Group Theory

Jacob Denson

Contents

CHAPTER 1

Introduction: What is a group?

In 1761, Leonard Euler proved the following theorem. Take the totient function $\phi : \mathbf{N} \to \mathbf{N}$, which maps a positive integer n to the number of positive integers less than or equal to n which are relatively prime to n [1]. A few examples are below, with the set of relatively prime integers included.

$$\varphi(1) = 1 \qquad \{1\}$$
$$\varphi(10) = 4 \qquad \{1,3,7,9\}$$

Euler showed that, for any two relatively prime integers a and b,

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

The idea of Euler's proof involves mapping the set of numbers relatively prime to b to integer multiples of a, showing that, modulo b, the arithmetical operations on the two sets behave the same way. The uncovered symmetry between the two sets then reveals the theorem as a byproduct.

Another problem which we have dealt with since high school is finding roots of polynomials; that is, finding real numbers x such that, for some specific specific sequence of other real numbers $(a_n, a_{n-1}, \ldots, a_1)$, the equation below holds.

$$a_n x^n + a_{n-1} x^{n-1} + \ldots a_1 x + a_0 = 0$$

Of course, for a polynomial of the form

$$a_2 x^2 + a_1 x + a_0$$

there is a simple strategy of finding a root. Simply plug the values of a_2, a_1, and a_0 into the formula

$$x = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2 a_0}}{2a_2}$$

and, if there will be exactly two solutions, or no solutions at all, for the polynomial. A natural question is whether there is a general 'quartic' formula to solve polynomials of the form

$$a_3 x^3 + a_2 x^2 + a_1 x + a_0$$

[1] A number is relatively prime to another number if the greatest common denominator of the two numbers is 1.

or even formulas for polynomials of higher degree. The answer to the first question is yes, but the formula is far too complicated to be memorized. The answer to the second is no. Everiste Galois in the 1800s explored these polynomial equations to determine the impossibility of a general formula to a polynomial of the form

$$x^4 + ax^3 + bx^2 + cx + d$$

via a consideration of symmetries between polynomial roots.

How do these problems connect? One way to see it is that these problems were solved via considering the interconnections between objects in a specific collection. The generalized consideration of this technique is known today as group theory. In this report we introduce the theory of groups, focusing mainly on the finite case. We assume the reader has at least had some experience with algebra before, especially of the linear variety, as well as general mathematical background that is covered in most introduction math subjects. The main reason these subjects are needed is because they are a guarentee that the reader possesses the mathematical rigour to appreciate the subject.

CHAPTER 2

Operations and Groups

Groups are a way to generalize the notion of a set with a well behaved operation. This operation mechanizes the ability to combine objects within that set. Examples include numbers, whose operations are inherent, and geometric symmetries of a shape. Since groups were defined in 1882 by Walter Dyck, the algebraic theory has arisen to become one of the largest fields in modern mathematics: abstract algebra. The power obtained from the simple definition of a group is that any proof about an abstract group instantly applies to the large number of concrete objects that fit the mold of a group's definition. Both problems in the prelude can be answered in the framework of group theory, and one can even say that group theory was created from a generalization of these specific problems. A mathematical trick to solving problems is to take many practical examples and attempt to abstract a minimal structure followed by all of them. It is then only natural then that Group theory resulted from a generalization of these problems.

Groups are intimately connected to the operations that define them. Thus, before we define a group, we must rigorously define what an operation is. Operations like addition and subtraction are mathematical structures that we use in everyday life. Thus, abstracting the general properties of operations is quite difficult for us to think of, as we are blinded by the everyday experience of using the operations. The formal definition is simple.

Definition 1. A **law of composition** or **assignment** on a set S is a function from $S \times S$ to S.

Think of an operation as a way of combining two elements of S into a new element in S. In our definition we have inherently incorporated the property of **closure**; the composition of any objects from the set S lies inside the set S.

If a and b are arguments to an assignment, we avoid using conventional symbols such as f or g for the assignment; formulas such as $f(a,b)$ and $g(a,b)$ are too clunky compared to the more elegant and suggestive notation of (ab), $(a \circ b)$, or $(a+b)$. We put the symbol in the middle rather than at the front to maintain our intuititive view of operations. A good notation can make working with complex ideas much simpler.

Some further shorthand suffices to write down formulas. Consider a given finite sequence of elements $(x_1, x_2, \ldots, x_n)$. Then another notation convenience arises, the aptly named 'Pi' notation, recursively defined as

$$\prod_{j=i}^{n} x_j = \left(\prod_{j=1}^{n-1} x_j\right) x_i$$

$$\prod_{j=i}^{i} x_j = x_j$$

The similarity to Σ notation used in arithmetical sums is intensional.

The problem with the above definition of assignment in our future studies is that it is too general. Mathematically, it is normally true that the more general a theory is, the less we can say about it. The set of all functions is too general to be studied in totality. Even when we narrow down functions to assignments, our study is too general. Thus we must specify subclasses of assignments which may prove more interesting. We take properties of common operations as inspiration.

The first problem with our definition of assignment is that it only considers combinations of pairs. Thus given a set of three elements abc, it is ambiguous whether to combine b and c first, then a, or to first form a and b together. For an arbitrary operation, this choice is integral as each choice may result in a different outcome. We refine the operations we focus on so this is not so.

