# Contents

1. **Some facts about integers**  
   1.1 Mathematical Induction  
   1.2 Least Integer Axiom or the Well-Ordering Principle  
   1.3 Division Algorithm  
   1.4 Greatest Common Divisors  
   1.5 Congruences  
   1.6 Solving Linear congruence equation  
   1.7 Appendix  

2. **Permutations**  
   2.1 Mappings  
   2.2 Permutations (Lots of Bijections)  
   2.3 Inverse of a permutation  
   2.4 Computation of $\alpha\beta\alpha^{-1}$  
   2.5 Disjoint cycles  

3. **Definition of a Group and Examples**  
   3.1 Binary operation  
   3.2 Definition of a Group  
   3.3 Laws of Exponents  
   3.4 Group of Isometries and the orthogonal group  
   3.5 Symmetry group of a figure and Dihedral groups  

4. **Subgroups, Cosets and Lagrange Theorem**  
   4.1 Subgroups  
   4.2 Cyclic groups  
   4.3 Cosets and Lagrange’s Theorem  
   4.4 Lagrange’s Theorem and the cyclic groups  
   4.5 Lagrange’s Theorem and Euler’s Theorem
## Contents

5 Homomorphisms 46
5.1 Definitions and Examples 46
5.2 Basic facts about homomorphisms 48
5.3 Kernel and Image of a homomorphism 49

6 Quotient Groups and the Isomorphism Theorems 53
6.1 Cosets of Normal Subgroup and Quotient group 53
6.2 The First Isomorphism Theorem 54
6.3 The Euler’s \( \varphi \) function revisited 56
6.4 Semi direct product 58
6.5 The Second Isomorphism Theorem 60
6.6 The Third Isomorphism Theorem and the Correspondence Theorem 63

7 Group Action 66
7.1 A Theorem of Cayley 66
7.2 Group Action 68
7.3 Orbits and Stabilizers 68
7.4 A relation between orbit and stabilizer 70
7.5 The class equation of a group and Cauchy’s Theorem 70
7.6 Counting with groups 73

8 Sylow’s Theorems 78
8.1 Introduction and Sylow’s Theorems Part 1 78
8.2 Sylow’s Theorems Part 2 and 3 79
8.3 Solvable groups 81
1
Some facts about integers

1.1 Mathematical Induction

The First Principle of Finite Induction states that if $S$ is a set of positive integers with the following properties:

(a) The integer 1 belongs to $S$.
(b) Whenever the integer $k$ is in $S$, the next integer $k + 1$ must also be in $S$.

Then $S$ is the set of all positive integers.

This is a very familiar “principle” we used to prove many mathematical results. It usually appears in the following form:

**Theorem 1.1.1** Given statements $S(n)$, one for each $n \geq 1$, suppose that

1. $S(1)$ is true and
2. if $S(n)$ is true then $S(n + 1)$ is true.

Then $S(n)$ is true for all $n \geq 1$.

The mathematical induction is an axiom which cannot be proved but we know that it is a “reasonable” assumption.

We now look at some problems which can be established using mathematical induction.

**Example 1.1.1** Find a formula for

$$1 + 3 + 5 + 7 + \cdots + (2n - 1).$$

Use induction to prove that your formula is correct.
Solutions. By working out the first few examples, we see that

\[ 1 + 3 + 5 + 7 + \cdots + (2n - 1) = n^2. \]

We prove this by induction. Clearly 1 = 1^2. Suppose the result is true for \( n - 1 \). Then

\[ 1 + 3 + \cdots + (2n - 3) + (2n - 1) = (n - 1)^2 + 2n = n^2. \]

1.2 Least Integer Axiom or the Well-Ordering Principle

Let \( \mathbb{N} \) denote the set of non-negative integers. We now state the Least Integer Axiom:

**Least Integer Axiom.** There is a smallest integer in every nonempty collection \( C \) of \( \mathbb{N} \).

Although this axiom cannot be proved, it is certainly plausible. To see this, consider the following procedure: First check if 1 belongs to \( C \). If not, check if 2 is in \( C \). Continue with this process and at some point, there must be an \( n \) such that \( n \) is in \( C \) since \( C \) is nonempty. This integer \( n \) is the least integer that is contained in \( C \).

It can be shown that the Least Integer Axiom implies the principle of Mathematical Induction. For more details, see page 3 of Rotman’s book.

We now show that the Least Integer Axiom follows from the Principle of Mathematical Induction.

Let \( T \) be a non-empty subset of \( \mathbb{N} \). We first show that if \( T \) is finite then it contains a smallest integer. For integer \( k \geq 1 \), let \( S(k) \) denote the statement “If \( T \) has \( k \) elements in \( \mathbb{N} \), then \( T \) has a smallest integer.” Note that \( S(1) \) is true. Suppose \( S(n) \) is true. Let \( T \) be a set with \( n + 1 \) elements. Let \( x \in T \). Then there exists a smallest integer \( a \in T - \{x\} \) by induction since \( |T - \{x\}| = n \). If \( x > a \) then \( a \) is the smallest integer in \( T \). If \( a > x \) then \( x \) is the smallest integer in \( T \). Hence \( S(n + 1) \) is true. So the Least Integer Axiom holds for all non-empty finite subset of \( \mathbb{N} \).

Suppose now that \( T \) is an infinite subset of \( \mathbb{N} \). Let \( x \in T \). Then \( [0, x] \cap T \) is finite and there exists a smallest integer \( a \in [0, x] \cap T \). Note that if \( y \notin [0, x] \cap T \), then \( y > x \geq a \). Hence, \( a \) is the smallest integer in \( T \). This completes the proof of the Least Integer Axiom.
The Least Integer Axiom is usually the starting point of the study of elementary number theory. In subsequent sections, we will use it extensively to derive some well-known results involving integers.

### 1.3 Division Algorithm

**Theorem 1.3.1 (Division Algorithm)** Given integers $a$ and $b$ with $b > 0$, there exist unique integers $q$ and $r$ with

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

**Proof.** Let

$$S = \{y : y = a - bx, x \text{ is an integer}, \ y \geq 0\}.$$  

This is a non-empty set of nonnegative integers (take $x = -|a|$) and so, by the least integer axiom, it has a smallest member, say $r = a - bq$. Then $a = bq + r$ and $r \geq 0$. Now we show that $r < b$. Suppose $r \geq b$. Then $r - b = a - b(q + 1)$ and hence $r - b \in S$. Since $r - b < r$, this contradicts the minimality of $r$. Hence, $r < b$. The pair $q, r$ is unique, for if there is another pair $q', r'$ such that $a = bq' + r'$, then $b(q' - q) = r - r'$. So, $|r - r'| = 0, b, 2b, 3b, \ldots$. If $|r - r'| \neq 0$ then $b \leq |r - r'|$, a contradiction. Therefore, $r' = r$ and $q = q'$.

□

**Example 1.3.1** Show that an integer of the form $4k + 3$ cannot be a sum of two squares.

**Solutions.** If $x$ and $y$ are integers then $x^2 + y^2$ has remainders 0, 1 or 2 when divided by 4. This implies that an integer with remainder 3 when divided by 4 cannot be a sum of two squares.

When $r = 0$, we have $a = bq$ and we say that $b$ divides $a$. We write

$$b|a.$$  

The integer $b$ is called a divisor of $a$.

**Example 1.3.2** Show that if $n|a$ and $n|b$ then $n|(ax + by)$ for any integers $x$ and $y$.  

Solutions. \( n \mid a \) implies that \( a = ns \) and \( n \mid b \) implies that \( b = nt \). Therefore \( ax + by = n(sx + ty) \). So \( n \mid (ax + by) \).

**Example 1.3.3** Let \( d \neq 0 \) and \( n \) be integers and \( d \mid n \). Then \( |d| \leq |n| \).

Solutions. \( n \mid a \) implies that \( a = ns \) and \( n \mid b \) implies that \( b = nt \). Therefore \( ax + by = n(sx + ty) \). So \( n \mid (ax + by) \).

**Example 1.3.4** Let \( d \) and \( d' \) be non-zero integers. If \( d \mid d' \) and \( d' \mid d \) then \( d = \pm d \).

Solutions. \( n \mid a \) implies that \( a = ns \) and \( n \mid b \) implies that \( b = nt \). Therefore \( ax + by = n(sx + ty) \). So \( n \mid (ax + by) \).

**Definition 1.3.1** We say that a positive integer is a prime if it has only two divisors. An integer that is not a prime is said to be composite.

**Example 1.3.5** Show that if \( p \) is a prime > 3 then \( p^2 + 2 \) is composite.

Solutions. A prime \( > 3 \) must be of the form \( 6k + 1 \) or \( 6k + 5 \). This implies that \( p^2 + 2 \) has remainder 3 when divided by 6 and this implies that \( p^2 + 2 \) is composite.

Given a positive integer \( n \), we can “partition” the set of integers \( \mathbb{Z} \) into disjoint subsets in the following way: Let

\[
[r]_n := \{ k \in \mathbb{Z} | n \mid (k - r) \}. \tag{1.3.1}
\]

Then

\[
\mathbb{Z} = \bigcup_{r=0}^{n-1} [r]_n.
\]

The set \([r]_n\) contains integers with the same remainder \( r \) when divided by \( n \). By Theorem 1.3.1, every integer has a unique remainder when divided by \( n \) and hence, every integer belongs to exactly one \([r]_n\). Hence the union is disjoint.

**Remarks.** Note that \([r]_n\) is a set and \( r \) is an element in the set. We could represent the same using any element in the set \([r]_n\). In other words,

\[
[r]_n = [nt + r]_n
\]
Some facts about integers

for any integer $t$.

Let

$$\mathbb{Z}/n\mathbb{Z} := \{[0]_n, [1]_n, \ldots, [n-1]_n\}.$$ 

On this set, we define two operations on its elements. The “addition” of two elements in $\mathbb{Z}/n\mathbb{Z}$ is defined by

$$[a]_n + [b]_n = [a + b]_n.$$ 

Here, the addition on the right hand side is the usual addition of integers. In a similar way, we define “multiplication” of two elements in $\mathbb{Z}/n\mathbb{Z}$ by

$$[a]_n \cdot [b]_n = [a \cdot b]_n.$$ 

The multiplication on the right hand side is again the product of two integers.

1.4 Greatest Common Divisors

A **common divisor** of integers $a$ and $b$ is an integer $c$ with $c|a$ and $c|b$. The **greatest common divisor** is a number $d$ with the following properties:

(a) $d \geq 0$
(b) $d|a$ and $d|b$
(c) $e|a$ and $e|b$ implies $e|d$

The greatest common divisor of $a$ and $b$ is unique (why?) and is denoted by

$$(a, b).$$

Our next result shows that the greatest common divisor of two integers exists.

**Theorem 1.4.1** Let $a$ and $b$ be integers. Then the smallest non-negative integer in the set

$$I := \{sa + tb | s \in \mathbb{Z}\}$$

is $(a, b)$.

**Proof.** If $a = b = 0$, then $I = \{0\}$. We may assume that at least one of $a$ and $b$ is not zero. Note that $I$ contains positive integers. Let $P$ be subset of $I$ containing positive integers. Then, by the Least Integer
Axiom, there is a smallest positive integer, say \(d\), in \(P\). We must show that 
\[d = (a, b)\]
for some integers \(x, y \in \mathbb{Z}\). We first show that \(d\) is a common divisor of \(a\) and \(b\). Suppose
\[a = dq + r, \quad 0 \leq r < d.\]
(This assumption is valid by Theorem 1.3.1.) Then
\[r = a - dq = a - (xaq + ybq) = a(1 - xq) + byq.\]
Therefore \(r \in P\) and it is smaller than \(d\). But \(d\) is the smallest element in \(P\). Hence \(r = 0\). In other words, \(d|a\). Similarly \(d|b\).

Finally if \(c|a\) and \(c|b\) then \(a = cu\) and \(b = cv\). Now,
\[xa + yb = c(ux + vy) = d,\]
and hence \(c|d\). Hence \(d\) is \((a, b)\).

\[\Box\]

**Definition 1.4.1** We say that two integers \(a\) and \(b\) are relatively prime if
\[(a, b) = 1.\]

Note that if \((a, b) = 1\) then \(1 = ax + by\) for some integers \(x\) and \(y\). Conversely, if \(1 = ax + by\), then \((a, b)|a\) and \((a, b)|b\), and so \((a, b)|1\) and so \((a, b) = 1\).

We thus have

**Theorem 1.4.2** Let \(a\) and \(b\) be integers. Then \((a, b) = 1\) if and only if 
\(1 = ax + by\) for some integers \(x\) and \(y\).

For any positive integer \(n\), the set
\[\left(\mathbb{Z}/n\mathbb{Z}\right)^* := \{[r]_n | 0 \leq r \leq n - 1 \text{ and } (r, n) = 1\},\]
where \([r]_n\) is given by (1.3.1), will also be an important example when we study groups.

**Definition 1.4.2** The number of elements in \((\mathbb{Z}/n\mathbb{Z})^*\), denoted by \(\varphi(n)\), is known as the Euler phi function.
Some facts about integers

In the next Theorem, we prove The Euclidean algorithm. This allows us to compute the greatest common divisor of two numbers.

**Theorem 1.4.3 (The Euclidean Algorithm)** Given positive integers \( a \) and \( b \), where \( b \nmid a \). Let \( r_0 = a, r_1 = b \), and apply the division algorithm repeatedly to obtain a set of remainders \( r_2, r_3, \ldots, r_n, r_{n+1} \) defined successively by the relations

\[
\begin{align*}
  r_0 &= r_1 q_1 + r_2 \quad 0 < r_2 < r_1 \\
  r_1 &= r_2 q_2 + r_3 \quad 0 < r_3 < r_2 \\
    &\vdots \\
  r_{n-2} &= r_{n-1} q_{n-1} + r_n \quad 0 < r_n < r_{n-1} \\
  r_{n-1} &= r_n q_n + r_{n+1} \quad r_{n+1} = 0.
\end{align*}
\]

Then \( r_n \), the last nonzero remainder in this process is \((a, b)\).

We need a simple lemma.

**Lemma 1.4.4** Let \( a, b, q, \) and \( r \) be integers such that

\[
b = aq + r,
\]

then

\[
(a, b) = (a, r).
\]

**Proof.** Let \( d = (a, b) \) and \( d' = (a, r) \). We want to show that \( d = d' \). We first observe that if \( a \mid b \) and both integers are positive then \( a \leq b \). To show that \( d = d' \), it suffices to show that \( d \mid d' \) and \( d' \mid d \). Note that since \( d \mid a \) and \( d \mid b \), we find that \( d \mid (b - aq) \) by Example 1.3.2. Hence, \( d \mid r \) and \( d \) is a common divisor of \( b \) and \( r \). Hence, \( d \) must divide the greatest common divisor of \( b \) and \( r \), which is \( d' \). Similarly, \( d' \mid b \) and \( d' \mid r \) implies that \( d' \mid (aq + r) \) by Example 1.3.2 and consequently, \( d' \mid b \). This implies that \( d' \) is a common divisor of \( a \) and \( b \) and hence it must divide \( d \), which is the greatest common divisor of \( a \) and \( b \). By the remark in the beginning of the proof, we conclude that \( d = d' \).

We now complete the proof of Theorem 1.4.3.

**Proof.**

There is a stage at which \( r_{n+1} = 0 \) because the \( r_i \) are decreasing and nonnegative. Next, applying Lemma 1.4.4, we find that

\[
(a, b) = (r_0, r_1) = (r_1, r_2) = \cdots = (r_n, r_{n+1}) = (r_n, 0) = r_n.
\]
This completes the proof of Theorem 1.4.3.

**Example 1.4.1** Find \((1011, 576)\) and find \(s\) and \(t\) such that
\[
1011s + 576t = (1011, 576).
\]

### 1.5 Congruences

We say that \(a\) is congruent to \(b\) modulo \(n\) when \(n \mid (a - b)\). The notation is
\[
a \equiv b \pmod{n}.
\]

Using this notation, we may rewrite \([r]_n\) as
\[
[r]_n := \{k \in \mathbb{Z} | k \equiv r \pmod{n}\}.
\]

**Theorem 1.5.1 (Basic Properties of Congruences)** Let \(n > 0\) be fixed and \(a, b, c, d \in \mathbb{Z}\). Then

(i) \(a \equiv a \pmod{n}\). (Reflexive)

(ii) If \(a \equiv b \pmod{n}\) then \(b \equiv a \pmod{n}\). (Symmetric)

(iii) If \(a \equiv b \pmod{n}\) and \(b \equiv c \pmod{n}\) then \(a \equiv c \pmod{n}\). (Transitive)

(iv) If \(a \equiv b \pmod{n}\) and \(c \equiv d \pmod{n}\) then \(a + c \equiv b + d \pmod{n}\)
and \(ac \equiv bd \pmod{n}\).

Note that if we let \(a = c, b = d\) in (4) we have a simple consequence that if \(a \equiv b \pmod{n}\) then \(a^2 \equiv b^2 \pmod{n}\). By iterating the process, we conclude that \(a^k \equiv b^k \pmod{n}\) if \(a \equiv b \pmod{n}\). Next, if we set \(c = d\) in (4), then we have \(a \equiv b \pmod{n}\) implies that \(ac \equiv bc \pmod{n}\).

Congruence behaves very much like equality “=”.

**Theorem 1.5.2** If \(ca \equiv cb \pmod{n}\) and \((c, n) = 1\), then \(a \equiv b \pmod{n}\).
Proof. Now, recall that if \((c,n) = 1\) then there exist \(x\) and \(y\) such that \(cx + ny = 1\). Multiplying \(a\) and \(b\) yields

\[ acx + any = a \]

and

\[ bcx + bny = b, \]

respectively. Since \(ac \equiv bc \pmod{n}\), we conclude that \(a - b \equiv (ac - bc)x \equiv 0 \pmod{n}\).

Theorem 1.5.2 can be used to prove the following result of Euclid.

**Corollary 1.5.3 (Euclid’s Lemma)** Let \(p\) be a prime. If \(p|ab\) then \(p|a\) or \(p|b\).

**Proof.** For any integer \(n\), \((n,p) = 1\) or \(p\) (\(p\) has only two divisors).

Suppose \(p \nmid a\). Then \((p,a)=1\). By Theorem 1.5.2, the relation

\[ ab \equiv 0 \pmod{p} \]

then implies that

\[ b \equiv 0 \pmod{p}. \]

\[ \square \]

One of the applications of Euclid’s Lemma is to prove the Fundamental Theorem of Arithmetic. For more details, see the Appendix.

The use of \(“a \equiv b \pmod{n}”\) to represent the statement “\(a\) and \(b\) has the same remainder when divided by \(n\)” proves to be very useful in solving problems. We illustrate with an example.

**Example 1.5.1** Show that a positive integer is divisible by 9 if and only if the sum of its digits is divisible by 9.

### 1.6 Solving Linear congruence equation

A linear congruence equation is of the form

\[ ax \equiv b \pmod{n}, \]

where \(a\), \(b\) and \(n\) are given integers and \(x\) is a variable. Our aim is to find an integer \(x\) (usually we want \(0 \leq x < n\)) that satisfies the equation. A linear congruence equation may not be solvable. For example if \((a,n)\) does not divide \(b\) then the equation is not solvable.
Example 1.6.1 (Solving linear congruence equations) Find an integer \( x \) so that

\[
8x \equiv 1 \pmod{13}.
\]

Solutions. By the Euclidean Algorithm, we find that

\[
13 = 8 + 5 \\
8 = 5 + 3 \\
5 = 3 + 2 \\
3 = 2 + 1.
\]

Hence

\[
1 = 3 - 2 = 3 - (5 - 3) = 2 \cdot (8 - 5) - 5 = 2 \cdot 8 - 3 \cdot (13 - 8) = 5 \cdot 8 - 3 \cdot 13.
\]

This implies that \( x \) may be taken to be 5. The above computation is important later for finding “inverses” of an element in the finite group \((\mathbb{Z}/n\mathbb{Z})^*\) for some positive integer \( n \).

1.7 Appendix

Theorem 1.7.1 (Fundamental Theorem of Arithmetic) Every positive integer \( n > 1 \) can be expressed as a product of primes; this representation is unique up apart from the order in which the factors occur.

We first need to show the following:

Theorem 1.7.2 Every integer \( n > 1 \) is either a prime or a product of prime numbers.

To prove the above result, we need the following version of mathematical induction. The proof of this result is the same as that of Theorem 1.1.1.

Theorem 1.7.3 Given statements \( S(n) \), one for each \( n \geq m \), suppose that

1. \( S(m) \) is true and
2. if \( S(k) \) is true for all \( m \leq k \leq n \) then \( S(n+1) \) is true.

