picture \( G \times X \) as rows indexed by \( g \in G \), each row being a copy of \( X \). We can also view this as columns indexed by \( x \in X \), each column being a copy of \( G \). Thus \((g, x)\) appears in the row indexed by \( g \) and the column indexed by \( x \). Construct a True/False table by placing a \( T \) in the cell labelled \((g, x)\) if \( g \cdot x = x \). Computing the number of \( T \)s in two ways gives (4-1): the number of \( T \)s in the row \( \{g\} \times X \) is \( |X^g| \), and the number of \( T \)s in the column above \( x \) is \( |G_x| \).

Now,
\[
\sum_{x \in X} |G_x| = |G| \cdot \sum_{x \in X} \frac{1}{|G_x|}
\]
\(^1\)There are many mis-attributions in mathematics, and this is one of long standing. In his book Burnside attributed this result to Frobenius, but Cayley knew it long before Frobenius. The only explanation I can think of for naming it after Burnside is because it appears in his book and that became a standard reference.
but the orbit of \( x \) gets counted exactly \( |G.x| \) times in this sum, so the right-hand side is exactly \( |G| \) times the number of distinct orbits. \( \square \)

**Example 4.11.** How many different flags can we make if we have 6 stripes and 3 colors of cloth? The difficulty in counting arises because we can turn a flag over to get a “new” flag, e.g., the flag with stripes colored \( ababab \) is the same as the flag with stripes colored \( ababaa \).

Let \( X \) be the set of flags where we consider the flags \( ababab \) and \( ababaa \) as different. Now the group \( G = \{1, \tau\} \cong \mathbb{Z}_2 \) acts on \( X \) where \( \tau \) is the operation “turn the flag over”. Thus \( \tau.ababab = ababaa \).

The number of orbits is the number of different flags.

We use Burnside’s formula in the form

\[
\text{the number of orbits} = \frac{1}{|G|} \sum_{g \in G} |X^g|.
\]

Since \( X \) is very large and \( G \) is very small it makes good sense to take a sum over the elements of \( G \) rather than a sum over the elements of \( X \), so we use the formula

\[
\text{the number of orbits} = \frac{1}{|G|} \sum_{g \in G} |X^g|.
\]

Hence we need to count the number of \( x \in X \) that are fixed by each \( g \in G \). Obviously \( X^1 = X \), and

\[
X^\tau = \{ x \mid \tau.x = x \} = \{ \text{flags } abcd ef \mid abcd ef = fedc ba \} = \{ \text{flags } abccba \}.
\]

Now there are 3 choices for each of \( a, b, \) and \( c \), and once those are chosen the flag \( abccba \) is determined; hence there are \( 3^3 \) flags \( abccba \), and \( |X^\tau| = 27 \).

Since \( |X| = 3^6 \), we have

\[
\text{the number of orbits} = \frac{1}{2}(3^6 + 3^3) = 378.
\]

\( \diamond \)

**3.1. Necklaces.** A standard illustration of Burnside’s formula is to count the number of necklaces that can be made from \( n \) beads having up to \( r \) different colors. Two necklaces are considered the same if one can be rotated to become the other. Thus there is an action of the cyclic group \( \mathbb{Z}_n \) on the set of necklaces with \( n \) beads and the number of genuinely different necklaces is the number of orbits under the action of \( \mathbb{Z}_n \).

**4. Applications to finite groups**

**4.1. The conjugation action.** We consider the conjugation action of \( G \) on itself: \( g \in G \) acts on \( x \in G \) according to the formula

\[
g \cdot x := gxg^{-1}.
\]

The **conjugacy class** of \( x \in G \) is

\[
C_G(x) = \{ gxg^{-1} \mid g \in G \},
\]
and the centralizer of \( x \in G \) is
\[
Z_G(x) = \{ g \in G \mid gx = xg \}.
\]
Thus \( C_G(x) \) is the orbit of \( x \), and \( Z_G(x) \) is the stabilizer of \( x \) for the conjugation action of \( G \) on itself. Stabilizers are always subgroups so \( Z_G(x) \) is a subgroup of \( G \).

### 4.1.1. Important observation.
Recall that the center of \( G \) is the subgroup
\[
Z(G) := \{ x \in G \mid gx = xg \text{ for all } g \in G \}.
\]
Thus \( x \in Z(G) \) if and only if \( |C_G(x)| = 1 \). In words, \( x \) commutes with everything in \( G \) if and only if its conjugacy class has size one.

The next result follows at once from the definition and Proposition

**Lemma 4.12.** Let \( x \) be an element of a finite group \( G \). Then
\[
|G| = |Z_G(x)| \times |C_G(x)|.
\]
In particular, the number of conjugates of \( x \) equals \( |G : Z_G(x)| \), which divides \( |G| \).