Then \( S(n) \) is true for all \( n \geq m \).
Some facts about integers

Proof of Theorem 1.7.2. We use induction on \( n \). The Theorem is true for \( n = 2 \). Suppose it is true for all integers \(< n\). Then if \( n \) is not a prime it has a positive divisor \( d \neq 1 \) and \( d \neq n \). Hence \( n = cd \), where \( c \neq n \). But both \( c < n \) and \( d < n \) and therefore, by induction hypothesis, they are products of primes. This implies that \( n \) is also a product of primes.

Proof of the Fundamental Theorem of Arithmetic. We use induction on \( n \). The theorem is true for \( n = 2 \). Assume, then that it is true for all integers greater than 1 and less than \( n \). We shall prove that it is also true for \( n \). If \( n \) is prime there is nothing to prove. Assume, then, that \( n \) is composite and that \( n \) has two factorizations, say

\[ n = p_1 p_2 \cdots p_s = q_1 q_2 \cdots q_t. \]  \hspace{1cm} (1.7.1)

Since \( p_1 \) divides the product \( q_1 q_2 \cdots q_t \), it must divide at least one factor. Relabel \( q_1, q_2, \ldots, q_t \) so that \( p_1 | q_1 \). Then \( p_1 = q_1 \) since both \( p_1 \) and \( q_1 \) are primes. In (1.7.1), we may cancel \( p_1 \) on both sides to obtain

\[ n/p_1 = p_2 \cdots p_s = q_2 \cdots q_t. \]

Now the induction hypothesis tells us that the two factorizations of \( n/p_1 \) must be identical, apart from the order of the factors. Therefore, \( s = t \) and the factorizations in (1.7.1) are also identical, apart from order. This completes the proof.
2
Permutations

2.1 Mappings

A set $X$ is a collection of elements (numbers, points, elephants, etc.); one writes

$$x \in X$$

to denote “$x$ belongs to $X$”. Two sets are equal whenever the statement “$x \in X$ if and only if $x \in Y$” holds.

A subset of a set $X$ is a set $S$ each of whose elements also belongs to $X$: if $s \in S$ then $s \in X$. If $S$ is a subset of $X$, we also say that $S$ is contained in $X$ and this is denoted by

$$S \subset X.$$  

Two sets $X$ and $Y$ are equal if one can show that

$$X \subset Y \text{ and } Y \subset X.$$  

Definition 2.1.1 Let $S$ and $T$ be two non-empty sets. A mapping $\sigma$ from $S$ to $T$ is a rule that assign to each $s \in S$, a unique $t \in T$. We write $t = \sigma(s)$.

Example 2.1.1 The rule $\sigma(x) = x^2$, where $x \in \mathbb{R}$ is a mapping. But the rule $\tau(x) = y$ with $y^2 = x$ is not a mapping since $\tau(1) = \pm 1$.

A mapping $\sigma$ from $S$ to $T$ is called an injection if the condition $\sigma(a) = \sigma(b)$ implies that $a = b$. A mapping $\sigma$ is called a surjection if for every $b \in T$, there exist an element $a \in S$ such that $\sigma(a) = b$. In other words, all the elements in $T$ are “hit” by $\sigma$. If $\sigma$ is both injective and surjective then $\sigma$ is called a bijection and there is a one to one correspondence between $S$ and $T$.  

15
Example 2.1.2 The map $\sigma_1 : \mathbb{Z} \to 2\mathbb{Z}$ defined by $\sigma(a) = 2a$ is a bijection. However the map $\sigma_2 : 2\mathbb{Z} \to \mathbb{Z}$ with $\sigma_2(2t) = 2t$ is injective but not surjective. The map $\sigma_3 : \mathbb{R} \to \mathbb{R}^+$ defined by $\sigma_3(x) = x^2$ is surjective but not injective.

2.2 Permutations (Lots of Bijections)

Definition 2.2.1 A permutation of a set $X$ is a bijection from $X$ to itself.

In high school, we learn about permutation of a set of say three letters or three numbers. How many ways can we arrange 1,2,3? The number of ways is 6 and they are:

\[123\quad 132\quad 213\quad 231\quad 312\quad 321.\]

We may now assign a bijection to each permutation of 1, 2, 3. For example, corresponding to the first arrangement, we define the function mapping 1 to 1, 2 to 2 and 3 to 3. The second arrangement corresponds to a function sending 1 to 1, 2 to 3 and 3 to 2. In general to obtain from each permutation a corresponding function, we write the number 1,2,3 above the permutation, for example, for last arrangement, we have

\[
\begin{pmatrix}
1 & 2 & 3 \\
3 & 2 & 1
\end{pmatrix}.
\]

Hence, the function corresponding to 321 sends 1 to 3, 2 to 2 and 3 to 1. It is clear that given a bijection from $X := \{1,2,3\}$ to itself, we obtain a permutation of 1,2 and 3. We denote a bijection obtained from a permutation of $\{1,2,\cdots,n\}$ by

\[
\alpha = \begin{pmatrix}
1 & 2 & 3 & \cdots & n \\
\alpha(1) & \alpha(2) & \alpha(3) & \cdots & \alpha(n)
\end{pmatrix}.
\]

Here it is understood that 1 is sent to $\alpha(1)$, 2 is sent to $\alpha(2)$ etc.

Definition 2.2.2 The set of all bijections on the set $X = \{1,2,\cdots,n\}$ is called the symmetric group on $X$. We denote this set as $S_n$. (The word “group” will be explained soon).
Example 2.2.1 The elements in $S_3$ are

$$\left\{ \begin{array}{c}
(1, 2, 3), (1, 2, 3), (1, 2, 3), (1, 2, 3), (1, 2), (1, 3, 2), (2, 1, 3), (2, 3, 1)
\end{array} \right\}$$

Given two functions $f$ and $g$ in $S_n$, we defined $f \circ g$ to be the function sending $k$ to $f(g(k))$. So for example, if

$$\alpha = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$

then $\alpha \circ \beta$ sends 1 to 3, 2 to 2 and 3 to 1. Similarly $\beta \circ \alpha$ sends 1 to 1, 2 to 3, 3 to 2. Note that $\beta \circ \alpha \neq \alpha \circ \beta$. Because of this, elements in $S_n$ do not satisfy the commutative law, i.e., $ab = ba$.

Note that from now on, we will suppress the $\circ$ and replace $\alpha \circ \beta$ by $\alpha \beta$. The representation of $\alpha \in S_n$ using two rows is very cumbersome. We now introduce a better way to represent $\alpha$.

Definition 2.2.3 Let $i_1, i_2, \cdots, i_r$ be distinct integers in $\{1, 2, \cdots, n\}$. If $\alpha \in S_n$ fixes the other integers and if

$$\alpha(i_1) = i_2, \alpha(i_2) = i_3, \cdots, \alpha(i_{r-1}) = i_r, \alpha(i_r) = i_1,$$

then we call $\alpha$ an r-cycle. We also say that $\alpha$ is a cycle of length $r$, and we denote it by

$$\alpha = (i_1 \ i_2 \cdots i_r).$$

Note that any $r$-cycle is an element of $S_n$ if $i_j \leq n$. A 2-cycle interchanges $i_1$ and $i_2$ and fixes everything else and it is called a transposition. A 1-cycle is in $S_n$ and it fixes every integer from $\{1, 2, \cdots, n\}$. We usually write a 1-cycle as $e$ or id.

We now describe a way to transform the two-rows representation of
$\alpha \in S_n$ into cycles. † This is best described using examples. Suppose

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 6 & 4 & 7 & 2 & 5 & 1 & 8 & 9 & 3 \end{pmatrix}. $$

We begin with 1. Note that $1 \mapsto 6 \mapsto 1$. This corresponds to a 2-cycle $(1 \ 6)$. Next we start with a number not in the first cycle, say 2. Now, $2 \mapsto 4 \mapsto 2$, corresponding to $(2 \ 4)$. Next, we look at the images of 3 under successive applications of $\alpha$, namely, $3 \mapsto 7 \mapsto 8 \mapsto 9 \mapsto 3$ and this is represented by $(3 \ 7 \ 8 \ 9)$. Finally 5 is fixed by $\alpha$ and corresponds to the one cycle $(5)$. Hence we find that

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

We do not usually write down 1-cycle and hence we may rewrite the above as

$$\alpha = (1 \ 6)(2 \ 4)(3 \ 7 \ 8 \ 9).$$

It is easy to retrieve the two-rows representation from the cycle representation of $\alpha$.

**Example 2.2.2** With the cycle notation, we can list all the elements of $S_3$ in a compact form, namely,

$$S_3 = \{(1), (1 \ 2), (2 \ 3), (1 \ 3), (1 \ 2 \ 3), (1 \ 3 \ 2)\}.$$  

The cycle notation allows us to multiply two elements $\alpha, \beta \in S_n$ but just looking at the cycles of $\alpha$ and $\beta$. For example if $\alpha = (1 \ 2)(3 \ 4)$ and $\beta = (1 \ 3)(2 \ 4)$ then

$$\alpha \beta = (1 \ 2)(3 \ 4)(1 \ 3)(2 \ 4).$$

We determine $\alpha \beta$ by determining the image of 1, 2, 3, and 4. To determine the image of 1, we read from right to left. We see that 1 first appears in the third cycle and is mapped to 3. Continuing to the right, we see that 3 is then mapped to 4. Hence 1 is mapped to 4. Now, we asked for the image of 4 under $\alpha \beta$. Note that 4 appears in the last cycle and mapped to 2, 2 is then mapped to 1. Hence 4 is mapped to 1. This implies that the cycle $(1 \ 4)$ is part of the cycle representation of $\alpha \beta$. Similarly we see that $(2 \ 3)$ is part of the cycle representation of $\alpha \beta$. Since we have exhausted the numbers 1,2,3,4 we conclude that

$$\alpha \beta = (1 \ 4)(2 \ 3).$$

† For a proof of the fact that all permutations are products of disjoint cycles, see Theorem 2.5.2.
2.3 Inverse of a permutation

The permutation $(1)(2)\cdots(n)$ that fixes all $i \in X = \{1, 2, \ldots, n\}$ is denoted by $e$.

An inverse of $\alpha \in S_n$, denoted by $\alpha^{-1}$, is the element such that $\alpha\alpha^{-1} = \alpha^{-1}\alpha = e$. If $\alpha$ is an $r$-cycle, it is easy to write down $\alpha^{-1}$. If $\alpha = (i_1 \ i_2 \ i_3 \cdots i_{r-1} \ i_r)$ then the inverse $\alpha^{-1} = (i_r \ i_{r-1} \cdots i_3 \ i_2 \ i_1)$. This is because if $i_j \mapsto i_{j+1}$ by $\alpha$ then $i_{j+1} \mapsto i_j$ by $\alpha^{-1}$ and hence $i_j$ is fixed by $\alpha\alpha^{-1}$.

Example 2.3.1 The inverse of $\alpha = (1 \ 2 \ 3 \ 4)$ is $(4 \ 3 \ 2 \ 1)$.

If $\alpha$ is a product of cycles given by

$$\tau_1\tau_2\cdots\tau_s,$$

then $\alpha^{-1}$ is given by (Check directly)

$$\tau_s^{-1}\tau_{s-1}^{-1}\cdots\tau_1^{-1}.$$  

The inverse of $\tau_i$ is given by the construction above and hence, we are able to determine the inverse of any product of cycles. We will prove in Theorem 2.5.2 that any permutation can be expressed as product of “disjoint cycles” (See Definition 2.5.1) and hence we have a way to determine the inverse of any permutation.

2.4 Computation of $\alpha\beta\alpha^{-1}$

Now that we know the meaning of inverse of a permutation and the computation of the inverse, we proceed to develop some computational tools. In this section, we will learn how to compute $\alpha\beta\alpha^{-1}$ for any two permutations $\alpha$ and $\beta$. The expression $\alpha\beta\alpha^{-1}$ is called the conjugate of $\beta$ by $\alpha$.

We prove the following result:

**Theorem 2.4.1** If $\beta(i) = j$ then $\alpha\beta\alpha^{-1}(\alpha(i)) = \alpha(j)$.

The proof is clear. It shows that in order to compute $\alpha\beta\alpha^{-1}$ it suffices to apply $\alpha$ into the elements in each cycle of $\beta$. For example, if $\beta = (1 \ 2 \ 3)$ then $\alpha\beta\alpha^{-1} = (\alpha(1) \ \alpha(2) \ \alpha(3))$. This also shows that if $\beta$ is an $r$-cycle so is $\alpha\beta\alpha^{-1}$. 
2.5 Disjoint cycles

Definition 2.5.1 Two cycles $\alpha, \beta \in S_n$, where $\alpha \neq e$ and $\beta \neq e$, are disjoint if every $i$ moved by one is fixed by the other. More precisely, the following conditions must hold:

(a) If $\alpha(i) \neq i$ then $\beta(i) = i$.

(b) If $\beta(j) \neq j$, then $\alpha(j) = j$.

Example 2.5.1 The cycles $(1 \ 2)$ and $(3 \ 4)$ are disjoint and the cycles $(1 \ 2)$ and $(1 \ 3)$ are not.

Suppose $\alpha$ and $\beta$ are cycles with length $> 1$. It is clear that $\alpha$ and $\beta$ are disjoint if and only if the elements from $\{1, 2, \cdots, n\}$ appearing in the cycles $\alpha$ and $\beta$ are different. For example, the number 1 appears in the cycles of both $\alpha$ and $\beta$ if and only if 1 is moved by both $\alpha$ and $\beta$.

Lemma 2.5.1 Disjoint cycles $\alpha, \beta \in S_n$ commutes.

Proof. Simply check that

$$\alpha\beta(i) = \beta\alpha(i) \quad (2.5.1)$$

for $1 \leq i \leq n$. It is clear that if $i$ is fixed by both $\alpha$ and $\beta$ then (2.5.1) holds. If $i$ is moved by $\alpha$, $i$ is fixed by $\beta$ and so, $\alpha\beta(i)$ is $\alpha(i)$. To complete our proof, we need to show that $\beta\alpha(i) = \alpha(i)$. We claim that $\alpha$ moves $\alpha(i)$. Suppose this is not the case, then $\alpha$ fixes $\alpha(i)$. This means that

$$\alpha(\alpha(i)) = \alpha(i).$$

Applying $\alpha^{-1}$ on both sides, we conclude that

$$\alpha(i) = i,$$

which is clearly a contradiction. Since $\alpha$ moves $\alpha(i)$, $\beta$ fixes $\alpha(i)$ and therefore,

$$\beta(\alpha(i)) = \alpha(i),$$

which is what we want to establish.

Definition 2.5.2 Let $k \in \mathbb{Z}$. If $\alpha \in S_n$, we define $\alpha^0 = e$, and for $k > 0$,

$$\alpha^{k+1} = \alpha\alpha^k.$$
If $k > 0$, define
\[ \alpha^{-k} = (\alpha^k)^{-1}. \]

Note that by the above definition, we have if \( k \geq 1 \),
\[ \alpha^{k-1} = \alpha^k \alpha^{-1} \]
and hence
\[ \alpha^{k-\ell} = \alpha^k \alpha^{-\ell}. \]

**Theorem 2.5.2** Every permutation $\alpha \in S_n$, where $\alpha \neq e$, is a cycle or product of disjoint cycles.

**Proof.**
Let $k$ be the number of elements in $X = \{1, 2, \cdots, n\}$ moved by $\alpha \in S_n$. Note that $k \geq 2$. We prove the assertion by induction. Suppose the assertion is true for permutations that move less than $k$ elements. Let $\alpha$ be a permutation that moves $k$ elements. Let $i$ be moved by $\alpha$. Then the set $\{i, \alpha(i), \alpha^2(i), \cdots\}$ is finite. Furthermore, there exists an $r$ such that $\alpha^r(i) = i$. This is because there are some integers $j, \ell$ such that $\alpha^\ell(i) = \alpha^j(i)$. This implies that if $j > \ell$, then $\alpha^{j-\ell}(i) = i$. If we set $T = \{s \in \mathbb{Z}| \alpha^s(i) = i\}$, then we see that $T$ is non-empty. By the least integer axiom, there is a smallest non-negative integer $r$ such that $\alpha^r(i) = i$. We claim that $Y = \{i, \alpha(i), \cdots, \alpha^{r-1}(i)\}$ contains distinct elements. For if not, then $\alpha^u(i) = \alpha^v(i)$, with $u, v < r$. This means that $\alpha^{u-v}(i) = i$ (assuming $u > v$) and this contradicts the minimality of $r$.

Now let $\tau = (i \alpha(i) \cdots \alpha^{r-1}(i))$ and consider the permutation $\alpha' = \alpha \tau$. If $\alpha' = e$ then $\alpha$ is a cycle and we are done. Suppose $\alpha' \neq e$. Note that $\alpha \tau^{-1}$ fixes all elements in $Y$ and so $\alpha'$ fixes these elements and thus moves less than $k$ elements in $X$. So $\alpha'$ is a cycle or a product of disjoint cycles. Now, if a cycle in the product of $\alpha'$ moves $j$, then $j \notin Y$ since $\alpha'$ fixes all elements in $Y$. Since $\tau$ moves only elements in $Y$, we see that $\tau$ fixes $j$.

If $\tau$ moves $j'$ then $\alpha'$ fixes $j'$ by our construction. This implies that the cycles in the representation of $\alpha'$ are disjoint from $\tau$. Therefore, $\alpha = \alpha' \tau$ is a product of disjoint cycles.

**Definition 2.5.3** A complete factorization of a permutation $\alpha$ is a factorization of $\alpha$ into disjoint cycles that contains one 1-cycle $(i)$ for every $i$ fixed by $\alpha$. 

\[2.5 \text{ Disjoint cycles}\]
The complete factorizations of elements in $S_3$ are
\[ S_3 = \{(1)(2)(3), (1\ 2)(3), (2\ 3)(1), (1\ 3)(2), (1\ 2\ 3), (1\ 3\ 2)\}. \]

Our next task is to show that the representation of a permutation into disjoint cycles is unique.

**Lemma 2.5.3** Suppose $\alpha \neq e$. Let $\alpha = \beta_1 \cdots \beta_t$ be a factorization into disjoint cycles. If $\beta_1$ moves $i$ then $\alpha^k(i) = \beta_1^k(i)$, $k \geq 1$.

**Proof.** Let $\delta = \beta_2 \cdots \beta_t$. Since $\beta_m$ and $\beta_1$ are all disjoint for $2 \leq m \leq t$, by Lemma 2.5.1, we find that $\beta_1 \beta_m = \beta_m \beta_1$. Therefore, we conclude that
\[ \beta_1 \delta = \beta_2 \beta_1 \beta_3 \cdots \beta_t = \beta_2 \beta_3 \cdots \beta_1 \beta_1 = \delta \beta_1. \] (2.5.2)

We now show a general statement that if $ab = ba$ then for positive integers $k \geq 1$,
\[ (ab)^k = a^k b^k. \]

We first show that $ab^k = b^k a$. Let $\ell \geq 1$ be an integer. Let $S(\ell)$ denote the statement “If $ab = ba$, then $(ab)^\ell = b^\ell a.$” $S(1)$ is true. Suppose $S(k)$ is true. Then
\[ ab^{k+1} = ab^k b = b^k ab = b^k ba = b^{k+1} a. \]

So $S(k+1)$ is true. Hence if $ab = ba$ and $k \in \mathbb{Z}^+$, $ab^k = b^k a$.

Next, let $S_1(\ell)$ denote the statement “If $ab = ba$, then $(ab)^k = a^k b^k.$” $S_1(1)$ is true. Suppose $S_1(k)$ is true. Then
\[ (ab)^{k+1} = (ab)^k ab = a^k b^k ab = a^k ab^k b = a^{k+1} b^{k+1}, \]
and hence the statement $S_1(k)$ is true for all integers $k \in \mathbb{Z}^+$. Using (2.5.2) and the above observation, we find that
\[ \beta_1^k \delta^k = (\beta_1 \delta)^k. \] (2.5.3)

Now,
\[ \beta_1^k \delta^k(i) = \beta_1^k(i) \]
and
\[ (\beta_1 \delta)^k(i) = a^k(i). \]

Using (2.5.3), we complete our proof.

**Lemma 2.5.4** If $\beta$ and $\gamma$ are cycles both of which move some $i$, and if $\beta^k(i) = \gamma^k(i)$ for all $k \geq 1$, then $\beta = \gamma$. 