**Proof.** This is a special case of the formula \( |G| = |G|/|\text{Stab}(x)| \) in Proposition 4.8. \( \square \)

### 4.1.3 (The Class Formula).
Let \( G \) be a finite group and let \( C_1, \ldots, C_n \) be the distinct conjugacy classes in \( G \). Then
\[
|G| = \sum_{i=1}^n |C_i|.
\]
(1) \( |G| = \sum_{i=1}^n |C_i| \);
(2) If \( Z(G) \) denotes the center of \( G \), then
\[
|G| = |Z(G)| + \sum_{|C_i|>1} |C_i|.
\]
(4-6)

**Proof.** Since \( G \) is the disjoint union of its orbits, \( |G| = \sum_{i=1}^n |C_i| \) and this can be written as
\[
|G| = \sum_{|C_i|=1} |C_i| + \sum_{|C_i|>1} |C_i|.
\]
However, \( Z(G) \) is the disjoint union of those \( C_i \) having cardinality one. \( \square \)

### 4.1.4. If the order of a group is a power of a prime, then its center is non-trivial; i.e., it contains a non-identity element.

**Proof.** Suppose that \( |G| = p^r \). If \( C_1, \ldots, C_m \) are the conjugacy classes with more than one element, then \( |C_i| = p^{r_i} \) for some \( r_i > 1 \). It therefore follows from (4-6) that \( p \) divides \( |Z(G)| \). \( \square \)

The next proof uses the “obvious” fact that if \( a, b, c, \ldots \) are elements in \( G \) that commute with each other then the subgroup they generate is abelian.

### Corollary 4.15. If \( p \) is a prime and \( |G| = p^2 \), then \( G \) is abelian.
4. ACTIONS OF GROUPS ON SETS

**Proof.** The center $Z$ of $G$ is not $\{1\}$, so has order either $p$ or $p^2$. If it is $p^2$, we are done, so suppose that $|Z| = p$. Let $z \in Z - \{1\}$ and $x \in G - Z$. Because $|Z| = p$, $Z \cong \mathbb{Z}/p\mathbb{Z}$ so $\langle z \rangle = Z$. The subgroup $\langle z, x \rangle$ contains $\langle z \rangle = Z$ and is strictly larger than $Z$ because it contains $x$ which is not in $Z$. Thus $|\langle z, x \rangle| > p$ and divides $p^2$ so must equal $p^2$. Hence $\langle z, x \rangle = G$. It follows that $G$ is abelian.

**Corollary 4.16.** If $p$ is a prime and $|G| = p^2$, then $G$ is isomorphic to either $\mathbb{Z}/p^2\mathbb{Z}$ or $\mathbb{Z}/p \times \mathbb{Z}/p$.

**Proof.** Use the classification of finite abelian groups.

**Theorem 4.17 (Cauchy’s Theorem).** Let $p$ be a prime. If $p$ divides the order of $G$, then $G$ has an element of order $p$.

**Proof.** We argue by induction on the order of $G$.

The result is true for a group of order $p$ because such a group is isomorphic to $\mathbb{Z}/p\mathbb{Z}$. If $H$ is a proper subgroup of $G$ such that $|H| < |G|$ and $p$ divides $|H|$, then we may apply the induction hypothesis to $|H|$ to obtain the result.

Suppose that $G$ is not abelian. Let $x \in G - Z(G)$; then the centralizer $C = \{g \in G \mid gxg^{-1} = x\}$ is a proper subgroup of $G$ so the result holds if $p$ divides $|C|$; so we can assume $p$ does not divide $|C|$; by Lemma 4.12, $p$ divides the order of the conjugacy class $\{gxg^{-1} \mid g \in G\}$. It now follows from (4-6) that $p$ divides $|Z(G)|$. But we are assuming that $Z(G) \neq G$, so the result follows from the induction hypothesis.

Now suppose $G$ is abelian. By the structure theorem for finite abelian groups $G$ is isomorphic to a direct product of cyclic abelian groups. Each of those cyclic abelian groups is isomorphic to a subgroup of $G$ so it suffices to show that a cyclic abelian group whose order is divisible by $p$ has an element of order $p$.

If $G$ is cyclic, then $G = \langle x \rangle$ for some $x$. The order of $x$ is $|G| = pn$ so $x^n$ has order $p$.

**PAUL** Turn the following into an exercise with hints.

Suppose we do not know the structure theorem for finite abelian groups. We can still show that a finite abelian group whose order is divisible by $p$ has an element of order $p$.