2.5 Disjoint cycles

Proof. This is clear since
\[ \beta = (i_1 \beta(i_1) \beta^2(i_1) \cdots \beta^{r-1}(i_1)) = (i_1 \gamma(i_1) \gamma^2(i_1) \cdots \gamma^{r-1}(i_1)). \]

Theorem 2.5.5 Let \( \alpha \neq e \in S_n \) and let \( \alpha = \beta_1 \cdots \beta_t \) be a complete factorization into disjoint cycles. This factorization is unique except for the order in which the cycles occur.

Proof. Every \( \alpha \) is a cycle or a product of disjoint cycles by Theorem 2.5.2. Inserting all the missing 1-cycles, we see that every \( \alpha \) has a complete factorization into disjoint cycles. For \( k \geq 2 \), we let \( S(m) \) denote the statement that “a permutation in \( S_n \) that moves \( k \) elements has a unique complete factorization into disjoint cycles.” We observe \( S(2) \) is true since if \( \alpha \) moves two elements, say 1 and 2, then \( \alpha = (1 \ 2)(3)(4) \cdots (n) \) and the representation is unique. Suppose \( S(k) \) is true for all \( 2 \leq k \leq m-1 \).

Let \( \alpha \) moves \( m \) elements. If \( \alpha \) is an \( m \)-cycle, then it must be of the form \( (a_1 \ a_2 \cdots a_m) \). Suppose \( \alpha \) is not an \( m \)-cycle. Let \( \alpha = \beta_1 \cdots \beta_t = \gamma_1 \cdots \gamma_s \) be two complete factorization of \( \alpha \) into disjoint cycles. Note that at least two of the cycles in each representation are not 1-cycle. Suppose \( \beta_t \) is not a one-cycle and that \( \beta_t \) moves \( i \). Then \( \alpha^k(i) = \beta_t^k(i), k \geq 1, \) by Lemma 2.5.3. Now, some \( \gamma_j \) must move \( i \). Since disjoint cycles commute by Lemma 2.5.1, we may assume that \( \gamma_j = \gamma_s \). Now, \( \gamma_s^k(i) = \alpha^k(i) \) by Lemma 2.5.3 and this implies that \( \gamma_s^k(i) = \beta_t^k(i) \). By Lemma 2.5.4, this implies that \( \gamma_s = \beta_t \). Suppose \( \beta_t \) moves \( \ell \) elements, then \( \alpha \beta_t^{-1} \) fixes these \( \ell \) elements and thus, \( \alpha \beta_t^{-1} \) moves \( k - \ell \) elements. Note that \( k - \ell \geq 2 \) since we assume that at least two cycles in each representation of \( \alpha \) are not 1-cycle. Therefore, by induction, the complete factorization of disjoint cycles of \( \alpha \beta_t^{-1} \) is unique and hence \( s = t \) and \( \gamma_j = \beta_j, 1 \leq j \leq t - 1 \). Therefore, the complete factorization of disjoint cycles of \( \alpha \) is unique.

We have just seen that any permutations in \( S_n \) is a product of disjoint cycles. We may also write permutations as product of transpositions, namely, product of two cycles (This shows in particular that writing permutations in terms of products of \( r \)-cycles is not unique).

Theorem 2.5.6 Any cycle \( (a_1 \ a_2 \cdots a_n) \) can be written as a product of transpositions, namely,
\[ (a_1 \ a_2 \cdots a_{n-1} \ a_n) = (a_1 \ a_n)(a_1 \ a_{n-1}) \cdots (a_1 \ a_2). \]
Permutations

The above can be checked directly.

Product of two disjoint cycles commute. The product of two transpositions, however, may not commute with each other. For example

\[(1 \ 2)(1 \ 3) = (1 \ 3 \ 2)\] while \[(1 \ 3)(1 \ 2) = (1 \ 2 \ 3)\]. Therefore the arrangement of the cycles in the representation of a permutation as transpositions is very important.

The representation of a permutation into a product of transpositions is not unique. For example

\[
(1 \ 2 \ 3) = (1 \ 3)(1 \ 2) = (2 \ 3)(1 \ 3) = (1 \ 3)(4 \ 2)(1 \ 2)(1 \ 4).
\]

Although there is no uniqueness in the above representations of the permutation \((1 \ 2 \ 3)\) by transpositions, the number of transpositions is always even. It is a fact that if a permutation is a product of even (odd) number of transpositions then it cannot be a product of an odd (even) number of transpositions.

**Definition 2.5.4** A permutation \(\alpha \in S_n\) is even if it can be factored into a product of an even number of transpositions; otherwise \(\alpha\) is odd.

We have not seen that if a permutation is even, then it cannot be written as a product of an odd number of transpositions. We only say that if a permutation cannot be written as a product of even number of transpositions then it is odd.

**Lemma 2.5.7** If \(k, l \geq 0\) and the letters \(a, b, c_i, d_j\) are all distinct, then

\[
(a\ b)(a\ c_1\cdots c_k\ b\ d_1\cdots d_l) = (a\ c_1\cdots c_k)(b\ d_1\cdots d_l),
\]

and

\[
(a\ b)(a\ c_1\cdots c_k)(b\ d_1\cdots d_l) = (a\ c_1\cdots c_k\ b\ d_1\cdots d_l).
\]

**Proof.** Check directly.

**Definition 2.5.5** If \(\alpha \in S_n\) and \(\alpha = \beta_1 \cdots \beta_t\) is a complete factorization into disjoint cycles, then signum \(\alpha\) is defined by

\[
\text{sgn}(\alpha) = (-1)^{n-t}.
\]

Note that since complete factorization of \(\alpha\) is unique, \(\text{sgn}(\alpha)\) is well-defined. If \(\tau\) is a transposition, \(\text{sgn}(\tau) = (-1)^{n-(n-1)} = (-1)\).
Lemma 2.5.8 If $\alpha, \tau \in S_n$ and $\tau$ is a transposition, then
\[
\text{sgn}(\tau\alpha) = -\text{sgn}(\alpha).
\]

Proof. Let $\alpha = \beta_1 \cdots \beta_t$ be a complete factorization of $\alpha$. If $\tau = (a \ b)$ and $a$ and $b$ occur in the same cycle $\beta_1$, say, then $\tau\alpha$ will have one extra cycle compared to the decomposition of $\alpha$ by Lemma 2.5.7. This shows that $\text{sgn}(\tau\alpha) = -\text{sgn}(\alpha)$. If $a$ and $b$ are contained in different cycles then $\tau\alpha$ will have one cycle less than that of the decomposition of $\alpha$ by Lemma 2.5.7. This again shows that $\text{sgn}(\tau\alpha) = -\text{sgn}(\alpha)$.

Theorem 2.5.9 For all $\alpha, \beta \in S_n$,
\[
\text{sgn}(\alpha\beta) = \text{sgn}(\alpha)\text{sgn}(\beta).
\]

Proof. Assume that $\alpha \in S_n$ is given and that $\alpha = \tau_1 \cdots \tau_m$ is a factorization of $\alpha$ into transpositions with $m$ minimal. We prove, by induction on $m$, that for every $\beta \in S_n$,
\[
\text{sgn}(\alpha\beta) = \text{sgn}(\alpha)\text{sgn}(\beta).
\]
The base step is given by Lemma 2.5.8. If $m > 1$, note that $\tau_2 \cdots \tau_m$ is also minimal. Therefore,
\[
\begin{align*}
\text{sgn}(\alpha\beta) &= \text{sgn}(\tau_1 \cdots \tau_m\beta) = -\text{sgn}(\tau_2 \cdots \tau_m\beta) \quad \text{(by Lemma 2.5.8)} \\
&= -\text{sgn}(\tau_2 \cdots \tau_m)\text{sgn}(\beta) \quad \text{(by Induction)} \\
&= \text{sgn}(\tau_1 \cdots \tau_m)\text{sgn}(\beta) \quad \text{(by Lemma 2.5.8)} \\
&= \text{sgn}(\alpha)\text{sgn}(\beta).
\end{align*}
\]
Now, if $\alpha = \tau_1 \cdots \tau_m = \gamma_1 \cdots \gamma_s$ are two representations as products of transpositions. Then $\text{sgn}(\alpha) = (-1)^m = (-1)^s$. Hence $s$ and $m$ must have the same parity. As a result, an even permutation can only be represented by a product of even number of transpositions and an odd permutation can only be represented by a product of odd number of transpositions.

† You should check that this is also true if $a$ or $b$ or both are contained in a 1-cycle.
3
Definition of a Group and Examples

3.1 Binary operation

We begin this Chapter with the definition of groups. We will then discuss examples of groups in details.

Definition 3.1.1 A binary operation on $A$ is a function $\sigma$ from $A \times A$ to $A$.

In other words, $\sigma((a, b)) = c$, where $a, b, c \in A$. Writing $\sigma((a, b))$ may be too cumbersome. We will replace the notation $\sigma((a, b))$ by $a \cdot b$.

Example 3.1.1 The addition and multiplication are binary operations on $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$.

Example 3.1.2 We have seen that given $N$, the set of integers can be partitioned into $N$ disjoint sets according to the remainders when divided by $N$. If we let $[i]_N$ denote the set of integers that gives a remainder of $i$ when divided by $N$, i.e.,

$$[i]_N := \{ n \in \mathbb{Z} | n \equiv i \pmod{N}\},$$

then we may write

$$\mathbb{Z} = [0]_N \cup [1]_N \cup \cdots \cup [N-1]_N.$$ 

Recall that given $i$, $[i+Nk]_N = [i]_N$. Let $\mathbb{Z}/\mathbb{N} = \{[0]_N, [1]_N, \cdots, [N-1]_N\}$. We may now define a binary operation on $\mathbb{Z}/\mathbb{N}$. We define

$$[a]_N + [b]_N := [a + b]_N,$$
and
\[ [a]_N \cdot [b]_N = [ab]_N. \]

All we need to check is that the above operations are well-defined, i.e., they are functions on \( S \times S \).

**Example 3.1.3** Let \( S_n \) be the set of permutations of \( \{1, 2, \cdots, n\} \). The composition of two permutation \( \alpha, \beta \in S_n \), namely, \( \alpha \beta \) is a binary operation.

**Example 3.1.4** Let \( M \) be the set of 2 by 2 matrices with entries in \( \mathbb{R} \). Matrix multiplication is a binary operation on \( M \times M \).

### 3.2 Definition of a Group

**Definition 3.2.1** A group is a set \( G \) equipped with a binary operation \( \cdot \) and an element \( e \in G \), called the identity, such that

(i) the **associative law** holds: for every \( x, y, z \in G \),
\[ x \cdot (y \cdot z) = (x \cdot y) \cdot z. \]

(ii) \( e \cdot x = x = x \cdot e; \)

(iii) for every \( x \in G \) there is \( x' \in G \) with \( x \cdot x' = e = x' \cdot x \).

**Example 3.2.1** The set of rational numbers \( \mathbb{Q} \) with operation \( + \) is a group. The identity is 0 and the inverse of \( n \) is \( -n \). The operation is clearly associative. The set \( \mathbb{R} = \mathbb{Q} - \{0\} \) with \( \cdot \) (multiplication) is also a group. The identity is 1 and the inverse of \( n \) is \( 1/n \).

**Example 3.2.2** We note that in the previous Chapter, our set \( S_n \) is indeed a group. Note that the binary operation is composition of permutations and the operation is associative. The identity is the 1 cycle. The inverse of the cycle \( c = (i_1 \ i_2 \cdots i_r) \) is \( c^{-1} = (i_r \ i_{r-1} \cdots i_2 \ i_1) \) and in general, the inverse of the product of cycles \( c_1 c_2 \cdots c_k \) is \( c_k^{-1} \cdots c_2^{-1}c_1^{-1} \). Hence \( S_n \) is a group and hence the name Symmetric Group on \( n \) elements.

**Example 3.2.3** The set \( S \) in Example 3.1.2 is a group under \( + \). The identity is \( [0]_N \) and the inverse of \( [i]_N \) is \( [N - i]_N \). We denote this group by \( \mathbb{Z}/N\mathbb{Z} \).
**Example 3.2.4** Let $GL(2, \mathbb{R})$ (called the general linear group of order 2 over $\mathbb{R}$) be the set of invertible $2 \times 2$ matrices with entries in $\mathbb{R}$, i.e.,

$$
GL(2, \mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{R} \quad \text{and} \quad ad - bc \neq 0 \right\}.
$$

The matrix multiplication on $GL(2, \mathbb{R})$ is a binary operation. The identity is

$$
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
$$

The inverse of

$$
\begin{pmatrix} a & b \\ c & d \end{pmatrix}
$$

is given by

$$
\frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.
$$

Hence, $GL(2, \mathbb{R})$ is a group under the matrix multiplication. The matrix multiplication is associative follows by direct computations. Here $\mathbb{R}$ may be replaced by $\mathbb{Q}$ or $\mathbb{C}$.

**Example 3.2.5** Let $A_n$ be the set of even permutations of $S_n$. Now, the identity (1-cycle) is an even permutation. If $\alpha \in A_n$ is even so is $\alpha^{-1}$. The composition is associative. The group $A_n$ is called the alternating group on $\{1, \cdots, n\}$.

**Example 3.2.6** The **affine group** $Aff(1, \mathbb{R})$ consists of all those functions $f : \mathbb{R} \to \mathbb{R}$ (called affine maps) of the form

$$
f(x) = ax + b,
$$

where $a, b \in \mathbb{R}$ with $a \neq 0$. Let us check that $Aff(1, \mathbb{R})$ is a group under composition. If $g : \mathbb{R} \to \mathbb{R}$ has the form $g(x) = cx + d$, where $c \neq 0$, then

$$
f(g(x)) = f(cx + d) = a(cx + d) + b = acx + (ad + b).
$$

Since $ac \neq 0$ this shows that $fg$ is an affine map and so composition is a binary operation on $Aff(1, \mathbb{R})$. The identity function is $e = x$, with $a = 1$ and $b = 0$. The inverse of $f$ is

$$
f^{-1}(x) = a^{-1}x - a^{-1}b.
$$

Here $\mathbb{R}$ may be replaced by $\mathbb{Q}$ or $\mathbb{C}$.  

3.2 Definition of a Group

**Example 3.2.7** Let $\Sigma(2, \mathbb{R})$ be all the matrices in $\text{GL}(2, \mathbb{R})$ whose column sums are 1, that is, $\Sigma(2, \mathbb{R})$ consists of all the nonsingular matrices

\[
\begin{pmatrix}
a & b \\
c & d
\end{pmatrix}
\]

with $a + c = 1 = b + d$. Matrices satisfying the conditions that column sums equal to 1 are called stochastic. We leave it as an exercise to show that the product of two stochastic matrices is stochastic and that the inverse of a stochastic matrix is stochastic.

**Example 3.2.8** For any positive integer $n$, let

\[ G = \{e^{2\pi ik/n} | 0 \leq k < n\} \]

be the set of all $n^{th}$ roots of unity, where

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

We leave it as an exercise to check that $G$ is a group under multiplication of complex numbers.

**Definition 3.2.2** A group $G$ is called abelian if $x \cdot y = y \cdot x$ for all $x, y \in G$.

**Question.** Which of the above groups are abelian?

Now that we have seen many examples of groups. We will now establish some basic properties of groups. Note that whatever standard results we prove for a group apply to all the examples. This is the advantage of studying groups in the abstract setting.

**Lemma 3.2.1 (Basic rules)**

(i) The cancelation laws hold: if either $x \cdot a = x \cdot b$ or $a \cdot x = b \cdot x$ then $a = b$.

(ii) The element $e$ is the unique element in $G$ with $e \cdot x = x \cdot e = x$ for all $x \in G$.

(iii) Given $x \in G$, there is a unique $x' \in G$ such that $x \cdot x' = x' \cdot x = e$.

This unique element $x'$ is called the inverse of $x$ and denoted by $x^{-1}$.

(iv) $(x^{-1})^{-1} = x$ for all $x \in G$.

**Proof of (i).** If $x \cdot a = x \cdot b$, then multiplying by $x^{-1}$ we find that $x^{-1}xa = x^{-1}xb$. But $x^{-1}x = e$ and $ea = a$, we therefore deduce that $a = b$. 

Definition of a Group and Examples

Proof of (ii). If \( e' \) and \( e \) is such that \( ex = xe \) and \( e' x = xe' \) for all \( x \in G \), then treating \( e \) as identity and letting \( x = e' \) we find that
\[
 ee' = e'e' = e'.
\]

Similarly treating \( e' \) as identity and letting \( x = e \), we find that
\[
 e'e = ee' = e.
\]

Comparing the two identities we deduce that \( e = e' \).

Proof of (iii). Let \( x' \) and \( x'' \) be the inverses of \( x \), i.e., \( xx' = x'x = e \) and \( xx'' = x''x = e \). But now,
\[
 x''xx' = (x''x)x' = x'
\]
and
\[
 x''xx' = x''(xx') = x''.
\]
Hence \( x' = x'' \).

Proof of (iv). Now
\[
 (x^{-1})^{-1}x^{-1} = e,
\]
and
\[
 xx^{-1} = e.
\]
Since inverse is unique and both \( x \) and \( (x^{-1})^{-1} \) are both inverses of \( x^{-1} \),
we find that
\[
 (x^{-1})^{-1} = x.
\]

3.3 Laws of Exponents

A group \( G \) can be written additively or multiplicatively. For the multiplicative notation, we usually use \( \cdot \) as the operation. From now on we will write \( xy \) instead of \( x \cdot y \) for multiplicative group. The identity in this case is usually denoted by \( e = 1 = 1_G \). For the additive notation, we write + as the operation and write \( x + y \) for any two elements \( x, y \in G \). Additive notation are usually reserved for abelian groups. The identity we use in this case is \( e = 0 = 0_G \). Most of the time, we will use the multiplicative notation.

The associative law tells us that
\[
 a(a(a)) = (a(a))a = aaa.
\]
3.3 Laws of Exponents

So we may represent the product of three identical elements by \( a^3 \). In fact, we define inductively

\[
a^1 = a \quad \text{and} \quad a^{n+1} = a a^n,
\]

where \( a \in G \) and \( n \) is a positive integer.

We also define \( a^0 = 1 \) and \( a^{-n} = (a^{-1})^n \) for positive integer \( n \). The following Theorem gives us the basic properties of “raising the power of \( a \”).

**Theorem 3.3.1 (Laws of Exponents)** Let \( G \) be a group, let \( a, b \in G \) and let \( r \) and \( s \) be integers.

(i) \((ab)^{-1} = b^{-1}a^{-1}\).

(ii) If \( a \) and \( b \) commute, then \((ab)^r = a^r b^r\).

(iii) \((a^{-1})^r = (a^r)^{-1}\).

(iv) \((a^r)^s = a^{rs}\).

(v) \(a^r a^s = a^{r+s}\).

The additive notation for \( a^n \) is \( na = a + a + \cdots + a \) (since we are adding \( a \) \( n \) times). Written additively, the above facts take the form

(i) \( -(a + b) = (-b) + (-a)\).

(ii) If \( a \) and \( b \) commute, then \( r(a + b) = ra + rb, r \in \mathbb{Z} \).

(iii) \( r(-a) = -(ra), r \in \mathbb{Z} \).

(iv) \( s(ra) = (rs)a, r, s \in \mathbb{Z} \).

(v) \( (ra + sa) = (r + s)a, r, s \in \mathbb{Z} \).

If \( G \) is a finite group, consider the subset \( \{1, a, a^2, a^3, \cdots \} \). Note that since \( G \) is finite, \( a^n = a^m \) for some \( n \) and \( m \). In other words, \( a^{n-m} = 1 \) and hence, there exist a \( k \) such that \( a^k = 1 \). The smallest such \( k \) is called the order of \( a \). If no such power exists, we say that \( a \) has infinite order. The argument given above for finite group shows that every element in a finite group has finite order.

**Lemma 3.3.2** Let \( G \) be a group and \( a \in G \). If \( k \) is the order of \( a \) and \( a^n = 1 \) then \( k | n \).

**Proof.** If \( a^n = 1 \) and \( k \) does not divide \( n \), then \( n = kq + r, 0 < r < k \) by the Division Algorithm. But \( a^k = 1 \) and \( a^n = 1 \) implies that \( a^r = 1 \). Since \( r < k \) this contradicts to the fact that \( k \) is minimal positive number for which \( a^m = 1 \).
Definition of a Group and Examples

Example 3.3.1 If $\alpha \in S_n$ is an $r$-cycle then the order of $\alpha$ is $r$. To see this, it is clear that $\alpha^r = 1$. For we may represent

$$\alpha = (i_1 \alpha(i_1) \cdots \alpha^{r-1}(i_1)).$$

Let $j$ be moved by $\alpha$ then $j = \alpha^k(i_1)$ for some $k$. Therefore $\alpha^r(j) = \alpha^k(\alpha^r(i_1)) = \alpha^k(i_1) = j$. Hence $\alpha^r = 1$. If the order of $\alpha$ is less than $r$, say $l$, then $\alpha^l = 1$. This shows that $\alpha = (i_1 \alpha(i_1) \cdots \alpha^{l-1}(i_1))$, which implies that $\alpha$ is an $l$-cycle with $l < r$, a contradiction.

3.4 Group of Isometries and the orthogonal group

Definition 3.4.1 An isometry of the plane is a function

$$\vartheta : \mathbb{R}^2 \to \mathbb{R}^2$$

that is distance preserving: for all points $P = (a, b)$ and $Q = (c, d)$ in $\mathbb{R}^2$,

$$\|\vartheta(P) - \vartheta(Q)\| = \|P - Q\|.$$  

Example 3.4.1 Given an angle $\theta$, the rotation $R_\theta$ about the origin $O$ is defined as follows: $R(O) = O$; if $P \neq O$ draw the line segment $PO$, rotate it $\theta$ counterclockwise if $\theta > 0$ to $P'O$ as shown in the following diagram. Define $R_\theta(P) = P'$. $R_\theta$ is an example of $\vartheta$.

![Rotation Diagram]

Theorem 3.4.1 Let $\vartheta$ be an isometry of the plane. If distinct points of $P, Q$ and $R$ in $\mathbb{R}^2$ are collinear, then $\vartheta(P), \vartheta(Q)$, and $\vartheta(R)$ are collinear. Hence, if $L$ is a line, then $\vartheta(L)$ is a line.

Proof. Suppose $P, Q$ and $R$ are collinear. Let us assume that $R$ is between $P$ and $Q$. Now,

$$\|P - Q\| = \|P - R\| + \|R - Q\|.$$  

(3.4.1)
3.4 Group of Isometries and the orthogonal group

If \( \vartheta(P), \vartheta(Q), \vartheta(R) \) are not collinear then they form vertices of a triangle and

\[ \|\vartheta(P) - \vartheta(Q)\| < \|\vartheta(P) - \vartheta(R)\| + \|\vartheta(R) - \vartheta(Q)\|. \]

Since \( \vartheta \) is an isometry, this gives

\[ \|P - Q\| < \|P - R\| + \|R - Q\| \]

and this contradicts (3.4.1).

**Theorem 3.4.2** Every isometry \( \vartheta \) of \( \mathbb{R}^2 \) fixing the origin is a linear transformation.

**Proof.** We need to show two things:

(i) \( \vartheta(rP) = r\vartheta(P) \) for all \( r \in \mathbb{R} \).

(ii) \( \vartheta(P + Q) = \vartheta(P) + \vartheta(Q) \).

Statement (i) is true for \( r = 0 \). Let \( r > 0 \). Then \( O, P \) and \( rP \) are collinear. By Theorem 3.4.1, \( \vartheta(rP) \) must lie on the line connecting \( \vartheta(P) \) and \( O \).

Now since

\[ \|\vartheta(rP)\| = \|rP\| = r\|P\| = r\|\vartheta(P)\|, \]

we find that \( \vartheta(rP) = -r\vartheta(P) \) or \( r\vartheta(P) \) (this is true since \( \vartheta(rP) \) and \( \pm r\vartheta(P) \) lie on the same line and is of equal distance away from \( O \)). If

\[ \vartheta(rP) = -r\vartheta(P), \]

then

\[ \|\vartheta(rP) - \vartheta(P)\| = \| - r\vartheta(P) - \vartheta(P)\| = |r + 1|\|P\|. \]

On the other hand,\n
\[ \|\vartheta(rP) - \vartheta(P)\| = \|rP - P\| = |r - 1|\|P\|. \]

This is impossible, since \( r-1 = r+1 \) implies that \( 0 = 2 \) and \( r-1 = -1-r \) implies that \( r = 0 \) (but \( r > 0 \)). Hence, \( \vartheta(rP) = rP \).

For \( r < 0 \), it suffices to show that \( \vartheta(-P) = -\vartheta(P) \). The result is clear if \( P = O \). Suppose that \( P \neq O \). Note that \( \|\vartheta(-P)\| = \|-P\| = \|P\| = \|\vartheta(P)\| \). This means that \( \vartheta(-P) = \pm \vartheta(P) \). If \( \vartheta(-P) = \vartheta(P) \), this would mean that

\[ \|\vartheta(-P) - \vartheta(P)\| = \|O\| = 0. \]

But since \( \vartheta \) is an isometry, \( \|\vartheta(-P) - \vartheta(P)\| = 2\|P\| \) Hence \( P = O \), a contradiction. Therefore, \( \vartheta(-P) = -\vartheta(P) \).
Assume $r < 0$. Let $r = -r'$, $r' > 0$. Then $\vartheta(r'P) = r'\vartheta(P)$. So $\vartheta(-rP) = -r\vartheta(P)$. But we have just shown that $\vartheta(-rP) = -\vartheta(rP)$ and hence $\vartheta(rP) = r\vartheta(P)$.

This proves (i).

Now, we prove (ii). If $P = -Q$ or $P = O$ then the result is obvious by (i). Suppose $P \neq -Q$ and $P \neq O$.

If $P, O$ and $Q$ are collinear, then $P + Q, P$ and $Q$ are collinear and

$$\|\vartheta(P + Q) - \vartheta(Q)\| = \|\vartheta(P)\|.$$  

Therefore,

$$\vartheta(P + Q) = \pm(\vartheta(P)) + \vartheta(Q).$$

If $\vartheta(P + Q) = -\vartheta(P) + \vartheta(Q)$ then since $\vartheta(-P) = -\vartheta(P)$,

$$\|\vartheta(P + Q) - (-\vartheta(P))\| = \|\vartheta(Q)\|.$$  

Simplifying using the fact that $\vartheta$ is an isometry we have

$$\|P + Q + P\| = \|Q\|.$$  

This implies that $P = -Q$ or $P = O$.

Finally, if $P, Q, O$ are not collinear, then let $S = P + Q$. The vertices $P, Q, S, O$ form a parallelogram. Under $\vartheta$, the images $\vartheta(P), \vartheta(Q), \vartheta(S)$ and $\vartheta(O)$ are vertices of a parallelogram. On the other hand, if a parallelogram has three vertices $\vartheta(O), \vartheta(P)$ and $\vartheta(Q)$, then the fourth vertex must be the point $\vartheta(P) + \vartheta(Q)$. Hence,

$$\vartheta(S) = \vartheta(P + Q) = \vartheta(P) + \vartheta(Q).$$

As a corollary, we have

**Corollary 3.4.3** Every isometry $\vartheta : \mathbb{R}^2 \to \mathbb{R}^2$ is a bijection and every isometry fixing $O$ is a nonsingular linear transformation.

**Proof.** First assume that $\vartheta(O) = O$. By Theorem 3.4.2, $\vartheta$ is a linear transformation. Let $e_1 = (1, 0)$ and $e_2 = (0, 1)$ be basis of $\mathbb{R}^2$. Since $e_1, O$ and $e_2$ are not collinear, $P := \vartheta(e_1), O$ and $Q := \vartheta(e_2)$ are also not collinear. Hence, $P$ and $Q$ form a basis for $\mathbb{R}^2$. Let $\psi$ be defined as

$$\psi(aP + bQ) = ae_1 + be_2, a, b \in \mathbb{R}.$$  

Then

$$\psi \circ \vartheta(\alpha e_1 + \beta e_2) = ae_1 + be_2,$$
and $\psi \circ \vartheta = \text{id}$. Also,

$$\vartheta \circ \psi(aP + bQ) = aP + bQ.$$ 

Hence $\psi$ is the inverse of $\vartheta$ and so $\vartheta$ is a bijection.

Suppose $\vartheta$ does not fix $O$, but instead, $\vartheta(O) = U$. Let $\tau(P) = P - U$. Then $\tau(U) = O$ and $\tau \circ \vartheta$ fixes $O$ and it is a bijection. Hence $\vartheta$ is a bijection.

**Definition 3.4.2** The set of isometries is denoted by $\text{Isom}(\mathbb{R}^2)$. The set of all isometries of the plane which fix the origin is denoted by $O(2, \mathbb{R})$.

**Exercise.** Check that $\text{Isom}(\mathbb{R}^2)$ and $O(2, \mathbb{R})$ are both groups under the composition of maps.

### 3.5 Symmetry group of a figure and Dihedral groups

If $\triangle$ is a triangle with vertices $P, Q,$ and $R$. If $\vartheta$ is an isometry such that $\vartheta(\triangle) = \triangle$, then we see that $\vartheta$ permutes the vertices of $\triangle$.

How many $\vartheta$’s are there?

We call the set of isometries $\vartheta$ fixing $\triangle$ the symmetry group of $\triangle$. In general, we have

**Definition 3.5.1** Given a figure $\Omega$. The symmetry group $\Sigma(\Omega)$ of a figure $\Omega$ in the plane is the set of all isometries $\vartheta$ of the plane with $\vartheta(\Omega) = \Omega$. The elements of $\Sigma(\Omega)$ are called the symmetries of $\Omega$.

**Example.** Determine the symmetry group of $\Omega$ when $\Omega$ is a pentagon.

In general, we have the following result for $\Omega_n$ when $\Omega_n$ is a regular polygon of $n$ sides.
Theorem 3.5.1 Let $\Omega_n$ be a regular polygon with $n$ sides and $n$ vertices given by $v_{[0]_n}, \ldots, v_{[n-1]_n}$ (labeled in an anticlockwise manner, and we are indexing the vertices using elements in $\mathbb{Z}/n\mathbb{Z}$). Let the center of $\Omega_n$ be the origin $O$. Let $a$ be the anticlockwise rotation about $O$ at an angle $2\pi/n$ and $b$ be the reflection about the line joining $O$ and one of the vertices, say, $v_{[0]_n}$. Then

$$\Sigma(\Omega_n) := \{a^j, a^jb | 1 \leq j \leq n\}.$$ 

Remarks. The distinct elements $a$ and $b$ satisfy $bab^{-1} = a^{-1}$.

Proof. It is clear that

$$a(v_{[j]_n}) = v_{[i+1]_n}.$$ 

Furthermore,

$$b(v_{[i]_n}) = v_{[n-i]_n}.$$ 

Note that the order of $a$ is $n$. For if $a^j(v_{[i]_n}) = v_{[i+j]_n}$ and $v_{[i]_n} = v_{[j]_n}$, then $n|j$ and the smallest such $j$ is $n$. The reflection $b$ has order 2. Note that for $0 \leq j \leq n-1$, $a^j$ and $a^jb$ are all distinct.

We now show that there are no more symmetries of $\Omega_n$. To prove this, we first show that if $O, P$ and $Q$ are noncollinear points, and if $\vartheta$ and $\psi$ are isometries of the plane such that $\vartheta(P) = \psi(P)$ and $\vartheta(Q) = \psi(Q)$ then $\vartheta = \psi$. To show the above, we note that an isometry is a linear transformation and a linear transformation is determined by its action on the standard basis. If $e_1 = mP + nQ$ and $e_2 = lP + kQ$ then the relation satisfied by $\psi$ and $\psi$ are $P$ and $Q$ shows that the action of $\vartheta$ and $\psi$ on $e_1$ and $e_2$ are the same. Hence $\vartheta = \vartheta$.

Returning to our proof of Theorem 3.5.1, let $\vartheta$ be a symmetry of $\Omega_n$. Then $\vartheta(v_{[0]_n}) = v_{[j]_n}$ for some $j$. Since $\vartheta$ preserves distance, $\vartheta(v_{[1]_n}) = v_{[j+1]_n}$ or $v_{[j-1]_n}$. Suppose $\vartheta(v_{[i]_n}) = v_{[j+1]_n}$. Since $a^j(v_{[0]_n}) = v_{[j]_n}$ and $a^j(v_{[1]_n}) = v_{[j+1]_n}$, by the previous paragraph, we conclude that $a^j = \vartheta$. Suppose $\vartheta(v_{[i]_n}) = v_{[j-1]_n}$. Since $a^j(v_{[0]_n}) = v_{[j]_n}$ and $a^j(v_{[1]_n}) = v_{[j+1]_n}$, by the previous paragraph, we conclude that $a^j = \vartheta$. Hence, $a^j = \vartheta$. In conclusion, we have found all the symmetries of $\Omega_n$.

It is straightforward to see that $\Sigma(\Omega_n)$ forms a group. Notice that the elements of $\Sigma(\Omega_n)$ are given by $a^ib^j$, $1 \leq i \leq n$, $j = 0, 1$. But we know that if $a^ib^j$ and $a^kb^j$ are both in $\Sigma(\Omega_n)$, then $a^ib^ja^kb^j \in \Sigma(\Omega_n)$ and so, $a^ib^ja^kb^j$ must be of the form $a^ib^j$. Therefore, it is necessary
3.5 Symmetry group of a figure and Dihedral groups

to find a relation that allows one to move $b^j$ to the right-hand side of $a^k$. The relation is $bab^{-1} = bab = a^{-1}$. Note that $bab^{-1}(v_{[0]}) = v_{[n-1]}$ and $bab^{-1}(v_{[1]}) = ba(v_{[n-1]}) = b(v_{[n]}) = v_{[0]}$. On the other hand, $a^{-1}(v_{[0]}) = v_{[n-1]}$ and $a^{-1}(v_{[1]}) = v_{[0]}$. Hence $bab = a^{-1}$. This means that $ba = a^{-1}b$ and allows us to rewrite expression such as $a^ib^ja^kb^\ell$ in the form $a^ib^\ell$.

**Definition 3.5.2** A group given by the relations

\[
\{a^n = 1, b^2 = 1, bab^{-1} = a^{-1}\}
\]

is called a **Dihedral group** of order $2n$. The notation for this group is $D_{2n}$ and it has $2n$ elements.

We have shown in Theorem 3.5.1 that $\Sigma(\Omega_n)$ is a Dihedral group of order $2n$. 
4

Subgroups, Cosets and Lagrange Theorem

4.1 Subgroups

Definition 4.1.1 A subset $H$ of a group $G$ is a **subgroup** if

1. $1 \in H$,
2. if $x, y \in H$ then $x \cdot y \in H$,
3. if $x \in H$, then $x^{-1} \in H$.

It is easy to see that $H$ is a group; the function $\cdot$ is a binary operation on $H$. The identity is in $H$, and so, $1 \cdot x = x = x \cdot 1$ for all $x \in H$. For every $x \in H$, the inverse is in $H$. Finally the associative law holds because it holds for the group $G$ and hence for the subset $H$.

In order to check if a subset $H$ of a group $G$ is a subgroup, we do not need to check the four axioms for group. Instead we only need to check one condition for any two arbitrary elements of $H$. This result is given as follow:

**Theorem 4.1.1** A subset $H$ of a group $G$ is a subgroup if and only if $H$ is nonempty and whenever $x, y \in H$, then $x \cdot y^{-1} \in H$.

**Proof.** If $H$ is a subgroup then it is nonempty. Let $x, y \in H$. Since $H$ is a group $y^{-1} \in H$ and $x \cdot y^{-1} \in H$.

Conversely, let $H$ be nonempty. The associative law holds for elements in $H$ since it holds for elements in $G$. The identity $e$ is in $H$ since if $x \in H$, $x \cdot x^{-1} = e \in H$. Also, for each $x \in H$, we have $e \cdot x^{-1} \in H$ since $e \in H$. Hence $x^{-1} \in H$. Finally, for $x, y \in H$, $y^{-1} \in H$ and so $x \cdot (y^{-1})^{-1} = x \cdot y \in H$. Hence $\cdot$ is a binary operation on $H$. $H$ is therefore a group.
Example 4.1.1 The alternating group $A_n$ in Example 3.2.5 is a subgroup of $S_n$.

Example 4.1.2 The four permutations

$$V = \{1, (1 \ 2)(3 \ 4), (1 \ 3)(2 \ 4), (1 \ 4)(2 \ 3)\}$$

forms a group. This is a subgroup of $S_4$ and it is called the Klein four group.

Example 4.1.3 The set

$$T = \{am + bn | a, b \in \mathbb{Z}\}$$

with $a, b \in \mathbb{Z}$ is a subgroup of $(\mathbb{Z}, +)$. First, $T$ is nonempty. One checks immediately that for $u = am + bn$, $v = as + bt \in T$,

$$u - v = a(m - s) + b(n - t) \in T.$$

4.2 Cyclic groups

Definition 4.2.1 If $G$ is a group and $a \in G$, write

$$< a >= \{a^n | n \in \mathbb{Z}\} = \{\text{all powers of } a\};$$

$< a >$ is called the cyclic subgroup of $G$ generated by $a$.

Note that $< a >$ is a subgroup of $G$.†

In the event when the cyclic subgroup $< a >$ is the group $G$, i.e. $G = < a >$, we say that $G$ is cyclic or that $G$ is a cyclic group. The element $a$ is called a generator of $G$. In this section, we give some basic properties of cyclic groups.

Theorem 4.2.1 Every subgroup $S$ of a cyclic group $G = < a >$ (not necessarily finite) is cyclic.

Proof. Let $S$ be a nontrivial subgroup of $G$. Then $S$ contains powers of $a$. In particular $S$ contains some positive powers of $a$. Let $T = \{k|a^k \in S\}$. By the least integer axiom, there is a least element $m$ in $T$. We claim that $S = < a^m >$. Clearly $< a^m > \subset S$. Suppose $s \in S$ and $s = a^j$ where $m \nmid j$ (if $m|j$ then $a^j \in < a^m >$). By the division algorithm,†

† Please verify this.
Subgroups, Cosets and Lagrange Theorem

\[ j = qm + r, \ 0 < r < m . \] Hence, \( a^r = a^{j-nq} \in S \). Since \( r \in T \) and \( r < m \), this contradicts the minimality of \( m \). Hence \( S = < a^m > \) is cyclic.

**Example 4.2.1** The group \((\mathbb{Z}, +)\) is a cyclic group generated by 1. The subgroup

\[ T = \{am + bn | a, b \in \mathbb{Z}\} \]

with \( a, b \in \mathbb{Z} \) is therefore cyclic. This implies that \( T = d\mathbb{Z} \) for some positive integer \( d \). We show that \( d = (a, b) \). Since \( a, b \in T (= d\mathbb{Z}) \), we see that \( d \mid a \) and \( d \mid b \) and hence, \( d \) is a common divisor of \( a \) and \( b \). Next, since \( d \in T \), \( d = au + bv \) for some integers \( u, v \). Suppose \( c \mid a \) and \( c \mid b \), then \( c \mid (au + bv) = d \). Hence \( d \) satisfies all three conditions in the definition of greatest common divisor of \( a \) and \( b \) and therefore, \( d = (a, b) \).

**Theorem 4.2.2** Let \( G \) be a finite cyclic group generated by \( a \). If the order of \( a \) is \( n \), then \( a^k \) is a generator of \( G \) if and only if \( (k, n) = 1 \).

**Proof.** If \( (k, n) = 1 \), then \( 1 = kx + ny \) for some integer \( x \) and \( y \). Since \( a^n = 1 \), we have \( a^{kx+ny} = a^{kx} = a \). Hence \( a \in < a^k > \) and \( G = < a^k > = < a > \). Therefore \( a^k \) is a generator of \( G \).

If \( a^k \) is a generator of \( G \), then \( a = a^{ks} \) for some integer \( s \). Hence, \( a^{1-ks} = 1 \). Since \( n \) is the order of \( a \), by Lemma 3.3.2, we conclude that \( n \mid (1 - ks) \). In order words

\[ 1 = ks + ny \]

for some integer \( y \). But this means that \( (k, n) = 1 \) by Theorem 1.4.2.

**Corollary 4.2.3** The number of generators of a cyclic group of order \( n \) is \( \varphi(n) \).

**Definition 4.2.2** If \( G \) is a finite group we write \( |G| \) for the number of elements in \( G \). This is called the order of the group \( |G| \).

**Warning!** Do not confuse the order of a group with the order of an element (see the paragraph before Lemma 3.3.2).

When \( G = < a > \) is a finite cyclic group, the order \( |G| \) is the same as the order of \( a \) (see the theorem below). Other than this, one should always treat the order of a group and order of element as two different quantities.
Theorem 4.2.4 Let $G = \langle a \rangle$. Then $|G|$ is the order of the element $a$.

Proof. Let $n$ be the order of $a$. Then the elements in the set $S = \{a, a^2, \cdots, a^n\}$ are all distinct. For otherwise, $a^i = a^j$ would imply that $a^{i-j} = 1$ and $i-j < n$ contradicting the fact that $n$ is the order of $a$. Clearly, $S \subseteq G$. Let $g \in G$. Then $g = a^m$. By division algorithm, $m = nq + r, 0 < r \leq n$. So, $a^m = a^{nq+r} = a^r$. Hence $G \subseteq S$. Therefore $S = G$ and $|G| = n$.