If $G$ is cyclic we can argue as in the last paragraph of the proof of Theorem 4.17 so suppose $G$ is abelian but not cyclic. We can choose a proper subgroup $H$ of largest possible order; if $p$ divides $m := |H|$, the theorem follows by the induction hypothesis; so we can assume that $p$ does not divide $m$; Since $p$ divides $G/H$, we may apply the induction hypothesis to $G/H$ to deduce that it has an element of order $p$, say $xH$. Thus $H = \langle xH \rangle^p = x^pH$, so $x^p \in H$. The maximality of $H$ implies that $\langle H, x \rangle = G$, so $G/H = \langle xH \rangle \cong \mathbb{Z}/p$. Thus, $|G| = pm$.

If $x^m \neq 1$, then

$$(x^m)^p = x^{mp} = x^{|G|} = 1,$$
5. FRACTIONAL LINEAR TRANSFORMATIONS

so $x^m$ has order $p$ so the theorem holds in this case.

Finally suppose that $x^m = 1$. By Theorem 1.6, there are elements $a, b \in \mathbb{Z}$ such that $ap + bm = 1$, so

$$x = x^{ap+bm} = (x^m)^a(x^p)^b \in H,$$

contradicting our choice of $x$.

5. Fractional linear transformations

The group $GL(2, \mathbb{C})$ (almost) acts on the complex plane by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z := \frac{az + b}{cz + d}$$

We say that $GL(2, \mathbb{C})$ acts by fractional linear transformations. You should check that this really is an action! But why did I say “almost”? The matrix

$$\begin{pmatrix} -1 & i \\ i & 1 \end{pmatrix}$$

has the remarkable property that it maps the upper half-plane

$$\mathbb{H} := \{ z \in \mathbb{C} \mid Im(z) > 0 \}$$

which is unbounded to the unit disc

$$\mathbb{D} := \{ z \in \mathbb{C} \mid |z| < 1 \}$$

which is bounded.

The subgroup

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \cong \mathbb{C}^\times$$

acts trivially on $\mathbb{C}$ by which I mean that it is in the stabilizer of every point in $\mathbb{C}$.

1. Show that no other elements in $GL(2, \mathbb{C})$ act trivially.
2. Determine the stabilizer of every point in $\mathbb{C}$.
3. Is the action of $GL(2, \mathbb{C})$ on $\mathbb{C}$ transitive? i.e., does $\mathbb{C}$ form a single orbit under the action of $GL(2, \mathbb{C})$?
4. Give a succinct and elegant geometric description of the image of the upper half-disc $\mathbb{H} \cap \mathbb{D}$ under the action of

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$
4. ACTIONS OF GROUPS ON SETS

(3) This is biggish question. Make your answer succinct and clear. Polish your initial answer. Write good prose.

Show that $S_4$ has a normal subgroup, $N$ say, that is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$, and that $S_4/N \cong S_3$ (use Theorem 3.7).

The way I want you to do this is to show that $S_4$ acts on a set $X$ with three elements. I will now tell you what $X$ is and what the action is. Each element of $X$ is a pair of subsets of $\{1, 2, 3, 4\}$. The three elements of $X$ are

$$x = \{\{1, 2\}, \{3, 4\}\}, \quad y = \{\{1, 3\}, \{2, 4\}\}, \quad z = \{\{1, 4\}, \{2, 3\}\}.$$  

Each $\sigma \in S_4$ permutes the elements of $\{1, 2, 3, 4\}$, so if $\{\{a, b\}, \{c, d\}\} \in X$, I define

$$\sigma.\{\{a, b\}, \{c, d\}\} := \{\{\sigma(a), \sigma(b)\}, \{\sigma(c), \sigma(d)\}\}.$$  

Explain why the action of $S_4$ on $X$ induces a group homomorphism $\psi: S_4 \rightarrow S_3$. Tell me explicitly what $\psi$ is. Show that $\psi$ is surjective and compute its kernel.

For each conjugacy class $C$ in $S_4$ choose one element $g \in C$ and determine $X^g$.

Determine the stabilizer of each element of $X$.

(4) Find all possible homomorphisms $\phi: D_4 \rightarrow \mathbb{Z}_4$.

(5) Prove that every normal subgroup of $G$ is a disjoint union of conjugacy classes.

(6) Let $S_4$ act on $\{1, 2, 3, 4\}$ in the natural way. What is the stabilizer of $\{2\}$? What is the stabilizer of $\{4\}$? Show that these two stabilizer subgroups are isomorphic by giving an explicit isomorphism (i.e., do not use the next problem).

(7) More generally, show that if a group $G$ acts on a set $X$ and $x$ and $y$ belong to the same orbit, then the stabilizer of $x$ is isomorphic to the stabilizer of $y$.

(8) Let $S_4$ act on the subsets of $\{1, 2, 3, 4\}$. What is the stabilizer of $\{1, 2\}$? What is the orbit of $\{1, 2\}$?

(9) Let $X$ be the set of all subgroups of $G$. Show that the rule

$$gH := ghg^{-1}, \quad H \in X, \text{ i.e., } H \text{ is a subgroup of } G, \text{ and } g \in G$$

is an action of $G$ on $X$.