We end this section with another property of cyclic groups.

Theorem 4.2.5 Let $G$ be a group of order $n$. If $G$ is cyclic then $G$ has a unique subgroup of order $d$ for each divisor $d$ of $n$.

Proof. Suppose $G = \langle a \rangle$ is a cyclic group. Let $d|n$. We claim that $|\langle a^{n/d} \rangle| = d$. By Theorem 4.2.4 it suffices to show that $d$ is the order of $a^{n/d}$, i.e., $d$ is the smallest positive integer such that $(a^{n/d})^k = 1$ (clearly $d$ satisfies this relation.) Let $r$ be any integer such that $a^{nr/d} = 1$. Since $n$ is the order of $a$, $n|r/d$, or $nr/d = nk$. This implies $r \geq d$.

To prove uniqueness, let $C$ be a subgroup of $G$ of order $d$. Then by Theorem 4.2.1, $C$ is cyclic. Suppose $C = \langle x \rangle$. Now $x = a^m$ for some $m$. Then $a^{md} = 1$ implies that $n|md$. Therefore,

$$md = nk.$$  

Hence,

$$x = (a^{n/d})^k.$$  

Therefore, $C \subset \langle a^{n/d} \rangle$. Since both $C$ and $\langle a^{n/d} \rangle$ have the same number of elements,

$$C = \langle a^{n/d} \rangle.$$  

4.3 Cosets and Lagrange’s Theorem

It turns out that the order of a subgroup $H$ of $G$ is not arbitrary. It is restricted by the order of the group $G$. We will show that if $H$ is a subgroup of $G$, then $|H|$ divides $|G|$. This is known as the Lagrange Theorem. To prove the Lagrange Theorem, we will need the concept of a coset.
Subgroups, Cosets and Lagrange Theorem

Definition 4.3.1 If $H$ is a subgroup of a group $G$ and $a \in G$ then the left coset $aH$ is the subset $aH$ of $G$, where

$$aH = \{ah | h \in H\}.$$  

Definition 4.3.2 If $H$ is a subgroup of a group $G$ and $a \in G$ then the right coset $Ha$ is the subset $Ha$ of $G$, where

$$Ha = \{ha | h \in H\}.$$  

Of course, $a = a1 \in aH$. Cosets are not subgroups in general. For example, if $a \notin aH$, then $1 \notin aH$; otherwise, $1 = ah$ for some $h$ and this implies that $a = h^{-1} \in H$.

Example 4.3.1 In additive notation, the cosets take the form $a + H$. Let $G = \mathbb{Z}$ is a group. Now $H = n\mathbb{Z}$ is a subgroup of $G$ (Check!). The cosets $a + H$ are therefore the set of numbers

$$a + H = \{a + nk | k \in \mathbb{Z}\}.$$  

This is the same as the subset $[a]$ defined in Example 3.1.2. In this case, there are exactly $n$ cosets.

Example 4.3.2 If $G = S_3$ and $H = \langle (1\ 2) \rangle$, there are exactly three cosets of $H$, namely,

$$H = \{(1\ 2)\},$$

$$(1\ 3)H = \{(1\ 3), (1\ 2\ 3)\}$$

$$(2\ 3)H = \{(2\ 3), (1\ 3\ 2)\},$$  

each of which has size 2.

The left coset $aH$ may not necessarily be equal to $Ha$, can you find an example?

Note that in our examples, our cosets are all disjoint. The next result shows that indeed different cosets of a given subgroup do not overlap.

Lemma 4.3.1 Let $H$ be a subgroup of a group $G$ and let $a, b \in G$.

(i) $aH = bH$ if and only if $b^{-1}a \in H$. In particular $aH = H$ if and only if $a \in H$.

(ii) If $aH \cap bH \neq \emptyset$, then $aH = bH$.

(iii) $|aH| = |H|$ for all $a \in G$. 

Proof of (i). Assume that \( aH = bH \). Then \( a = bh \) for some \( h \in H \). Hence \( b^{-1}a = h \in H \). Conversely suppose \( b^{-1}a \in H \). Let \( x \in aH \). Then \( x = ah_1 \) for some \( h_1 \in H \). But \( b^{-1}x = b^{-1}ah_1 \in H \) and hence, \( b^{-1}x \in H \) and \( x \in bH \). This shows that \( aH \subseteq bH \). Next if \( y \in bH \), then \( y = bh_2 \) for some \( h_2 \in H \). This shows that

\[
a^{-1}y = a^{-1}bh_2 = (b^{-1}a)^{-1}h_2 \in H
\]
since \( b^{-1}a \in H \). Hence \( y \in aH \) and \( aH \subseteq bH \).

Proof of (ii). If \( aH \cap bH \neq \emptyset \) then there exist \( h, h' \) such that \( ah = bh' \) and so, \( b^{-1}a = h'h^{-1} \in H \). By (i), we conclude that \( aH = bH \).

Proof of (iii). The function \( f : H \to aH \) which sends \( h \mapsto ah \) is a bijection. (Check!) Therefore \( H \) and \( aH \) has the same number of elements.

**Theorem 4.3.2 (Lagrange’s Theorem)** If \( H \) is a subgroup of a finite group \( G \), then \(|H| \) is a divisor of \(|G| \).

**Proof.** Let \( G = \cup_{a \in G} aH \). Note that the union is disjoint for if \( aH \cap bH \neq \emptyset \), \( aH = bH \). Since \( G \) is finite, the right hand side is a union of finite number of sets. Now by counting the number of elements on both sides and assuming that there are altogether \( t \) cosets, we find that

\[
|G| = \sum_{i=1}^{t} |a_i H|.
\]

Now by Lemma 4.3.1 (iii), we conclude that

\[
|G| = t|H|.
\]

Hence \(|H| \) divides \(|G| \).

**Definition 4.3.3** The **index** of a subgroup \( H \) in \( G \), denoted by \([G : H]\), is the number of cosets of \( H \) in \( G \).

Note that when \( G \) is a finite group,

\[
\]

We also note that the index may still be defined when \( G \) is infinite. For example, \([\mathbb{Z} : n\mathbb{Z}] = n \) if \( n \) is a positive integer.
4.4 Lagrange’s Theorem and the cyclic groups

We can now prove the converse of Theorem 4.2.5.

**Theorem 4.4.1** Let \( G \) be a group of order \( n \). If there is at most one cyclic subgroup of order \( d \) where \( d \mid n \), then \( G \) is cyclic.

**Proof.** Let \( C_d = \{ g \in G \mid g \) has order \( d \} \). Note that \( d \mid n \) by Lagrange’s Theorem. If \( C_d \) is nonempty, then by Corollary 4.2.3 \( |C_d| = \varphi(d) \). Hence

\[
n = \sum_{d \mid n} |C_d| \leq \sum_{d \mid n} \varphi(d)
\]

since there is at most one cyclic group of order \( d \). We will prove later that

\[
\sum_{d \mid n} \varphi(d) = n. \quad (4.4.1)
\]

If the above is true, then each \( |C_d| = \varphi(d) \). In particular \( |C_n| \neq 0 \) and so \( G \) is cyclic.

We now show (4.4.1). Let \( S = \{1, 2, \cdots, n\} \). Then

\[
S = \cup G_d,
\]

where

\[
G_d = \{ k \mid (k, n) = d \}.
\]

One checks that \( G_d \cap G_d' = \phi \) if \( d \neq d' \). Also \( |G_d| = \varphi(n/d) \). Hence,

\[
n = \sum_{d \mid n} \varphi\left(\frac{n}{d}\right) = \sum_{d \mid n} \varphi(d).
\]

4.5 Lagrange’s Theorem and Euler’s Theorem

We end by giving a new proof of Fermat’s Little Theorem and its generalization, namely, Euler’s Theorem.

We first prove the following:

**Theorem 4.5.1** Let \( G \) be a finite group. If \( g \in G \), then

\[
g^{[G]} = 1_G.
\]

**Proof** Let \( g \neq 1_G \) be an element in the finite group \( G \). Then the cyclic subgroup generated by \( g \), namely, \( \langle g \rangle \) has the same order as the order of \( g \) by Theorem 4.2.4. Let the order of \( g \) be \( d \). Since \( d \) is also the
order of the subgroup $\langle g \rangle$, $d | |G|$ by Lagrange’s Theorem. Therefore, $|G| = dk$ for some positive integer $k$. Now,

$$g^{|G|} = g^{dk} = (g^d)^k = 1^k_G = 1_G,$$

since $d$ is the order of the element $g$.

\[\Box\]

**Theorem 4.5.2 (Euler’s Theorem)** Let $N$ be a positive integer and $a$ be such that $(a, N) = 1$. Then

$$a^{\varphi(N)} \equiv 1 \pmod{N}.$$

**Proof** Let

$$S^* = \{[a] | (a, N) = 1, 1 \leq a \leq N\}.$$

Note that each element in $S^*$ has an inverse, i.e., for each $[i] \in S^*$ there is a $[j] \in S^*$ such that $ij = 1$. The existence of $[j]$ follows from the solvability of

$$ij \equiv 1 \pmod{N}.$$

The congruence is solvable since $(i, N) = 1$ implies the existence of integers $j$ and $v$ such that $ij + Nv = 1$. Hence, $S^*$ is a group. Note that by definition of $\varphi(N)$, the number of elements in $S^*$ is $\varphi(N)$. By Theorem 4.5.1,

$$[a]^{\varphi(N)} = [1],$$

and this is Euler’s Theorem. Note that when $N = p$, $\varphi(N) = p - 1$ and we recover Fermat’s Little Theorem. \[\Box\]
5
Homomorphisms

5.1 Definitions and Examples
We begin with an example.

Example 5.1.1 The group \( \mathbb{Z}/N\mathbb{Z} \) in Example 3.2.3 under addition is cyclic of order \( N \). It is generated by \([1]\). The group in Example 3.2.8 is also cyclic (but multiplicative) of order \( N \), generated by \( e^{2\pi i/n} \). They look similar via the map of the generator \([1] \mapsto e^{2\pi i/n} \), hence, \([k] \mapsto e^{2\pi ik/n}\). The notion of homomorphism allows us to compare different groups and identify those which are “similar”.

Definition 5.1.1 If \((G, \cdot)\) and \((H, \circ)\) are groups, then a function \( f : G \to H \) is a homomorphism if for all \( x, y \in G \),

\[
f(x \cdot y) = f(x) \circ f(y).
\]

If \( f \) is also a bijection then \( f \) is called an isomorphism. Two groups \( G \) and \( H \) are said to be isomorphic, denoted by \( G \simeq H \), if there exists an isomorphism \( f : G \to H \) between them.

Example 5.1.2 Let \( \alpha \in G \). The map

\[
\gamma_\alpha : G \to G
\]

defined by

\[
\gamma_\alpha(g) = \alpha g \alpha^{-1}
\]
is an isomorphism.
**Example 5.1.3** Let \( \mathcal{A} \) be the set of all \( 2 \times 2 \) matrices of the form
\[
A = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix},
\]
where \( a \neq 0 \). It is easy to see that \( \mathcal{A} \) is a subgroup of \( GL(2, R) \), the group of invertible \( 2 \times 2 \) matrices with entries in \( R \). In Example 3.2.6, we define the group \( \text{Aff}(1, R) \) whose elements are the functions \( f(x) = ax + b \), where \( a \neq 0 \). We claim that
\[
\varphi : \text{Aff}(1, R) \to \mathcal{A}
\]
defined by \( f \mapsto A \) is an isomorphism. If \( g : R \to R \) has the form \( g(x) = cx + d \) then
\[
fg(x) = acx + (ad + b),
\]
so that
\[
\varphi(fg) = \begin{pmatrix} ac & ad + b \\ 0 & 1 \end{pmatrix}.
\]
On the other hand,
\[
\varphi(f)\varphi(g) = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c & d \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} ac & ad + b \\ 0 & 1 \end{pmatrix}.
\]
Therefore, \( \varphi \) is a homomorphism. To see that \( \varphi \) is an isomorphism, we simply observe that the inverse of \( \varphi \) is given by
\[
\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \mapsto f,
\]
where \( f(x) = ax + b \). It follows that
\[
\mathcal{A} \cong \text{Aff}(1, R).
\]

**Example 5.1.4** We now show that the stochastic group in Example 3.2.7 is isomorphic to the affine group in Example 3.2.6. In the light of the previous example, we may identity \( \text{Aff}(1, R) \) with a subgroup of \( GL(2, R) \). Now suppose
\[
M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}
\]
is stochastic, then \( a + c = 1 = b + d \). Let
\[
Q = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.
\]
Then
\[
QMQ^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ a+c & b+d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} 
= \begin{pmatrix} a & b \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a-b & b \\ 0 & 1 \end{pmatrix}.
\]

Now, since \(QMQ^{-1}\) is invertible, this implies that \(a - b \neq 0\). Hence
\[QMQ^{-1} \in \mathcal{A}.\]

If \(B \in \mathcal{A}\), then
\[
Q^{-1}BQ = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} a+b & b \\ 1-a-b & 1-b \end{pmatrix} \in \Sigma(2, \mathbb{R}).
\]

### 5.2 Basic facts about homomorphisms

**Lemma 5.2.1** Let \(f : G \to H\) be a homomorphism of groups.

(i) \(f(1) = 1\);

(ii) \(f(x^{-1}) = f(x)^{-1}\);

(iii) \(f(x^n) = f(x)^n\) for all \(n \in \mathbb{Z}\).

**Proof of (i).** Since \(f\) is a homomorphism,
\[
f(1) = f(1 \cdot 1) = f(1)f(1).
\]
Therefore
\[
f(1)^{-1}f(1) = f(1)^{-1}f(1)f(1),
\]
which implies that \(f(1) = 1\).

**Proof of (ii).**
\[
f(x)f(x^{-1}) = f(x \cdot x^{-1}) = f(1) = 1.
\]
Hence, \(f(x^{-1}) = f(x)^{-1}\).

**Proof of (iii).** If \(n \geq 0\), then
\[
f(x^n) = f(x)f(x) \cdots f(x),
\]
where the right hand side consists of \(n\) copies of \(f(x)\). Hence \(f(x^n) = f(x)^n\). If \(n < 0\), then \(-n > 0\) and
\[
f(x^{-n}) = f(x)^{-n} = (f(x)^n)^{-1}.
\]
But 
\[ f(x^{-n}) = f((x^n)^{-1}) = f(x^n)^{-1}. \]

Hence, 
\[ f(x^n)^{-1} = (f(x)^n)^{-1} \]
and we complete the proof.

**Example 5.2.1** We have seen in the beginning of this Chapter that \( \mathbb{Z}/N\mathbb{Z} \) and the group generated by \( e^{2\pi i/N} \) are isomorphic. We now show that every cyclic group is isomorphic to \( \mathbb{Z}/N\mathbb{Z} \). We only need to define \( f([k]_N) = g^k \) where \( G = \langle g \rangle \) is a cyclic group of order \( N \). We then check that the map is a well defined and that it is a bijective homomorphism.

Although cyclic groups of the same order are isomorphic, groups with the same order are not isomorphic in general. The following example illustrate this point.

**Example 5.2.2** Let \( G \) be a cyclic group of order 4. Let \( H \) be the group 
\[ H = \{(1), (1 \ 2)(3 \ 4), (1 \ 4)(2 \ 3), (1 \ 3)(2 \ 4)\}. \]
These two groups are not isomorphic because all the elements in \( H \) has order 2 but there exist an element in \( G \) of order 4. More precisely if there is an isomorphism \( f \) from \( G \) to \( H \), then \( f \) must be injective. Let \( g \) be a generator of \( G \) and \( f(g) = a \). If \( a = 1 \) then \( f(1) = f(g) \) and \( f \) is not injective. Hence \( a \neq 1 \) and so \( a^2 = 1 \) since every element in \( H \) is of order 2. But this means that \( f(g^2) = 1 = f(1) \). Since \( g^2 \neq 1 \), \( f \) is not injective. Hence, \( G \nsubseteq H \).

### 5.3 Kernel and Image of a homomorphism

**Definition 5.3.1** If \( f \) is a homomorphism from \( G \) to \( H \), define the kernel of \( f \) to be 
\[ \ker f := \{ x \in G | f(x) = 1 \}. \]
Define the image of \( f \) to be 
\[ \text{im } f := \{ h \in H | h = f(g) \text{ for some } g \in G \}. \]
Note that $f$ is surjective if $\text{im } f = H$. We now show that $f$ is injective if and only if the $\ker f$ is the identity. If $f$ is injective, then $f(x) = 1$ has only one solution, i.e., $x = 1$. Suppose $\ker f = \{1\}$. Let $f(x) = f(y)$.

Then

$$f(xy^{-1}) = f(x)f(y)^{-1} = 1.$$ 

This implies that $xy^{-1} \in \ker f$ and hence, $x = y$. Hence $f$ is injective.

**Example 5.3.1** Let $H = \{\pm 1\}$ be the multiplicative group of two elements. We see from Chapter 4 that the function $\text{sgn}$ from $S_n$ to $H$ is a homomorphism. The kernel contains those permutations such that $\text{sgn}(\alpha) = 1$, or in other words, those $\alpha$’s in $A_n$, the alternating group. The image is clear $\{\pm 1\}$ since odd permutations are mapped to $-1$ while even permutations are mapped to $1$.

**Example 5.3.2** The determinant of a matrix in the group $GL(2, \mathbb{R})$ sends a matrix to $\mathbb{R} - \{0\}$. The kernel is $SL(2, \mathbb{R})$, the group of matrices with determinant equal to $1$ and the image is the whole of $\mathbb{R} - \{0\}$.

**Example 5.3.3** The homomorphism from $\mathbb{Z}$ to $\mathbb{Z}/N\mathbb{Z}$ sending $i$ to $[i]$ has kernel $N\mathbb{Z}$ and image $\mathbb{Z}/N\mathbb{Z}$.

**Definition 5.3.2** A subgroup $K$ of a group $G$ is called a normal subgroup if $k \in K$ and $g \in G$ implies that $gkg^{-1} \in K$. If $K$ is a normal subgroup of $G$, we write $K \triangleleft G$.

**Remarks.**

1. Note that if $G$ is an abelian group then any subgroup is normal in $G$.

2. If $K$ is normal in $G$ then $gK = Kg$, i.e., the left coset and the right coset containing $g$ coincide. Conversely, if $Kg = gK$ for all $g \in G$, then $K$ is normal. So sometimes we also say that if a group $K$ satisfies $Kg = gK$ for all $g \in G$, then $K$ is a normal subgroup of $G$.

3. There are subgroups which are not normal. Let $H = \langle (1 \ 2) \rangle$ of $S_3$. If $\alpha = (1 \ 2 \ 3)$ then

$$\alpha(1 \ 2)\alpha^{-1} = (2 \ 3) \notin H.$$ 

Hence $H$ is not normal.
Example 5.3.4 Define the center of a group $G$, denoted by $Z(G)$, to be

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

So, $Z(G)$ consists of all elements commuting with everything in $G$. We first show that $Z(G)$ is a subgroup. We check this using definitions of groups. First note that $1 \in Z(G)$ since $1 \cdot g = g \cdot 1 = g$. 

If $z \in Z(G)$ and $g \in G$ then

$$zg^{-1} = g^{-1}z$$

and upon taking inverses on both sides, we conclude that

$$gz^{-1} = z^{-1}g$$

and hence $z^{-1} \in Z(G)$. Finally we show that $\cdot$ is a binary operation on $Z(G)$. Take $x, y \in Z(G)$. Then $xgy = xgy = gxy$ since $x, y \in Z(G)$. Hence $xy \in Z(G)$. Associative laws hold for $Z(G)$ since it holds for $G$. It is clear that $Z(G)$ is normal since $gzg^{-1} = z$ and hence $gZ(G) = Z(G)g$ for all $g \in G$.

Example 5.3.5 The Klein four group

$$V = \{(1), (1 2)(3 4), (1 3)(2 4), (1 4)(2 3)\}$$

is a normal subgroup of $S_4$.

Our major Theorem in this section is as follow:

**Theorem 5.3.1** Let $f : G \to H$ be a homomorphism of groups. Then the image of $f$ is a subgroup of $H$ and the kernel of $f$ is a normal subgroup of $G$.

**Proof.** Let $I := \text{im } f$. We first show that $I$ is a subgroup of $H$. Note that $f(1) = 1$ and so, $1 \in I$. If $a, b \in I$, then $f(x) = a$ and $f(y) = b$ for some $a, b \in G$. This implies that $f(xy) = ab$. Hence $ab \in I$. Finally if $a \in I$, then $f(x) = a$ for some $x \in G$. Hence $f(x^{-1}) = a^{-1}$ by the basic property of homomorphism. Hence $a^{-1} \in I$ and we complete the proof.

Next let $K = \ker f$. We show that $K$ is a subgroup of $G$. Note that $1 \in K$, since $f(1) = 1$. Suppose $x, y \in K$. Then $f(x) = f(y) = 1$. Hence $f(xy) = f(x)f(y) = 1$ and hence $xy \in K$. Finally, if $x \in K$, then $f(x) = 1$. This implies that $f(x^{-1}) = 1$ and $x^{-1} \in K$. 


Our final task is to show that $K$ is a normal subgroup of $G$. Let $g \in G$ and $k \in K$. Then

$$f(gkg^{-1}) = f(g)f(k)f(g)^{-1} = f(g)f(g)^{-1} = 1,$$

since $f(k) = 1$. Hence $gkg^{-1} \in K$ and $K \triangleleft G$.

**Example 5.3.6** (Rotman, p. 108, Exercise 2.65(a)). Show that $A_n$ is a normal subgroup of $S_n$ of order $n!/2$.

This follows from the homomorphism

$$\text{sgn} : S_n \to \{ \pm 1 \}.$$ 

The kernel of this map is $A_n$ and since the kernel of a homomorphism is a normal subgroup, we conclude that $A_n \triangleleft S_n$. 
6 Quotient Groups and the Isomorphism Theorems

6.1 Cosets of Normal Subgroup and Quotient group
Recall that a subgroup $H$ is normal in $G$ if for every $g \in G$ and $h \in H$, $ghg^{-1} \in H$.

**Theorem 6.1.1** Let $H \triangleleft G$. Then the collection of cosets $\{a, H\}$ forms a group under the operation

$$aH \circ bH = (a \ast b)H,$$

where $\ast$ is the operation from the group $G$. This group is denoted by $G/H$ and is called the quotient group of $G$ by $H$.

Before we continue, let us show that the operation defined above for cosets of a “non-normal” subgroup may not always be well-defined, i.e., it could be a many to one map.

**Example 6.1.1** Consider $H = \{(1), (1\ 2 \ 3), (1\ 3 \ 2)\} \leq A_4$. The cosets are

- $H$,
- $(1\ 2)(3\ 4)H = \{(1\ 2)(3\ 4), (2\ 4\ 3), (1\ 4\ 3)\}$
- $(1\ 3)(2\ 4)H = \{(1\ 3)(2\ 4), (1\ 4\ 2), (3\ 4\ 2)\}$,
- $(1\ 4)(3\ 2)H = \{(1\ 4)(3\ 2), (1\ 3\ 4), (1\ 2\ 4)\}$.

Now,

$$(1\ 3)(2\ 4)H \circ (1\ 4)(3\ 2)H = (1\ 3)(2\ 4)(1\ 4)(3\ 2)H = (1\ 2)(3\ 4)H.$$
Quotient Groups and the Isomorphism Theorems

One the other hand, if we choose a different representative for the coset 
\((1 \ 3)(2 \ 4)H\), namely, \((1 \ 4 \ 2)H = (1 \ 3)(2 \ 4)H\), we find that

\[(1 \ 4 \ 2) \circ H (1 \ 4)(3 \ 2)H = (1 \ 2 \ 3)H = H.\]

This shows that the operation is inconsistent since \(H \neq (1 \ 2)(3 \ 4)H\).

This is precisely the reason for introducing normal subgroups. When
\(H\) is normal in \(G\), the operation is well-defined.

Proof of Theorem 6.1.1 Let \(G/H\) denote the set of cosets. We first show
that \(\circ\) is a well-defined operation. Suppose \(aH, bH \in G/H\) and

\[aH \circ bH = (a \ast b)H.\]

We want to show that this operation is independent of the element we
choose to represent to cosets. Now by definition, any element in \(aH\) is
of the form \(a \ast h\) for some \(h \in H\). So instead of picking \(a\) to represent
the cosets, we may pick \(a \ast h\). We have to show that

\[a \ast h \ast b \ast h' \circ (b \ast h')H = (a \ast b)H.\]

Now, by definition of the map, \((a \ast h)H \circ (b \ast h')H = (a \ast h \ast b \ast h')H.

Note that we have

\[a \ast h \ast b \ast h' = a \ast b \ast (b^{-1} \ast h \ast b) \ast h'.\]

Since \(H\) is normal, \(b^{-1} \ast h \ast b \in H\), i.e., \(b^{-1} \ast h \ast b = h''\), say. Hence,

\[(a \ast h \ast b \ast h')H = (a \ast b \ast h'' \ast h')H = (a \ast b)H.\]

The identity is clearly \(H\). The inverse of a coset \(aH\) is \((a^{-1})H\) and
associativity follows from the associativity of \(\ast\). Hence, \((G/H, \circ)\) is a
group.

Example 6.1.2 Let \(G\) be an abelian group. Then any subgroups of \(G\)
are normal. Hence \(G/H\) is always a group. Furthermore it is abelian.
For example \(\mathbb{Z}\) is abelian and its subgroups are \(n\mathbb{Z}\). The quotient groups
\(\mathbb{Z}/n\mathbb{Z}\) are cyclic of order \(n\).

6.2 The First Isomorphism Theorem

Theorem 6.2.1 (First Isomorphism Theorem) Let \(\varphi : (G, \ast) \to
(G', \ast')\) be a homomorphism with kernel \(K\). Then \((G/K, \circ)\) is isomorphic
to \(\varphi(G) := (\text{Im} \, \varphi, \ast')\). Here, \(aK \circ bK = (a \ast b)K\).
Proof. Define $\psi : G/K \to \varphi(G)$ by

$$\psi(aK) = \varphi(a).$$

Now $\psi$ is a homomorphism since

$$\psi(aK \circ bK) = \psi((a \circ b)K) = \varphi(a \circ b) = \varphi(a) \circ' \varphi(b) = \psi(aK) \circ' \psi(bK).$$

We must show that the map is a bijection. Let $aK, bK \in G/K$. Suppose $\psi(aK) = \psi(bK)$. Then $\varphi(a) = \varphi(b)$, which implies that $\varphi((a \circ b^{-1})) = e'$, the identity of $G'$. Hence, $a \circ b^{-1} \in K$, and therefore, $a \in bK$, which implies that $aK = bK$. The map is clearly surjective since $\psi(aK) = \varphi(a)$ for any element $\varphi(a) \in \varphi(G)$. Hence the map is bijective. Therefore $G/K$ is isomorphic to $\varphi(G)$.

Example 6.2.1 Let us revisit the fact that all cyclic groups are isomorphic. Consider the homomorphism

$$f : (\mathbb{Z}, +) \to G$$

defined by $f(k) = g^k$, where $G = \langle g \rangle$, $|G| = m$. Note that the kernel of $f$ is $m\mathbb{Z}$ and the image is $G$. By the First Isomorphism Theorem,

$$\mathbb{Z}/m\mathbb{Z} \simeq G.$$ 

Since $G$ is arbitrary cyclic group of order $m$, we conclude that all cyclic groups of order $m$ are isomorphic to the quotient group $\mathbb{Z}/m\mathbb{Z}$ and they are therefore isomorphic to each other.

Example 6.2.2 Let $G = \text{Aff}(1, \mathbb{R})$. Let $H$ be the group of non-zero real numbers under multiplication. Define

$$\varphi : G \to H,$$

by

$$\varphi(f_{a,b}) = a,$$

where $f_{a,b}(x) = ax + b$. Note that the kernel of $\varphi$ is $N := \{f_{1,b} | b \in \mathbb{R}\}$. Also the map $\varphi$ is surjective. Hence,

$$G/N \simeq H.$$
6.3 The Euler’s $\varphi$ function revisited

Let $(H, \ast)$ and $(K, \circ)$ be two groups. Consider the set 

$$H \times K = \{(h, k) | h \in H, k \in K\}.$$

One can make this set into a group using the operation $\cdot$ on $H \times K$ defined by 

$$(h, k) \cdot (h', k') = (h \ast h', k \circ k').$$

We now recall that for a fixed $n$, let $[r]_n$ denote the set of integers which have remainder $r$ when divided by $n$. The set $\mathbb{Z}/n\mathbb{Z} = \{[r]_n\}$ is an additive group of order $n$. It is identified with the quotient group $\mathbb{Z}/n\mathbb{Z}$. (This is the cyclic group in Example 6.2.1.)

We now show that

**Theorem 6.3.1** If $m$ and $n$ are relatively prime, then 

$$\mathbb{Z}/mn\mathbb{Z} = \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}.$$ 

**Proof.** Define 

$$f : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z},$$

by 

$$f(a) = ([a]_m, [a]_n).$$

The kernel of this map is the set of elements $a$ satisfying $a \equiv 0 \pmod{m}$ and $a \equiv 0 \pmod{n}$, which means that $a \equiv 0 \pmod{mn}$ since $m$ and $n$ are relatively prime. Hence, the kernel is $mn\mathbb{Z}$. By the First Isomorphism Theorem, we find that 

$$\text{Im } f \simeq \mathbb{Z}/mn\mathbb{Z}.$$ 

Next, note that $\text{Im } f$ is contained in $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ and so there is map 

$$\psi : \mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$

given by the composition 

$$i_1 : \mathbb{Z}/mn\mathbb{Z} \to \text{Im } f$$

followed by 

$$i_2 : \text{Im } f \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}.$$
It is clear that

\[ |\mathbb{Z}/mn\mathbb{Z}| = |\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}|. \]

In order to show that \( \psi \) is bijective, it suffices to show that \( \psi \) is injective. But this is true since \( \psi \) is the compositions of two injections \( i_1 \) and \( i_2 \).

As a corollary, we have the Chinese Remainder Theorem:

**Corollary 6.3.2** Let \( m \) and \( n \) be relatively prime and let \( a \) and \( b \) be any two integers. Then there exist an integer \( x \) such that

\[ x \equiv a \pmod{m} \quad \text{and} \quad x \equiv b \pmod{n}. \]

We have seen that the set

\[ (\mathbb{Z}/n\mathbb{Z})^* = \{[r]|\gcd(r, n) = 1\} \]

is a group under multiplication. Let this group be denoted by \( U_n \). Note that \( |U_n| = \varphi(n) \).

By the previous Theorem we find that

\[ f : \mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \]

is an isomorphism. We now consider \( f|_{U_{mn}} \) be the restriction of \( f \) on \( U_{mn} \). Let \([a]_{mn} \in U_{mn}\). Then

\[ f([a]_{mn}) = ([a]_m, [a]_n). \]

Since \([a]_{mn} \in U_{mn}\), there exist a \( b \) such that \( ab \equiv 1 \pmod{mn} \). This means that there exist a \( b \) such that \( ab \equiv 1 \pmod{m} \) and \( ab \equiv 1 \pmod{n} \) and so, \([a]_m \in U_m \) and \([a]_n \in U_n \). Hence \( f \) is a map from \( U_{mn} \) to \( U_m \times U_n \). Note that if \([a]_m = [1]_m \) and \([a]_n = [1]_n \) then \([a]_{mn} = [1]_{mn} \) and hence, the kernel of \( f \) is \([1]_{mn}\). Let \([a]_m \in U_m \) and \([b]_n \in U_n \). Since \( f \) is surjective on \( \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \), there is \([c]_{mn} \in \mathbb{Z}/mn\mathbb{Z} \) such that \( f([c]_{mn}) = ([a]_m, [b]_n) \). We need to show that \([c]_{mn} \in U_{mn} \) and this will imply that \( f|_{U_{mn}} \) is an isomorphism. Now, \( f([c]_{mn}) \) is sent to \([c]_m, [c]_n \) and since \([c]_m = [a]_m \), \( c \) is relatively prime to \( m \) by choice of \( a \). Similarly \( c \) is relatively prime to \( n \). But this means that \( c \) is relatively prime to \( mn \) (recall that \( \gcd(c, n) = \gcd(c, m) = 1 \) implies that \( \gcd(c, mn) = 1 \)). Hence \([c]_{mn} \in U_{mn} \). Hence \( f|_{U_{mn}} \) is an isomorphism from \( U_{mn} \) to \( U_m \times U_n \). By counting the number of elements, we conclude that

\[ \varphi(mn) = \varphi(m)\varphi(n), \quad (6.3.1) \]
if \( \gcd(m, n) = 1 \).

To continue, note that if \( m = p^\alpha \) then the numbers from 1 to \( p^\alpha \) that are not prime to \( m \) are \( p, 2p, \ldots, p^{\alpha-1}p \). Hence there are \( p^{\alpha-1} \) numbers which are not coprime to \( m \). Hence, the total numbers less than \( p^\alpha \) that are relatively prime to \( m \) is \( p^\alpha - p^{\alpha-1} \), or

\[
\varphi(p^\alpha) = p^\alpha - p^{\alpha-1}.
\]  

(6.3.2)

Now, any integer \( n \) greater than 1 can be written as

\[
n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k}.
\]

By (6.3.1) and the fact that \( p_i^{\alpha_i} \) and \( p_j^{\alpha_j} \) are relatively prime, we find that

\[
\varphi(n) = \varphi(p_1^{\alpha_1}) \varphi(p_2^{\alpha_2}) \cdots \varphi(p_k^{\alpha_k})
= (p_1^{\alpha_1} - p_1^{\alpha_1-1})(p_2^{\alpha_2} - p_2^{\alpha_2-1}) \cdots (p_k^{\alpha_k} - p_k^{\alpha_k-1})
\]

by (6.3.2)

\[
= n \prod_{p|n} \left(1 - \frac{1}{p}\right).
\]

### 6.4 Semi direct product

If \((H, \circ)\) and \((N, \bullet)\) are groups, we can obtain the direct product of these groups by using the natural operation

\[
(h, n) \circ (h', n') = (h \circ h', n \bullet n').
\]

In this section, we learn another way of obtaining new groups from \( H \) and \( N \).

Recall that an isomorphism of a group from \( N \) to \( N \) is called an automorphism of \( N \). The set of automorphisms of \( N \) formed a group, denoted by \( \text{Aut}(N) \), called the automorphism group of \( N \).

Suppose now that there is a homomorphism \( \xi \) from \( H \) to \( \text{Aut}(N) \). This means that for \( h \in H \), \( \xi(h) \) is an automorphism of \( N \). Let \((H, \circ)\) and \((N, \bullet)\) be groups and we define

\[
(h, n) \circ (h', n') = (h \circ h', n \bullet \xi(h)(n')).
\]

Note that since \( \xi(h) \) is an automorphism \( \xi(h)(n') \in N \) and \( n \bullet \xi(h)(n') \in N \). So \( \circ \) is a binary operation on \( H \times N \).

Now, \( \circ \) is an associative operation. We check this as follow:
\[(h, n) \diamond ((h', n') \diamond (h'', n'')) = (h, n) \diamond (h' \circ h'', n' \cdot (\xi(h')(n'')))
= (h \circ h' \circ h'', n \cdot (\xi(h')n' \cdot (\xi(h')(n'')))
= (h \circ h' \circ h'', n \cdot (\xi(h)h'')(n''))
= (h \circ h' \circ h'', n \cdot (\xi(h)h')(n'')) = (h \circ h' \circ h'', n \cdot (\xi(h)h')(n'')).\]

and

\[((h, n) \diamond (h', n')) \diamond (h'', n'') = (h \circ h', n \cdot (\xi(h)(n'))) \diamond (h'', n'')
= (h \circ h' \circ h'', n \cdot (\xi(h)(n') \cdot (\xi(h'h'')(n''))).\]

We leave it as an exercise for the reader to check that the identity is 
\((1_N, 1_N)\) and the inverse of 
\((h, n)\) is 
\((h^{-1}, \xi(h^{-1})(n^{-1}))\).

**Example 6.4.1**

We can check that \(\text{Aff}(1, \mathbb{R})\) can be expressed as a semi-direct product. We note that we may view
\[f_{a,b} \circ f_{a',b'} = f_{aa', ab' + b}\]
as an operation on \(\mathbb{R}^+ \times \mathbb{R}\) with operation
\[(a, b) \circ (a', b') = (aa', b + ab').\]

Note that
\[\xi(a)(b) = ab.\]

**Example 6.4.2**

The smallest non-abelian group of odd order is a group of 21 elements. This can be constructed using semi-direct product. Let \(N = \langle b \rangle\) be a cyclic group of order 7 and \(H = \langle a \rangle\) be a cyclic group of order 3. Note that if \(G\) is a semi-direct product of \(N\) and \(H\), then \(N \cong \{1_H\} \times N = \tilde{N}\)
appears as a copy of a subgroup of \(G\) and \([G : \tilde{N}] = 3\). Set
\[\xi(a)(b) = aba^{-1} = b^2.\]

With this \(\xi\), we may construct a non-abelian group of order 21.
6.5 The Second Isomorphism Theorem

Let $H$ and $K$ be subgroups of a given group $G$. Define

$$HK = \{hk | h \in H, k \in K\}.$$  

In general $HK$ is not a group. For example, if we let $G = S_3$, $H = \langle (1\,2) \rangle$ and $K = \langle (1\,3) \rangle$, then

$$HK = \{(1), (1\,2), (1\,3), (1\,3\,2)\}.$$  

This is not a subgroup since $(1\,3\,2)$ has no inverse.

In the next theorem, we show that $HK$ is a group when either $H$ or $K$ is a normal subgroup of $G$.

**Theorem 6.5.1**

(a) If $H$ and $K$ are subgroups of $G$ and if one of them is a normal subgroup then $HK$ is a subgroup of $G$; moreover, $HK = KH$ in this case.

(b) If both $H$ and $K$ are normal subgroups then $HK$ is a normal subgroup of $G$.

**Proof.** Suppose $K \triangleleft G$. This means for any $g \in G$ and $k \in K$,

$$gkg^{-1} \in K,$$

or

$$gk = k'g \quad \text{for some } k' \in K.$$  

To prove (a), we only need to show that for any $hk, h_1k_1 \in HK$, $hk(h_1k_1)^{-1} \in HK$. Now, since $K$ is normal in $G$, there exist some $k'$ such that

$$hkk_1^{-1}h_1^{-1} = hh_1^{-1}k' \in HK.$$  

Hence $HK$ is a subgroup of $G$.

To prove the second part, observe that if $K$ is normal in $G$, then for some $k'$,

$$hk = k'h \in KH.$$  

Hence

$$HK \subseteq KH.$$  

Conversely, for any $h \in H$ and $k \in K$,

$$h^{-1}kh \in K$$
and hence, for some $k' \in K$,

$$kh = hk' \in HK.$$ 

To prove (b), we observe that for any $g \in G$, $h \in H$ and $k \in K$,

$$ghkg^{-1} = ghg^{-1}k' \quad \text{since } K \text{ is normal}$$

$$= gg^{-1}h'k' \quad \text{since } H \text{ is normal}$$

$$= h'k' \in HK.$$ 

Hence $HK$ is normal in $G$.

**Theorem 6.5.2 (Second Isomorphism Theorem)** If $H$ and $K$ are subgroups of a group $H \triangleleft G$, then $HK$ is a subgroup, $H \cap K \triangleleft K$, and

$$K/(H \cap K) \simeq HK/H.$$ 

**Proof.** Since $H$ is a normal subgroup of $G$, $H$ is normal subgroup $HK$ (a group since $H$ is normal). Hence $HK/H$ is a group. Define

$$\varphi : K \rightarrow HK/H$$

by

$$\varphi(k) = kH.$$ 

One checks directly that this is a homomorphism. To show that $\varphi$ is surjective, let $hkH \in HK/H$. Since $H \triangleleft G$, $k^{-1}hk \in H$, or $hk \in kH$. Hence we may write $hkH = kH$. The preimage of $hkH = kH$ under $\varphi$ is then given by $k$. This shows that $\varphi$ is surjective. Hence $\text{Im } \varphi = HK/H$.