Show that $H \subseteq \text{Stab}_G(H)$? Show that $H$ is normal if and only if $\text{Stab}_G(H) = G$.

(10) Let $G$ be a finite group and $N$ a subgroup such that $p = [G : N]$ is the smallest prime that divides $|G|$. Show that $N$ is normal in $G$. Hint: use the previous exercise.

(11) Let $G$ be a finite group. Show that $G$ is isomorphic to a subgroup of some symmetric group $S_n$ by viewing $G$ as acting on itself by left multiplication: that is, write $X = G$ and let $G$ act on $X$ by $g.x = gx$ for $g \in G$ and $x \in X$; check that this really is an action
according to our definition; how does this action show us that $G$ is isomorphic to a subgroup of some $S_n$?
Small Groups

A major problem in every branch of mathematics is the Classification Problem: classify the objects being studied. One is only concerned about classification up to isomorphism. It most branches of mathematics this is an unattainable goal. Still, the problem is wonderfully fruitful because to tackle the problem one must develop methods and techniques that are valuable far beyond their application to the classification problem.

For example, a sensible subproblem is the classification of finite abelian groups, and this can be settled by elementary methods as we shall see later.

In this chapter we classify some groups of small order and in doing so develop a few results of wider applicability.

To classify all groups of order \( n \) we must make a list of groups, each of order \( n \), such that no two on the list are isomorphic to one another, and such that if \( G \) is any group of order \( n \) it is isomorphic to one of those on our list.

All simple groups of finite order have been classified. Remember that a group \( G \) is simple if its only normal subgroups are itself and \( \{1\} \). This classification was achieved around 1980 and the its proof requires some 10,000 pages that represented the cumulative work of hundreds of mathematicians over the preceding century. There are infinitely many finite simple groups. Most fall into one of a finite number of infinite families. For example, the cyclic groups of prime order \( \mathbb{Z}_p \) form one such infinite family, and the alternating groups \( A_n \) for \( n \geq 5 \) form another. There are several other infinite families, each of which is more complicated than the two families just mentioned. There are also 26 finite simple groups that do not belong to an infinite family but seem to exist for reasons unique unto themselves. These are called the sporadic groups.

The only group of order one is the trivial group, \( \{1\} \), consisting of the identity element and nothing more.

By ???, if \( p \) is a prime number there is only one group of order \( p \) up to isomorphism, the cyclic group \( \mathbb{Z}_p \). Hence \( \mathbb{Z}_2 \) and \( \mathbb{Z}_3 \) are the only groups of order 2 and 3 respectively.
5. SMALL GROUPS

1. Groups of order 4

The abelian groups $\mathbb{Z}_2 \times \mathbb{Z}_2$ and $\mathbb{Z}_4$ are not isomorphic because $\mathbb{Z}_4$ has an element of order 4 but $\mathbb{Z}_2 \times \mathbb{Z}_2$ does not.

To see that these are the only groups it suffices to show that every group of order 4 is abelian. If $G$ is a group of order 4 having an element of order 4 it is isomorphic to $\mathbb{Z}_4$. So, suppose $G$ has order 4 and no element of order 4. Since the order of an element divides the order of the group, every non-identity element has order two. It will follow from the next result that $G$ is abelian.

Proposition 5.1. Let $G$ be a group in which every non-identity element has order two. Then $G$ is abelian. In fact, $G \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \times \cdots \times \mathbb{Z}_2$.

Proof. Let $x, y \in G$. Then $1 = x^2 = y^2 = (xy)^2 = xyxy$. Hence $xy = x1y = x(xyxy)y = x^2yxy^2 =yx$. Thus $xy = yx$, so $G$ is abelian.

2. Groups of order 6

3. Groups of order 8

4. Groups of order 9

The next result shows that there are only two groups of order 9, $\mathbb{Z}_3 \times \mathbb{Z}_3$ and $\mathbb{Z}_9$.

Proposition 5.2. Let $p$ be a prime number. The only groups of order $p^2$ are $\mathbb{Z}_p \times \mathbb{Z}_p$ and $\mathbb{Z}_{p^2}$.

Proof. Suppose $|G| = p^2$. By Corollary 4.15 $G$ is abelian. If $G$ has an element of order $p^2$ it is isomorphic to $\mathbb{Z}_{p^2}$ so suppose $G$ does not have such an element. Let $1 \neq x \in G$. Then $\langle x \rangle \cong \mathbb{Z}_p$. Pick any $y \notin \langle x \rangle$ and write $H = \langle y \rangle$ and $K = \langle x \rangle$. Then $H \cong \mathbb{Z}_p$ too, and $H \cap K = \{1\}$. Because $G$ is abelian, $HK$ is a subgroup; it is strictly larger than $H$ so must equal $G$. It now follows from Proposition 3.18 that $G \cong H \times K \cong \mathbb{Z}_p \times \mathbb{Z}_p$.

5. Groups of order 10

6. Groups of order 12

Some Exercises.

(1) Show that the subgroup of $D_n$ consisting of the powers of $\tau$ is a normal subgroup.

(2) Is the subgroup of $D_n$ generated by $\sigma$ normal?

(3) View $D_4$ as a subgroup of $S_4$. Is it a normal subgroup?

(4) Find all subgroups of $D_n$.

In the next few exercises we write $\overline{i}$ for the element of $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ that is the coset containing $i$. 
(5) Which elements of \( \mathbb{Z}_{16} \) belong to \( \langle 6 \rangle \). What are the elements in each coset of \( \langle 6 \rangle \)?

(6) Repeat the previous exercise but now for the cosets of \( \langle 4 \rangle \).

(7) Is \( \mathbb{Z}_2 \times \mathbb{Z}_4 \) isomorphic to \( \mathbb{Z}_8 \)? Why?

(8) Is \( \mathbb{Z}_3 \times \mathbb{Z}_4 \) isomorphic to \( \mathbb{Z}_{12} \)? Why?

(9) Find the order of the element \( (6, 5, 4, 3) \) in \( G = \mathbb{Z}_4 \times \mathbb{Z}_5 \times \mathbb{Z}_6 \times \mathbb{Z}_7 \).

(10) Is \( \mathbb{Z}_8 \) generated by \( \{4, 6\} \)? Is it generated by \( \{4, 5, 6\} \)? Give reasons.
The symmetric groups, again

1. Conjugacy classes

1.1. Partitions. A partition of a positive integer $n$ is a collection of positive integers $n_1, \ldots, n_k$ such that $n_1 + \cdots + n_k = n$. The order of the integers is not important. It is often convenient to denote a partition by writing, for example, $(1^32^35)$ to denote the partition $1, 1, 2, 3, 5$ of 16.

Each element of $S_n$ determines a partition of $n$ by taking the size of its orbits. For example, the orbits associated to the element

$$g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 4 & 1 & 5 & 6 & 3 & 7 & 9 & 8 \end{pmatrix}$$

are $\{1, 2, 4, 5, 6, 3\}$, $\{7\}$, $\{8, 9\}$, so the partition of 9 associated to $g$ is $(1, 2, 6)$.

**Lemma 6.1.** Two elements of $S_n$ are conjugate if and only if they determine the same partition of $n$; that is, if and only if they have orbits of the same size.

**Proof.** Suppose that $\sigma$ and $\tau$ yield the same partition of $n$. Then, we can write $\{1, \ldots, n\}$ as a disjoint union in two ways, say

$$\{1, \ldots, n\} = A_1 \sqcup \ldots \sqcup A_r = B_1 \sqcup \ldots \sqcup B_r,$$

where $|A_i| = |B_i|$ for all $i$, and the elements of each $A_i$ (resp., each $B_i$) consist of a single $\sigma$-orbit (resp., $\tau$-orbit). Fix elements $a_i \in A_i$ and $b_i \in B_i$ for all $i$. It is obvious that there is an element $\eta \in S_n$ such that $\eta(A_i) = B_i$ for all $i$, and even more precisely $\eta(\sigma^j(a_i)) = \tau^j(b_i)$ for all $i$ and $j$. In particular, $\eta(a_i) = b_i$, so $\tau = \eta \sigma \eta^{-1}$.

The converse is obvious. \qed

The bijection between conjugacy classes and partitions is fundamental to the analysis of the symmetric group.
Example 6.2. The conjugacy classes in $S_5$ are as follows:

<table>
<thead>
<tr>
<th>Partition</th>
<th>Element in the conjugacy class</th>
<th>Size of conjugacy class</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>(12345)</td>
<td>24</td>
</tr>
<tr>
<td>1, 4</td>
<td>(1234)</td>
<td>30</td>
</tr>
<tr>
<td>2, 3</td>
<td>(12)(345)</td>
<td>20</td>
</tr>
<tr>
<td>1, 1, 3</td>
<td>(123)</td>
<td>20</td>
</tr>
<tr>
<td>1, 2, 2</td>
<td>(12)(34)</td>
<td>15</td>
</tr>
<tr>
<td>1, 1, 1, 2</td>
<td>(12)</td>
<td>10</td>
</tr>
<tr>
<td>1, 1, 1, 1, 1</td>
<td>1</td>
<td>1.</td>
</tr>
</tbody>
</table>

We will use this list to find the normal subgroups of $S_5$ in Proposition 6.7.

1.2. Generating $S_n$. The next result is useful for recognizing when a subgroup of the symmetric group is actually the whole group.

Proposition 6.3. $S_n$ is generated by $(12)$ and $(12\cdots n)$.