The kernel of $\varphi$ is given by all $k \in K$ such that $kH = H$. In other words, $k \in H$. Hence $k \in H \cap K$. By the First Isomorphism Theorem, we deduce that

$$K/(H \cap K) \simeq HK/H.$$ 

The additive version of the second isomorphism theorem is

$$(H + K)/H \simeq K/K \cap H.$$ 

Let $m$ and $n$ be positive integers. Then one can show the following

$$m\mathbb{Z} + n\mathbb{Z} = (m,n)\mathbb{Z},$$

$$m\mathbb{Z} \cap n\mathbb{Z} = [m,n]\mathbb{Z},$$
and
\[ [a\mathbb{Z} : b\mathbb{Z}] = (b/a)\mathbb{Z}, \quad \text{if } a \mid b. \]

Let \( H = m\mathbb{Z} \) and \( K = n\mathbb{Z} \). Now,
\[ (H + K)/H = (m,n)\mathbb{Z}/m\mathbb{Z} \]
and
\[ K/(K \cap H) = n\mathbb{Z}/[m,n]\mathbb{Z}. \]

By the second isomorphism theorem, we have
\[ (m,n)\mathbb{Z}/m\mathbb{Z} \cong n\mathbb{Z}/[m,n]\mathbb{Z}. \]

Consider the order of the groups of both sides, we recover the identity
\[ [m,n](m,n) = mn. \]

The Second Isomorphism Theorem shows that
\[ |HK|/|H| = |K|/|H \cap K| \quad (6.5.1) \]
when \( H \) is a normal subgroup of \( G \). It turns out (6.5.1) is also true when \( H \) is not normal, that is, when \( HK \) is not a subgroup of \( G \). This result is contained in the next Theorem.

**Theorem 6.5.3 (Product Formula)** If \( H \) and \( K \) are subgroups of \( G \) then
\[ |HK|/|H| = |K|/|H \cap K| \]

**Proof.** Define a map
\[ \varphi : H \times K \to HK \]
by
\[ \varphi((h,k)) = hk. \]

Note that \( \varphi \) is surjective. We now determine the number of elements in \( H \times K \) that are mapped to a fixed element \( hk \). Define
\[ F(hk) = \{(hd,d^{-1}k) | d \in H \cap K \}. \]

Note that \( F(hk) \subseteq \varphi^{-1}(hk) \). On the other hand if
\[ \varphi((h',k')) = h'k' = hk, \]
6.6 The Third Isomorphism Theorem and the Correspondence Theorem

then

\[ h^{-1}h' = kk'^{-1} = d \]

and \( d \in H \cap K \). Hence

\[ \varphi^{-1}(hk) = F(hk). \]

Now, since

\[ H \times K = \bigcup_{hk \in HK} \varphi^{-1}(hk), \]

we conclude that

\[ |HK||H \cap K| = |H \times K| = |H||K|. \]

\[ \square \]

6.6 The Third Isomorphism Theorem and the Correspondence Theorem

**Theorem 6.6.1 (Third Isomorphism Theorem)** If \( H \) and \( K \) are normal subgroups of a group \( G \) and \( K \) is a subgroup of \( H \), then \( H/K \triangleleft G/K \) and

\[ (G/K)/(H/K) \cong G/H. \]

**Proof.** Define

\[ \varphi : G/K \to G/H \]

by

\[ \varphi(aK) = aH. \]

Note that \( \varphi \) is well defined. To show this, let \( a' \) be such that \( a'K = aK \), then \( a'a^{-1} \in K \subset H \). Hence \( \varphi(a'K) = a'H = aH \).

The kernel of \( \varphi \) contains \( aK \) such that \( \varphi(aK) = aH = H \). This means that \( a \in H \). Hence the kernel is \( H/K \). Since \( \varphi \) is surjective, by the First Isomorphism Theorem, we complete the proof of the Third Isomorphism Theorem.

\[ \square \]

The next Theorem is simple but useful.

**Theorem 6.6.2 (The Correspondence Theorem)** Let \( G \) be a group and \( K \triangleleft G \) be a normal subgroup. Let \( \text{Sub}(G;K) \) denote the family of all those subgroups \( S \) of \( G \) containing \( K \), and let \( \text{Sub}(G/K) \) denote the family of all the subgroups of \( G/K \).
(i) The function $\varphi : S \to S/K$ is a bijection from $\text{Sub}(G; K) \to \text{Sub}(G/K)$.

(ii) $T \triangleleft S$ in $\text{Sub}(G; K)$ if and only if $\overline{T} \triangleleft \overline{S}$ in $\text{Sub}(G/K)$, in which case $S/T \cong \overline{S}/\overline{T}$. Note that $[S : T] = [\overline{S} : \overline{T}]$.

(iii) Denoting $S/K$ by $\overline{S}$ we have $T \leq S \leq G$ in $\text{Sub}(G; K)$ if and only if $\overline{T} \leq \overline{S}$ in $\text{Sub}(G/K)$. Furthermore, $[S : T] = [\overline{S} : \overline{T}]$.

Proof. We check that if $S$ is a subgroup of $G$ containing $K$ then $\overline{S}$ is a subgroup of $G/K$. Take $sK, s'K \in G/K$. Then $ss'^{-1}K \in S$ since $ss'^{-1} \in S$.

We now prove (i). We first prove that $\varphi$ is injective. Suppose $\overline{S} = \overline{S}'$. This means that if $s \in S$ then $sK = s'K$ for some $s' \in S'$. This implies that $s \in S'$. So $S \subseteq S'$. Similarly $S' \subseteq S$ and hence $S = S'$.

Next, we show that $\varphi$ is surjective. The key is to show that every subgroup of $G/K$ is of the form $H/K$ for some subgroup $H$ containing $K$.

Let $H$ be a subgroup of $G/K$. Define $H = \{h| hK \in H\}$. Now, if $h, h' \in H$ then $hK, h'K \in H$ and so, $hh'^{-1}K \in H$ and so, $hh'^{-1} \in H$. Hence $H$ is a subgroup of $G$.

Note also that $K \triangleleft H$ since $K \in H$. Furthermore, $K \triangleleft H$ since it is normal in $G$.

By the definition of $H$ we see that $H/K = H$. (For if $hK \in H/K$ then $hK \in H$ by definition of $H$. Conversely if $hK \in H$ then $h \in H$ and hence $hK \in H/K$.)

Since every subgroup of $\text{Sub}(G/K)$ is of the form $H/K$ with $H$ containing $K$, it is clear that the preimage of $H/K$ under $\varphi$ is $H$.

Next, we prove (ii). Observe that if $T \triangleleft S$ then $\overline{T} \triangleleft \overline{S}$. The isomorphism follows from the Third Isomorphism Theorem. Suppose $\overline{T} \triangleleft \overline{S}$. Let $t \in T$ and $s \in S$. Then $sts^{-1}K(\overline{T}) \in \overline{T}$. This implies that $sts^{-1}K = t'K$ for some $t' \in T$. Therefore $T \triangleleft S$. From the isomorphism, we see that $[S : T] = [\overline{S} : \overline{T}]$.

Finally we prove (iii). If $T \leq S$ then clearly $\overline{T} \leq \overline{S}$. Conversely, if $\overline{T} \leq \overline{S}$ then $tK \in S/K$ and so $tK = sK$ for some $s \in S$. Hence $ts^{-1} \in K \subseteq S$ and hence $t \in S$.

The equality of the indices does not follow from third isomorphism since $S/T$ may not be a group. To show that equality of the group indices, we show that there is a bijection from $S/T$ to $\overline{S}/\overline{T}$. Define $\psi : S/T \to \overline{S}/\overline{T}$.
by

\[ \psi(sT) = sK\overline{T}. \]

If \( sK\overline{T} = s'K\overline{T} \), then \( ss'^{-1}K \in \overline{T} \). This implies that \( ss'^{-1} \in T \), or \( sT = s'T \). Hence \( \psi \) is injective. Take \( sK\overline{T} \in S/\overline{T} \). Then the pre-image is \( sT \). Therefore \( \psi \) is bijective.

We conclude this Chapter with an application of the Correspondence Theorem.

**Theorem 6.6.3** Let \( G \) be a finite abelian group.

(i) If \( p | |G| \) then \( G \) contains an element of order \( p \).

(ii) \( G \) has a subgroup of order \( d \) for every divisor \( d \) of \(|G|\).

**Proof.** We first prove (i) using induction on \( n = |G| \). The base step \( n = 2 \) is clear. Since \( G \) contains two elements and the non-trivial element is the element of order 2. Suppose \( |G| = n \). Choose \( a \in G \) to be an element of order \( k \). If \( p | k \) then \( k = pl \) and \( a^l \) has order \( p \) (See tutorial 4). If \( p \nmid k \), then consider the cyclic subgroup \( H = \langle a \rangle \). Now \( H \triangleleft G \) since \( G \) is abelian. The group \( G/H \) has order less than \( n \) and is divisible by \( p \). Hence there exists an element \( bH \in G/H \) such that \( bH \) has order \( p \) by induction. If the order of \( b \) is \( m \), then \( b^mH = H \) and \( p|m \). We are therefore back to the first case.

Next, we prove (ii) by induction on \(|G|\). If \(|G| = 2\), again this is true. Let \(|G| = n \) and let \( d | |G| \). Suppose \( p | d \). Then \( G \) contains an element of order \( p \) and this element generates a group \( H \) of order \( p \). This implies that \( G/H \) has order less than \( n \) and so \( G/H \) has a subgroup of order \( (d/p) \) since \( (d/p)|(n/p) \). By the correspondence theorem, there exist a subgroup \( S \) such that \( |S/H| = d/p \). Now, \(|S/H| = |S|/p = d/p \) implies that \(|S| = d \). Hence the result.
7
Group Action

7.1 A Theorem of Cayley

In this section, we show that every finite group of order $|G|$ can be identified as a subgroup of $S_{|G|}$.

**Theorem 7.1.1** Every finite group $G$ is isomorphic to a subgroup of $S_{|G|}$, where $S_{|G|}$ is the set of bijections from $G$ to $G$.

**Proof.**
For each $a \in G$, define $\tau_a : G \to G$ by

$$\tau_a(g) = ag$$

for all $g \in G$. Note $\tau$ is a bijection on $G$ but for $a \neq 1$, $\tau$ is not a homomorphism.

For $a, b \in G$,

$$\tau_a \tau_b(g) = abg = \tau_{ab}(g).$$

Hence,

$$\tau_a \tau_b = \tau_{ab}. \quad (7.1.1)$$

Define $\varphi : G \to S_{|G|}$ by

$$\varphi(a) = \tau_a.$$

By (7.1.1), $\varphi$ is a homomorphism from $G$ to $S_{|G|}$. By the first isomorphism theorem, we have

$$G/\text{Ker}\varphi \simeq \text{Im}\varphi.$$

But if $\varphi(a) = \tau_a = 1_{S_{|G|}}$ then $\tau_a(g) = ag = g$ for all $g \in G$. This implies that $a = 1$. Hence the Kernel of $\tau$ is $1_G$. As a result we see that

$$G \simeq \text{Im}\varphi \leq S_{|G|}.$$
Remarks. One can show that $S_G \simeq S_{|G|}$. Take $f : G \to \{1, 2, \cdots, n\}$ be a bijection from $G$ to $\{1, 2, \cdots\}$. Define $\alpha : S_G \to S_n$ by

$$\alpha(\tau) = f\tau f^{-1}.$$ 

This is an isomorphism.

The next result we are going to show is similar to the above.

**Theorem 7.1.2** Let $G$ be a group and let $H$ be a subgroup of $G$ having finite index $n$. Then there exists a homomorphism $\varphi : G \to S_{G/H}$ with $\ker \varphi \leq H$.

**Proof.** Denote the set of cosets by $G/H$. Define $\tau_a : G/H \to G/H$ by $\tau_a(gH) = agH$. Note that $\tau_a$ permutes elements in $G/H$ and hence may be identified with an element of $S_{G/H}$. As in the previous theorem, $\varphi(a) := \tau_a$ is a homomorphism from $G$ to $S_{G/H}$. Now, the kernel $\ker \varphi$ is given by $\varphi(a) = 1_{S_{G/H}}$. In other words $\tau_a(gH) = agH = gH$ for all $gH \in G/H$. In particular $aH = H$ and so, $a \in H$. Hence the kernel of $\varphi$ is contained in $H$.

Remarks. Note that $S_{G/H} \simeq S_{|G/H|}$.

We now use the above theorems to find all non-isomorphic groups of order 6.

**Theorem 7.1.3** If $G$ is a group of order 6, then $G$ is isomorphic to either $\mathbb{Z}/6\mathbb{Z}$ or $S_3$. Moreover, $\mathbb{Z}/6\mathbb{Z}$ and $S_3$ are non-isomorphic.

**Proof.** By Lagrange’s Theorem, the only possible orders of non-identity elements are 2, 3 and 6. If $G$ has an element of order 6 then $G \simeq \mathbb{Z}/6\mathbb{Z}$.

Now the order of $G$ is even. Hence, there is an element of order 2 in $G$ (see Tutorial 4). Let $T = \langle t \rangle$ where $t$ is one of the elements of order 2. Then $[G : T] = 3$ and there is a homomorphism $\rho$ from $G$ to $S_3$ by the above theorem with $\ker \rho \leq T$. Hence $\ker \rho = \{1\}$ or $\ker \rho = T$. If $\ker \rho = \{1\}$ then $G \simeq S_3$.

If $\ker \rho = T$, then $T \triangleleft G$ and $G/T$ is cyclic of order 3. Let $aT$ be the generator for $G/T$. Let the order of $a$ be $m$. Then $(aT)^m = T$ and so $3|m$. If $m = 6$ then $G = \langle a \rangle$ and $G$ is cyclic.

Now, assume that the order of $a$ is 3. We note first that $ata^{-1} \in T$ since $T \triangleleft G$. So, $ata^{-1} = t$ or $ata^{-1} = 1$. But $ata^{-1} = 1$ implies that $t = 1$, a contradiction. Hence, $ata^{-1} = t$ and $at = ta$. Hence $(at)^k = a^k t^k$. If $(at)^k = 1$ then $a^k = t^{-k}$. Then $a^{3k} = t^{-3k} = 1$. This implies that $t^{-3k} = 1$ and hence $2|(-3k)$, or $2|k$. Similarly, $3|k$. The
7.2 Group Action

**Definition 7.2.1** If $X$ is a set and $G$ is a group, then $G$ acts on $X$ if there exists a function $\alpha : G \times X \rightarrow X$, called an action, such that

1. For $g, h \in G$, then $\alpha_g \circ \alpha_h = \alpha_{gh}$.
2. $\alpha_1 = 1_{S_X}$ the identity function.

**Theorem 7.2.1** If $\alpha : G \times X \rightarrow X$ is an action of a group $G$ on a set $X$, then $g \mapsto \alpha_g$ defines a homomorphism $G \rightarrow S_X$. Conversely, if $B : G \rightarrow S_X$ is a homomorphism, then $\beta : G \times X \rightarrow X$ defined by $\beta(g, x) = B(g)(x)$ is an action.

**Proof.** We claim that $\alpha_g$ is a permutation on $X$. Note that $\alpha_g \alpha_g^{-1} = \alpha_{gg^{-1}} = \alpha_1 = 1_{S_X}$. Hence $\alpha_g$ has an inverse and $\alpha_g \in S_X$. Define $A : G \rightarrow S_X$ by $A(g) = \alpha_g$. Now $A$ is a homomorphism since

$$A(gh) = \alpha_{gh} = \alpha_g \alpha_h = A(g)A(h).$$

Conversely, given a homomorphism $B : G \rightarrow S_X$. Let $\alpha_g = B(g)$. We show that $\alpha$ is an action. Clearly, $\alpha : G \times X \rightarrow X$. For $g, h \in G$,

$$\alpha_{gh} = B(gh) = B(g)B(h) = \alpha_g \alpha_h.$$

Finally, $\alpha_1 = B(1) = 1_{S_X}$ since $B$ is a homomorphism.

Theorem 7.1.1 shows that $G$ acts on $G$ via left multiplication and Theorem 7.1.2 shows that $G$ acts on $G/H$ by left multiplication of $g$ to elements in $G/H$.

**Example 7.2.1** We now see one more group action. Define $\alpha_g(h) = ghg^{-1}$. Note that $\alpha_g$ is a bijection from $G$ to $G$. We easily check that $\alpha$ is an action from $G \times G \rightarrow G$.

**Example 7.2.2** Let $X = \{v_1, v_2, v_3\}$ be set of vertices of a triangle. Let $G = S_3$. Then the map $\alpha_g(v_i) = v_{g(i)}$ is an action from $G \times X \rightarrow X$.

7.3 Orbits and Stabilizers

From now on, we will write the action of $g$ on $x$ as $g \cdot x$ instead of $\alpha_g(x)$. (We will replace the symbol $\alpha$ by $\cdot$.)
7.3 Orbits and Stabilizers

Definition 7.3.1 If $G$ acts on a set $X$, then the orbit of $x$ ($x \in X$), denoted by $\mathcal{O}(x)$, is the subset of $X$ given by

$$\mathcal{O}(x) = \{g \cdot x\}.$$ 

The stabilizer of $x$, denoted by $G_x$, is the subgroup of $G$ given by

$$G_x = \{g \in G|g \cdot x = x\}.$$ 

Remarks. Let $g, g' \in G_x$. Let $1_G \in G_x$ and so $G_x$ is non-empty. Observe that $g^{-1}g \cdot x = 1_G \cdot x = x$. But $g \cdot x = x$ and so $g^{-1} \cdot x = x$. Hence $g'g^{-1} \cdot x = x$ and $G_x$ is a subgroup of $G$.

Example 7.3.1 Consider the action of $G$ on $G$ given by $g \cdot h = gh$. Then $\mathcal{O}(h) = G$. This is because the equation $gh = g'$ for any $g' \in G$ can be solved and hence for any $g' \in G$, $g'$ can be written as $g \cdot h = g'$ and hence $g' \in \mathcal{O}(h)$. A action that gives only a single orbit is called a transitive action.

Given $g \in G$, what is the stablizer of $g$? By definition $G_g = \{g' \in G|g'g = g\}$. So $G_g = \{1\}$.

Example 7.3.2 Let $G$ acts on $G/H$ by $g \cdot xH = gxH$. Again this action is transitive since the equation $gxH = x'H$ is solvable. Therefore there is only one orbit. Given $xH \in G/H$, $G_{xH} = \{g \in G|gxH = xH\}$. This implies that $x^{-1}gx \in H$, or $g = xhx^{-1}$ for some $h \in H$. Conversely if $g$ is a conjugate of $x$ then it is in $G_{xH}$. This implies that $G_{xH} = \{xhx^{-1}|h \in H\} = xHx^{-1}$.

Example 7.3.3 Let $G$ acts on $G$ by conjugation, that is, $g \cdot x = gxg^{-1}$. In general, this is not transitive. For example, when $G$ is an abelian group then $\mathcal{O}(x)$ contains only one element $x$. Now, $\mathcal{O}(x) = \{gxg^{-1}\}$ and this called the set of conjugates of $x$, denoted $x^G$. The stabilizer is given by $G_x = \{g \in G|gxg^{-1} = x\} = \{g \in G|gx = xg\}$. This is called the centralizer of $x$, denoted by $C_G(x)$.

Let $G$ acts on $X$ with more than one orbit. Suppose $\mathcal{O}(x)$ and $\mathcal{O}(x')$ are two orbits and $x' \notin \mathcal{O}(x)$. Note that if $g \cdot x = g' \cdot x'$ then $g^{-1}g \cdot x = x'$ and so $x' \in \mathcal{O}(x)$. Hence, $\mathcal{O}(x) \cap \mathcal{O}(x') = \emptyset$. This implies that $X$ is a disjoint union of orbits and we have

Theorem 7.3.1 If $G$ acts on $X$, then $X$ is a disjoint union of the orbits. If $X$ is finite, then $|X| = \sum |\mathcal{O}(x_i)|$, where one $x_i$ is chosen from each orbit.
7.4 A relation between orbit and stabilizer

We now prove an important relation between the stabilizer of \(x\) and the orbit containing \(x\).

**Theorem 7.4.1** If \(G\) acts on a set \(X\) and \(x \in X\), then

\[ |O(x)| = |G : G_x|. \]

*Proof.* Define a map \(\psi : G/G_x \to O(x)\) by

\[ \psi(gG_x) = g \cdot x. \]

We show that this is well defined. Pick \(gg'G_x\) where \(g' \in G_x\). We find that \(\psi(gg'G_x) = gg' \cdot x = g \cdot x\) since \(g' \cdot x = x\). Suppose \(g \cdot x = g' \cdot x\). Then \(g^{-1}g' = x\) and hence, \(gG_x = g'G_x\). Hence \(\psi\) is injective. Next, given \(g \cdot x \in O(x)\), we see that its pre-image is \(gG_x\) and hence, \(\psi\) is surjective. Therefore \(\psi\) is bijective and

\[ |G : G_x| = |O(x)|. \]

\[ \square \]

**Corollary 7.4.2** If a finite group \(G\) acts on a set \(X\), then the number of elements in any orbit is a divisor of \(|G|\).

*Proof.* This is true since \(|O(x)||G : G_x|\) and \(|G : G_x||G|\).

**Corollary 7.4.3** If \(x\) lies in a finite group \(G\), then the number of conjugates of \(x\), denoted by \(x^G\), is the index of its centralizer:

\[ |x^G| = |G : C_G(x)|, \]

and hence it is a divisor of \(|G|\).