Proof. Let $H$ be the subgroup generated by $a = (12)$ and $b = (12\cdots n)$. Then $H$ contains $bab^{-1} = (23)$ and hence, by induction, $(ii + 1)$ for all $i$. Thus $H$ contains $(12)(23)(12) = (13)$ and $(13)(34)(13) = (14)$, and so on. That is, $(1i) \in H$ for all $i$. Hence, if $i \neq j$, $H$ contains $(1i)(1j)(1i) = (ij)$. Since every element of $S_n$ is a product of transpositions we conclude that $H = S_n$, as claimed.

1.3. Even and odd permutations.

Lemma 6.4. The identity in $S_n$ cannot be written as a product of an odd number of transpositions.

Proof. We will show that if $\sigma \in S_n$ and $\tau = (ij)$ is a transposition, then the number of orbits for $\sigma$ and $\tau \sigma$ differ by one. Since the identity has exactly $n$ orbits, a product of an odd number of transpositions cannot have $n$ orbits, so the lemma follows.

We write $\sigma = \sigma_1 \sigma_2 \cdots \sigma_t$ as a product of disjoint cycles, and order them so that neither $i$ nor $j$ appears in any of the cycles $\sigma_3, \ldots, \sigma_t$. Notice that $t$ is the number of orbits of $\sigma$ having size $> 1$. We write $\ell(\sigma_i)$ for the the length of $\sigma_i$, which is the same as the size of the orbit of $\sigma$ that corresponds to $\sigma_i$. Hence the number of orbits of $\sigma$ is

$$t + n - \ell(\sigma_1) - \cdots - \ell(\sigma_t).$$

Case 1. Suppose that both $i$ and $j$ are fixed by $\sigma$.

Theorem 6.5. If $\sigma$ is written as a product of transpositions in two different ways, say $\sigma = \alpha_1 \cdots \alpha_m = \beta_1 \cdots \beta_r$, then $m \equiv r \pmod 2$.

Proof. Since

$$1 = \sigma \sigma^{-1} = \alpha_1 \cdots \alpha_m \beta_r^{-1} \cdots \beta_1^{-1},$$

Lemma 6.4 implies that $m + r$ is even, and hence $m \equiv r \pmod 2$. 

\[\square\]
A permutation is **even** if it is a product of an even number of transpositions, and is **odd** if it is a product of an odd number of transpositions. The previous lemma ensures that this definition is unambiguous. The set of even permutations form a subgroup of $S_n$ called the **alternating group** and denoted by $A_n$.

**Proposition 6.6.** There is a group homomorphism

$$\text{sgn} : S_n \to \{\pm 1\}$$

defined by

$$\text{sgn}(\sigma) = \begin{cases} +1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$$

The kernel of this homomorphism is obvious: the alternating group $A_n$. Hence $A_n$ is a normal subgroup of $S_n$ of index 2.

We now use the list of conjugacy classes in Example 6.2 to find the normal subgroups of $S_5$.

**Proposition 6.7.** The only normal subgroups of $S_5$ are $A_5$, $S_5$, and $\{1\}$.

**Proof.** Let $H$ be a normal subgroup that is neither $S_5$ nor $\{1\}$. Since $gHg^{-1} \subset H$ for every $g \in S_5$, $H$ is a union of conjugacy classes. The conjugacy classes have sizes 1, 10, 15, 20, 24, and 30. The class of size one must belong to $H$ because it consists of the identity element.

The order of $H$ is a divisor of $|S_5| = 120$. Since 120 is not divisible by 1 + 10, or 1 + 15, or 1 + 20, or 1 + 24, or 1 + 10 + 15, or 1 + 30, the only possibilities for $|H|$ are 40 = 1 + 15 + 24 and 60 = 1 + 15 + 24 + 20.

If $|H| = 40 = 24 + 15 + 1$, then $H$ contains every 5-cycle, so contains (12345), and contains every element corresponding to the partition 1, 2, 2, so contains (12)(34). Hence $H$ contains their product (135). This implies that $|H| > 40$. So $|H| = 40$ is impossible.

Thus $|H| = 60$ and $H$ contains the conjugacy classes of (12345), (12)(34), and (135). But the union of these conjugacy classes is $A_5$. Hence $H = A_5$. \hfill \square

**Lemma 6.8.** $A_n$ is generated by $(123), (124), \ldots, (12n)$.

**Proof.** By its very definition $A_n$ is generated by the elements $(ab)(cd) = (cad)(abc)$ and $(ac)(ab) = (abc)$. So it suffices to show that each $(abc)$ belongs to the subgroup generated by $\{(12m) \mid 3 \leq m \leq n\}$. If $\{b, c\} \cap \{1, 2\} = \emptyset$, then $(1bc) = (12c)^{-1}(12b)(12c)$ and $(2bc) = (12b)(12c)(12b)^{-1}$. If $\{a, b, c\} \cap \{1, 2\} = \emptyset$, then $(abc) = (12a)(2bc)(12a)^{-1}$. The result now follows. \hfill \square

**Remark.** Lemma 6.8 says more than it says. For example, it also says that $A_n$ is generated by $(521), (523), (524), \ldots, (52n)$; and $A_n$ is generated by $(n31), (n32), (n34) \ldots, (n3n − 1)$. And so on. In the next proof we make use of these variations of Lemma 6.8.
Proposition 6.9. If \( n \geq 5 \), then \( A_n \) is a simple group.