*Proof.* This follows by specifying the action, the group \(G\) and the set \(X\). See Example 7.3.3.

7.5 The class equation of a group and Cauchy’s Theorem

**Lemma 7.5.1** Let \(G\) be a finite group. Let \(x \in G\). Then \(x^G\) contains exactly one element if and only if \(x \in Z(G)\), where \(Z(G)\) is the center of \(G\).
Proof. If \( x^G \) contains one element then since \( x \in x^G \), \( gxg^{-1} = x \) for all \( g \in G \). Hence \( gx = xg \) for all \( g \in G \) and \( x \in Z(G) \). Conversely if \( x \in Z(G) \), then \( x \) commutes with every \( g \in G \) and so \( gxg^{-1} = x \) for all \( g \in G \) and hence, \( x^G = \{x\} \).

**Lemma 7.5.2** Let \( G \) be a finite group. Then

\[
|G| = |Z(G)| + \sum_i |G : C_G(x_i)|,
\]

(7.5.1)

where one \( x_i \) is selected from each conjugacy class having more than one element.

**Proof.** From Theorem 7.3.1, with \( X = G \) and the action being the conjugation, we find that

\[
|G| = \sum_i |\mathcal{O}(x_i)|.
\]

By Theorem 7.4.1,

\[
|\mathcal{O}(x_i)| = |G : C_G(x_i)|.
\]

Hence we have

\[
|G| = \sum_{x_i} |G : C_G(x_i)|
= \sum_{|G : C_G(x_i)| = 1} |G : C_G(x_i)| + \sum_{|G : C_G(x_i)| > 1} |G : C_G(x_i)|
= |Z(G)| + \sum_{x_i \notin Z(G)} |G : C_G(x_i)|,
\]

by Lemma 7.5.1.

The equation (7.5.1) is called the class equation of a finite group \( G \).

**Theorem 7.5.3** If \( G \) is a finite group whose order is divisible by a prime \( p \), then \( G \) contains an element of order \( p \).

**Proof.** If \( |G| = 2 \) then clearly \( G \) has an element of order 2. Suppose the statement is true for all groups with order less than \( n \). Let \( G \) be a group of order \( n \). Let \( p \) be a prime that divides \( n \). First note that if \( x \in Z(G) \) then \( x^G = \{x\} \).

Let \( x \notin Z(G) \) be an element of \( G \). Then \( C_G(x) \) is a subgroup of \( G \). If \( p | C_G(x) \) for some element \( x \in G \), then by induction, \( C_G(x) \) has an element of order \( p \) and we are done.
If not, then since \( p|G : C_G(x)| : |C_G(x)| \), by Euclid’s lemma, \( p||G : C_G(x)| \) for all \( x \notin Z(G) \). Now from (7.5.1), we have
\[
|G| = \sum_{x_i \in G} |x_i^G| = \sum_{x_i \notin Z(G)} |G : C_G(x_i)| + |Z(G)|.
\]
Now, \( p||G : C_G(x)| \) for all \( x \notin Z(G) \) and \( p||G| \). The above equation then implies that \( p||Z(G)| \). But \( Z(G) \) is an abelian group and so, by Theorem 6.6.3, we deduce that \( Z(G) \) contains an element of order \( p \) and hence \( G \) contains an element of order \( p \).

We now give another proof of Cauchy’s Theorem.

**Example 7.5.1** Let \( G \) be a finite group and let \( p||G| \). Let
\[
X = \{(a_1, a_2, \ldots, a_p)| a_i \in G, a_1a_2 \cdots a_p = 1\}.
\]
Note that \( |X| = |G|^{p-1} \). Let \([i]_p \in Z/pZ\) and define
\[
[i]_p(a_1, a_2, \ldots, a_p) = (a_{i+1}, a_{i+2}, \ldots, a_p, a_1, \ldots, a_i).
\]
Note that \((a_{i+1}, a_{i+2}, \ldots, a_p, a_1, \ldots, a_i) \in X\) since
\[
a_{i+1}a_{i+2} \cdots a_p(a_1 \cdots a_i) = (a_1 \cdots a_i)^{-1}a_1a_2 \cdots a_p(a_1 \cdots a_i) = 1.
\]
Clearly,
\[
[i + j]_p(a_1, a_2, \ldots, a_p) = [i]_p([j]_p(a_1, a_2, \ldots, a_p)).
\]
Also, \([0]_p(a_1, a_2, \ldots, a_p) = (a_1, a_2, \ldots, a_p)\). Hence we have an action of \( Z/pZ \) on \( X \). If \( a_i = a \) for \( 1 \leq i \leq p \), then the orbit containing \( a := (a_1, a_2, \ldots, a_p) \) consists of one element. Suppose \( a_i = a \neq 1 \) for all \( i \) then this implies that there is an \( a \) such that \( a^p = 1_G \) and so, \( G \) contains an element of order \( p \).

Suppose not. That is, there is no \( a \) such that \( a_i = a \neq 1 \) for all \( i \). Then for each \( a, a_i \neq a_j \) for some \( i \) and \( j \). The orbit containing \( a \) must contain more than one element and since \( |O(a)| \) divides the order of \( Z/pZ \), it must be \( |O(a)| = p \). Hence, by Theorem 7.3.1, we have
\[
|G|^{p-1} = \{|(1, 1, 1, \ldots, 1)\} + kp \equiv 1 \pmod{p}.
\]
But \( p||G| \) and hence, \( 0 \equiv 1 \pmod{p} \), which is a contradiction. Therefore, there must be an element of order \( p \) in \( G \).

Let the order of \( G \) be \( n \). Suppose now that \( p \) is not a divisor of \( n \).
By Lagrange’s Theorem, $G$ has no element of order $p$. By the above computations, this means that
\[ n^{p-1} \equiv 1 \pmod{p} \]
and this is Fermat’s little Theorem.

**7.6 Counting with groups**

**Lemma 7.6.1**

(i) Let a group $G$ act on a set $X$. If $x \in X$ and $\sigma \in G$ then $G_{\sigma x} = \sigma G_x \sigma^{-1}$.

(ii) If a finite group $G$ acts on a finite set $X$ and if $x$ and $y$ lie in the same orbit, then $|G_y| = |G_x|$.

**Proof.**

(i). Let $g \in G_{\sigma x}$. Then

\[ g \sigma \cdot x = \sigma \cdot x. \]

Hence,

\[ \sigma^{-1} g \sigma \cdot x = x \]

and hence

\[ \sigma^{-1} g \sigma \in G_x, \quad \text{and} \quad G_{\sigma x} \subset \sigma G \sigma^{-1}. \]

The reverse inclusion is similar.

(ii). If $x$ and $y$ lies in the same orbit then there exist $\sigma \in G$ such that $\sigma \cdot x = y$. Therefore

\[ |G_y| = |G_{\sigma x}| = |\sigma G_x \sigma^{-1}| = |G_x| \]

since

\[ \sigma G_x \sigma^{-1} \cong G_x. \]

**Theorem 7.6.2** Let $G$ act on a finite set $X$. If $N$ is the number of orbits, then

\[ N = \frac{1}{|G|} \sum_{\sigma \in G} |\text{Fix} (\sigma)|, \]

where $\text{Fix} (\sigma)$ is the set of elements $x \in X$ fixed by $\sigma$. 
Proof. Let $X = \{x_1, x_2, \cdots, x_m\}$ and $G = \{\sigma_1, \sigma_2, \cdots, \sigma_n\}$. Let

$$\alpha_{i,j} = \begin{cases} 1 & \text{if } \sigma_i \text{ fixes } x_j, \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{i,j} = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{i,j}.$$ 

Translating this in an appropriate way, we see that

$$\sum_{i=1}^{n} \text{number of elements of } x \text{ fixed by } \sigma_i = \sum_{j=1}^{m} \text{number of } \sigma \in G \text{ fixing } x_j.$$ 

In other words,

$$\sum_{i=1}^{n} |\text{Fix}(\sigma_i)| = \sum_{j=1}^{m} |G_{x_j}|. \quad (7.6.1)$$

Let the number of orbits be $t$. This means that there exists $y_1, y_2, \cdots, y_t$ such that

$$X = \mathcal{O}(y_1) \cup \cdots \cup \mathcal{O}(y_t).$$

Therefore,

$$\sum_{j=1}^{m} |G_{x_j}| = \sum_{k=1}^{t} \sum_{x \in \mathcal{O}(y_k)} |G_x|.$$ 

By Lemma 7.6.1 (ii), we see that $|G_x| = |G_{y_k}|$ whenever $x \in \mathcal{O}(y_k)$. Hence,

$$\sum_{j=1}^{m} |G_{x_j}| = \sum_{k=1}^{t} |G_{y_k}||\mathcal{O}(y_k)|.$$ 

By Theorem 7.4.1, we have

$$|G_{y_k}||\mathcal{O}(y_k)| = |G|.$$ 

This shows that

$$\sum_{j=1}^{m} |G_{x_j}| = |G|t,$$

and together with (7.6.1), we conclude that $t$, the number of orbits, is

$$t = \frac{1}{|G|} \sum_{\sigma \in G} |\text{Fix}(\sigma)|.$$
Example 7.6.1
We now consider the following problem: In how many ways can we design a flag having three stripes with three colors, namely, red, white and yellow? Two examples of such flags are as follow:

![Red-White-Yellow Flag](image1)

To solve the above problem, we number the stripes as 1, 2 and 3 and denote white by $w$, yellow by $y$ and yellow by $y$. First, note that the following two flags are the “same” since we can flip one to the other.

![Red-White-Yellow Flag](image2)

In other words if $X$ is the set of triplets of the form $(a, b, c)$ where $a, b, c \in \{w, y, r, b\}$, we see that $(a, b, c)$ is identified with $(c, b, a)$. Now, let $G = (1\ 3)$. Then $G$ permutes the first stripe and the third stripe. We let $G$ acts on $X$. Then we see that the number of orbits is exactly the number of distinct designs of flags we are seeking. (For example, one of the orbit would contain $(y, w, r)$ and $(r, w, y)$ but this is counted as one design).

By Theorem 7.6.2, we see that this number is equal to

$$\frac{1}{2} \left( |\text{Fix}((1)(2)(3))| + |\text{Fix}((1\ 3)(1)(1))| \right).$$

The number of elements fixed by $(1)(2)(3)$ is $4^3$ since every $x \in X$ is fixed by the identity. The number of elements fixed by $(1\ 3)(2)$ is $4^2$. This is because to be fixed by $(1\ 3)$, both stripes 1 and 3 can only take the same color. There is no restriction for stripe 2 and hence the total $x \in X$ fixed by $(1\ 3)$ is $4^2$.

Hence the total number of flags is

$$\frac{1}{2}(3^3 + 3^2) = 18.$$

This is listed as follow:
Example 7.6.2 (Another proof of Fermat’s Little Theorem) Consider a \( p \)-gon where \( p \) is a prime. Suppose we want to color the vertices of the \( p \)-gon using \( n \) colors. How many ways can we do this?

Let \( X \) contains the \( p \)-tuples \( (a_1, a_2, \cdots, a_p) \) where \( a_i \) takes one of the \( n \) colors. Let \( G \) be the cyclic group generated by \( \sigma := (1, 2, \cdots, p) \). Let \( G \) acts on \( X \) by

\[
\sigma^j \cdot (a_1, a_2, \cdots, a_p) = (a_{\sigma^j(1)}, a_{\sigma^j2}, \cdots, a_{\sigma^jp}).
\]

Note that under this action, the orbit containing \( (a_1, a_2, \cdots, a_p) \) are

\[
\{(a_{i+1}, a_{i+2}, \cdots, a_p, a_1, \cdots, a_i), 0 \leq i \leq p - 1\}.
\]

This means that if a colored polygon can be obtained from another colored polygon via rotation, we only count the polygon once. As such, the number of orbits is the number of polygons with distinct colorings. Hence the number of orbits is

\[
\frac{1}{p} (|\text{Fix}((1)(2) \cdots (p))| + |\text{Fix}(\sigma)| + \cdots + |\text{Fix}(\sigma^{p-1})|).
\]

Each \( \sigma^i, i \neq 0 \), is a \( p \)-cycle and so the total number of \( x \) fixed by \( \sigma^i \)
is $n$. The number of $x$ fixed by $(1)(2)\cdots(p)$ is $n^p$. Since the number of orbits is an integer, we find that

$$n^p \equiv n \pmod{p},$$

which is Fermat’s Little Theorem.
8
Sylow’s Theorems

8.1 Introduction and Sylow’s Theorems Part 1

Lagrange’s Theorem states that every subgroups $H$ of a group $G$ has the property that $|H| | |G|$. The converse of Lagrange’s Theorem is true for finite abelian groups but not true in general for non-abelian groups. The non-abelian group $A_4$ is an example for which the converse of Lagrange’s Theorem fails.

On the other hand, Cauchy’s Theorem shows us that if we consider a subset of divisors of $|G|$, then a subgroup of order that is in the subset may still exist. Cauchy’s Theorem states that if $p$ is a prime number dividing $|G|$ then there is a subgroup of order $p$. Is there any other divisor $d$ of $|G|$ that would be the order of a subgroup of $G$?

In this chapter, we will show that if $p^\alpha | |G|$ but $p^{\alpha+1} \nmid |G|$, then there is a subgroup of order $p^\alpha$ in $G$. This is the main theorem of Sylow.

Definition 8.1.1 We use $p^\alpha \parallel m$ to denote $p^\alpha | m$ but $p^{\alpha+1} \nmid m$.

Theorem 8.1.1 (Sylow’s Theorem (Part 1)) If $p$ is a prime such that $p^\alpha || |G|$, then $G$ has a subgroup of order $p^\alpha$.

Proof

Let $|G| = p^\alpha m$, where $(p, m) = 1$. Let $\mathcal{S}$ be a set of subsets of $G$ (these subsets may not be groups) with exactly $p^\alpha$ elements, namely,

$$\mathcal{S} = \{ M \in G | |M| = p^\alpha \}.$$ 

Note that $\mathcal{S}$ has exactly $\binom{p^\alpha m}{p^\alpha}$ sets. Let $G$ acts on $\mathcal{S}$ via

$$g \cdot M = gM, M \in \mathcal{S}.$$
Note that
\[ |S| = \sum_{M \in S} \frac{|G|}{|\text{Stab}_G(M)|}. \]

Now check that \( p \nmid \left( \frac{p^\alpha m}{p^\alpha} \right) \) (Prove this) and hence there exists an orbit \( \mathcal{O}_Y \) with number of sets not divisible by \( p \). This implies that
\[ p \nmid \frac{|G|}{|\text{Stab}_G(Y)|}. \]

Therefore, \( |\text{Stab}_G(Y)| \) must be divisible by \( p^\alpha \) or the group \( P = \text{Stab}_G(Y) \) has order \( |P| \geq p^\alpha \).

Let \( a \in Y \). Since \( P \) stabilize \( Y \), \( aP \in Y \). Hence, \( |aP| = |P| \leq |Y| \).

But \( |Y| = p^\alpha \) and hence,
\[ |P| \leq p^\alpha. \]

Together with the previous inequality, we see that we constructed a group \( P \) of order exactly \( p^\alpha \).

\[ \square \]

Note that as a consequence of Theorem 8.1.1, we immediately obtain Cauchy’s Theorem. This is because the existence of the group \( P \) with order \( p^\alpha \) implies that \( Z(P) \) is non-trivial. We only need to apply Cauchy’s Theorem for abelian group since \( Z(P) \) is abelian to obtain an element of order \( p \) in \( G \).

**Definition 8.1.2** If \( p^\alpha ||G| \), and \( P \) is a subgroup of order \( p^\alpha \) then we call \( P \) a Sylow \( p \)-subgroup of \( G \).

### 8.2 Sylow’s Theorems Part 2 and 3

**Theorem 8.2.1 (Sylow’s Theorem (Part 2))** Let \( n_p \) be the number of conjugates of \( P \) in \( G \). Then
\[ n_p \equiv 1 \pmod{p}. \]

**Proof**

Let \( \mathcal{T} \) be the set of conjugates of a Sylow \( p \)-subgroup in \( G \). Write \( n_p = |\mathcal{T}| \). The group \( P \) acts on \( \mathcal{T} \) via
\[ s \cdot M = sMs^{-1}, M \in \mathcal{T}. \]
Suppose there is a one element $P$-orbit of $\mathcal{T}$, say $\{P_1\}$. If $s \in P$, then $sP_1s^{-1} = P_1$. This implies that

$$PP_1 = P_1P.$$ 

This implies that $P_1P$ is a subgroup of $G$. Since

$$|PP_1| = \frac{|P_1| \cdot |P_2|}{|P_1 \cap P_2|},$$

we see that $PP_1$ is a group with order divisible by $p$. Also

$$P \leq PP_1.$$ 

But the order of $P$ is maximal power of $p$ since it is a Sylow $p$-subgroup and hence,

$$PP_1 = P,$$

or

$$P_1 \subseteq P,$$

and hence $P_1 = P$ since the two groups have the same number of elements. Therefore, if there is an orbit with one element, the orbit must be $\{P\}$.

The rest of the orbits must contain more than one element. Since

$$|O_M| = \frac{|P|}{|\text{Stab}_P(M)|},$$

we conclude that the rest of the orbits has order dividing $p$, or

$$n_p = 1 + pL \equiv 1 \pmod{p}.$$

\[\square\]

**Theorem 8.2.2 (Sylow’s Theorem (Part 3))** Let $P$ be a Sylow $p$-subgroup. Every subgroup of $G$ with number of elements equal to a power of $p$ must be contained in some conjugate of $P$. In particular, any two Sylow $p$-subgroup are conjugates.

**Proof**

Let $P_2$ be a subgroup of $G$ with order a power of $p$. Let $|P_2| = p^r$. Let $P_2$ acts on $\mathcal{T}$, the set of conjugates of $P$ in $G$. If $P_2 \subseteq P'$ for some $P' \in \mathcal{T}$ then the proof is complete.
Suppose this is not the case and suppose that there is an orbit \(\{P_3\}\) with only one element. Then
\[
yP_3y^{-1} = P_3.
\]
for all \(y \in P_2\). This implies that
\[
P_2P_3 = P_3P_2,
\]
or
\[
P_2P_3 \leq G.
\]
But \(|P_2P_3|\) is a power of \(p\) and we must have
\[
P_2P_3 = P_3,
\]
or
\[
P_2 \subset P_3
\]
and this contradicts our assumption that \(P_2\) is not contained in any conjugates of \(P\).

Therefore there is no one element orbit. Number of elements in each orbit must be a divisor of \(|P_2|\). But this also means that the size of each orbit is divisible by \(p\). In other words \(n_p \equiv 0 \pmod{p}\) and this contradicts Theorem 8.2.1.

We thus conclude that all Sylow \(p\)-subgroups are conjugate to each other in \(G\).

\[\square\]

A consequence of the above theorem is that when \(G\) acts on \(T\), then there is only one orbit. But we have seen that if \(G\) acts on a set \(T\), then the number of elements in an orbit divides \(|G|\). Here we have only one orbit and this implies that
\[
n_p || G |.
\]

### 8.3 Solvable groups

**Definition 8.3.1** A finite group is solvable (or soluble) if it has a series of the form
\[
1 = G_0 \triangleleft G_1 \triangleleft G_2 \triangleleft \cdots \triangleleft G_n = G
\]
such that
\[
G_i/G_{i-1} \text{ is abelian.}
\]
Theorem 8.3.1 All groups of order $p^m$ and $pq$ are solvable.

Proof Suppose this is true for $p^m$ with $m < k$. Let $|G| = p^k$. We know that $|Z(G)| > 1$. Now both $G/Z(G)$ and $Z(G)$ are solvable by induction. We must have

$$G_0/Z(G) < \cdots < G_{i-1}/Z(G) < G/Z(G).$$

By the correspondence theorem, we have

$$Z(G) = G_0 < G_1 < \cdots < G_{i-1} < G,$$

such that

$$G_{i}/G_{i-1} \cong (G_{i}/Z(G))/(G_{i-1}/Z(G))$$

are abelian (by the third isomorphism theorem). We can complete the series for $G$ by using the series for $Z(G)$ on the left hand side of the series.

Hence $G$ is solvable.

Next, let $p > q$. Then $n_p \equiv 1 \pmod{p}$ and $n_p | pq$. This implies that $n_p | q$ or $n_p = 1$ or $q$. But $n_p = 1 + \ell p$ for some $\ell$ and $p > q$. Therefore $n_p \neq q$ since $p > q$. Hence $n_p = 1$. This implies that $P < G$ and $|G/P| = q$ and so $G$ is solvable.