Proof. Let \( H \) be a non-trivial normal subgroup of \( A_5 \). Suppose that \( H \) contains a 3-cycle, say \((123)\). If \( i \not\in \{1,2,3\} \), then \( H \) contains \((12)(3i)(123)(3i)(12) = (1i2)\) and its square \((12i)\). It now follows from Lemma 6.8 that \( H = A_n \).

Thus, combining the previous paragraph and the remark prior to this proposition, we see that if \( H \) contains a 3-cycle, then \( H = A_n \).

Choose \( 1 \neq \alpha \in H \) fixing as many elements of \( \{1,\ldots,n\} \) as possible. Since \( \alpha \) is even it is not a transposition. Write \( \alpha \) as a product of disjoint cycles.

Suppose that only 2-cycles occur in the cycle decomposition of \( \alpha \). Suppose that \( \alpha \) is a product of two disjoint 2-cycles, say \( \alpha = (12)(34) \). Then \( H \) contains \( (543)\alpha(543)^{-1}\alpha^{-1} = (345) \); since \( H \) contains a 3-cycle \( H = A_n \).

Now suppose that \( \alpha \) is a product of more than two disjoint 2-cycles, say \( \alpha = (12)(34)(56)(78) \cdots \). Then \( H \) contains \( (543)\alpha(543)^{-1}\alpha^{-1} = (36)(45) \) and applying the previous argument to \((36)(45)\) in place of \((12)(34)\) we see that \( H \) contains a 3-cycle, and is therefore equal to \( A_n \).

We may now assume that the cycle decomposition for \( \alpha \) contains a \( d \)-cycle with \( d \geq 3 \), say \( \alpha = (123\cdots) \cdots \).

If \( \alpha \) moves just 3 elements then \( \alpha = (123) \), and again we deduce that \( H = A_n \). We also see that \( \alpha \) cannot move exactly four elements, because then it must be \((123i)\) and this is not even.

So \( \alpha \) moves at least 5 elements, say \( 1,2,3,4,5 \). Since \( H \) is normal it contains \( (543)\alpha(543)^{-1} \), and hence \( \beta := (543)\alpha(543)^{-1}\alpha^{-1} \). If \( \alpha \) fixes \( i \), then \( i \) does not appear in the cycle decomposition for \( \alpha \), so \( i \geq 6 \) and \( \beta \) also fixes \( i \). But \( \beta \) also fixes 3, so \( \beta \) fixes more elements than \( \alpha \); this contradicts our choice of \( \alpha \), so we conclude that \( H \) must equal \( A_n \).

Example 6.10. Notice that \( S_n \) is a subgroup of \( S_{n+1} \); it is the subgroup consisting of those permutations \( \sigma \) such that \( \sigma(n+1) = n+1 \). Hence we have inclusions

\[
S_1 \subset S_2 \subset S_3 \subset \cdots.
\]

We can form the “union” of these groups. Let \( G \) be the group consisting of those permutations of \( \{1,2,3,\ldots\} \) that move only a finite number of elements. If \( \sigma \in G \), then there is some \( n \) such that \( \sigma(i) = i \) for all \( i > n \). Hence \( \sigma \) belongs to \( S_n \), and it follows that \( G \) is the union of all \( S_n \), \( n \geq 0 \). Notice that \( G \) is infinite but every element of \( G \) has finite order because it belongs to some \( S_n \).

2. Generating \( S_n \)

Lemma 6.11. Every permutation can be written as a product of transpositions.

Proof. Every cycle is a product of transpositions because, for example,

\[
(1 2 \ldots m - 1 m) = (1 m)(1 m - 1) \cdots (1 3)(1 2).
\]
But every permutation is a product of cycles, so the result follows.

**Proposition 6.12.** $S_n$ is generated by the transpositions $(12)$, $(23)$, \ldots, $(n-1\,\,n)$.

**Proof.** After Lemma 6.11 it suffices to prove that every transposition is a product of transpositions of the form $(i\,i+1)$. If $i < j$, then

$$(ij) = (j-1\,j)(j-2\,j-1) \cdots (i\,i+1) \cdots (j-2\,j-1)(j-1\,j).$$

Check!

Lemma 6.11 can be read as saying that $S_n$ is generated by transpositions. However, one can be efficient and generate it with just $n-1$ transpositions. Show that $S_n = \langle (1\,2), (2\,3), \ldots, (n-1\,n) \rangle$; here and elsewhere the notation $\langle x, y, \ldots, z \rangle$ is used to denote the smallest subgroup containing the elements $x, y, \ldots, z$; we call it the subgroup generated by $x, y, \ldots, z$.

**Some Exercises.**

1. Let $X$ be an infinite set and $X!$ the group consisting of all permutations of $X$. Let $H \subseteq X!$ consist of the permutations that move only a finite number of elements of $X$. Show that $H$ is a subgroup of $X!$.

2. Let $G$ act on a set $X$. Let $Y$ be a subset of $X$. Show that

$$H := \{ g \in G \mid gy = y \text{ for all } y \in Y \}$$

is a subgroup of $G$.

3. Show that a cycle is an odd permutation if and only if it has even length.

4. If $\sigma \in S_n$ is an odd cycle, show that $\sigma^2$ is a cycle.

5. Let $G$ be an abelian group and $H$ the set consisting of the elements of order two and the identity. Show that $H$ is a subgroup. Give an example to show this fails if $G$ is not abelian.

6. Prove that $S_5$ does not have a subgroup with 30 elements. (Hint: use the action of $S_5$ on the cosets of such a subgroup together with the fact that $A_5$ is simple.)

7. Prove that the alternating group $A_5$, which has 60 elements, does not have a subgroup having 30 elements.

8. Does $A_5$ have a subgroup with 10 elements?

9. Make a bracelet with 10 beads, each of which can be any of four colors. How many different bracelets can you make?

10. Let $G$ be a group of order $np$ where $p$ is prime and $1 < n < p$. Show that $G$ is not simple. (Hint: show that $G$ has a subgroup $H$ of order $p$, and look at the action of $G$ on the cosets of $H$.)

**Some More Exercises.**
(1) Fix a positive integer \( n \) and let \( \mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z} \) be the cyclic group of order \( n \). For each integer \( a \), write \( \bar{a} \) for \( a + n\mathbb{Z} \in \mathbb{Z}_n \). Write
\[
U_n := \{ \bar{a} \mid (a, n) = 1 \}
\]
(where \((a, n)\) denotes the greatest common divisor of \( a \) and \( n \)) and define a product on \( U_n \) by
\[
\bar{a} \cdot \bar{b} = \bar{ab}.
\]
Show that this multiplication is well-defined and that it makes \( U_n \) a group—you will have to think carefully about why inverses exist in \( U_n \) (Theorem 4.5 might be helpful).

We define Euler’s \( \phi \)-function by
\[
\phi(n) = \text{the number of integers } a \text{ such that } 1 < a < n \text{ and } (a, n) = 1.
\]
Thus \(|U_n| = \phi(n)\).

(2) Suppose that \( p \) is a prime number and \( n \) an integer. Why does \( p \) divide \( n^p - n \)? Hint: how many elements are in \( U_p \)? do a calculation in \( U_p \). This is called Fermat’s Little Theorem and was proved some 350 years ago.

(3) An \textit{automorphism} of a group \( G \) is an isomorphism \( \psi : G \to G \). We write \( \text{Aut} \; G \) for the set of all automorphisms of \( G \). Show that \( \text{Aut} \; G \) is itself a group. What is the multiplication in \( \text{Aut} \; G \)?

(4) Write \( \mathbb{Z}_p = \{0, 1, 2, \ldots, p-1\} \). Let \( \psi \in \text{Aut} \; \mathbb{Z}_p \). Explain why \( \psi \) is completely determined by \( \psi(1) \).

(5) What is \(|\text{Aut} \; \mathbb{Z}_p|\)? Explain.

(6) (a) If \( p = 17 \) and \( \psi(1) = 3 \), what is \( \psi^4(6) \)? (b) If \( p = 19 \) and \( \psi(2) = 3 \), what is \( \psi(1) \)? (c) If \( p = 133 \) and \( \psi(26) = 26 \), is \( \psi \) the identity map? Why or why not?

(7) Show there are only two groups of order 4 up to isomorphism.

(8) Let \( Q = \{\pm 1, \pm i, \pm j, \pm k\} \) with multiplication table
\[
i^2 = j^2 = k^2 = -1, \quad ij = k, \quad jk = i, \quad ki = j, \quad -1\]multiplying in the obvious way, and other products are consequences of these rules. Compute \( ji, kj, ik \). Determine the center \( Z \) of \( Q \). What is \( Q/Z \) isomorphic to—use the previous exercise.

(9) Show every subgroup of \( Q \) is normal.

(10) Why is \( Q \) not isomorphic to \( D_4 \)?