Definition 2. An assignment on a set S is **associative** if, for any three elements a, b, and c in the set S,

$$a(bc) = (ab)c$$

so that we may formulate them in any order we choose.

Our first algebraic structure is a preliminary algebraic structure to a group. The aim of its introduction is to make statements of further theorems more elegant. Think of it as a warmup structure to get used to the objects we will study later on.

Definition 3. A **semigroup** is a set possessing an associative operation.

Ultimately, the property of associativity means brackets in an equation are irrelevant. We prove this rigorously, and then avoid using brackets for the rest of the report, unless for reasons of emphasis.

Theorem 2.1. *Let there be given a semigroup S, and a finite sequence $(x_1, x_2, \ldots, x_n)$ of elements in S. Then, for any positive integers l and m such that $l + m = n$, it can*

be shown that

$$\Big(\prod_{k=1}^{l} x_k\Big)\Big(\prod_{k=1}^{m} x_{l+k}\Big) = \prod_{k=1}^{n} x_k$$

Proof. We prove by induction on n, the number of elements in the sequence $(x_1,\ldots,x_n)$. When $n = 1$, the statement is vacously true; there are no positive integers l and m such that $l + m = 1$. When $n = 2$, we have two elements x_1 and x_2. It of course follows that $l = 1$, $m = 1$, and

$$\Big(\prod_{k=1}^{1} x_k\Big)\Big(\prod_{k=1}^{1} x_{k+1}\Big) = x_1x_2 = \prod_{k=1}^{2} x_k$$

Now suppose for our inductive argument that for any sequence of $n-1$ elements, the statement to be proved holds, where $n - 1 \geqslant 2$. Consider a sequence of n elements $(x_1, x_2, \ldots, x_n)$ and two positive numbers l and m such that $l + m = n$. The following calculation below proves the statement needed, where $m \geqslant 2$. The case where $m = 1$ follows the same strategy, but is left for the reader to fill in.

$$\Big(\prod_{k=1}^{l} x_k\Big)\Big(\prod_{k=1}^{m} x_{l+k}\Big) = \Big(\prod_{k=1}^{l} x_k\Big)\Big(\Big(\prod_{k=1}^{m-1} x_{l+k}\Big)\, x_m\Big) \tag{1}$$

$$= \Big(\prod_{k=1}^{l} x_k \prod_{k=1}^{m-1} x_{l+k}\Big)\, x_m \tag{2}$$

$$= \Big(\prod_{k=1}^{n-1} x_k\Big)\, x_m \tag{3}$$

$$= \prod_{k=1}^{n} x_k \tag{4}$$

Here (1) follows by definition of Pi notation, (2) follows from the associativity law, (3) results from the inductive hypothesis, and (4) from the definition of the notation again. □

Additional symbology aids in the elegance of our writing. Given a positive integer n, we can describe the exponential of the assignment on an element a in the following way:

$$a^n = \underbrace{aa\ldots aa}_{\text{n times}}$$

A recursive definition, more mathematically elaborate, is that

$$a^1 = a$$

$$a^n = (a^{n-1}a)$$

The following two lemmas related to the exponential are obvious results of the exponential, though they will be used inherently in the sequel, and thus must be explicity stated.

LEMMA 2.2. *For any two integers n and m, and any element a in an associative assignment,*

$$a^{n+m} = a^n a^m$$

Proof. Let n be an arbitrary number. We prove by induction on the other number m. For $m = 1$,

$$a^{n+1} = a^n a = a^n a^1$$

This follows directly from the definition. Now suppose this is true for $m = r-1$. We will show it must hold for $m = r$.

$$a^{n+r} = a^{n+r-1}a = (a^n a^{r-1})a = a^n(a^{r-1}a) = a^n a^r$$

It is important for our understanding to verify where each hypothesis and assumption was used in the calculation above. □

LEMMA 2.3. *For any two positive integers n and m, and for any element a,*

$$(a^n)^m = a^{nm}$$

Proof. As in the last proof, we use induction on m for a fixed number n. For $m = 1$,

$$(a^n)^1 = a^n = a^{n\cdot 1}$$

And by supposing that the lemma holds for $m = r - 1$

$$(a^n)^r = (a^n)^{r-1}a^n = a^{n(r-1)}a^n = a^{n(r-1)+n} = a^{nr}$$

Note that here lemma (2.2) is used implicitly. □

Another property of assignments is a very powerful characteristic.

DEFINITION 4. An assignment on a set is **commutative** if, for any elements a and b in the set,

$$ab = ba$$

thus allowing pairs of elements to permute between one another.

The power of commutativity is that, given an associative and commmutative operation, we can permute any elements in an equation. Let us rigorously prove this.

THEOREM 2.4. *For any finite sequence of elements $(x_1, x_2, \ldots, x_n)$ from a set upon which an associative and commutative assignment is defined, and for any permutation π on the numbers 1 to n,*

$$\prod_{k=1}^{n} x_k = \prod_{k=1}^{n} x_{\pi(k)}$$

Proof. We again prove by induction on the number of elements in the sequence. When the number of elements is one, the statement is obvious; the only permutation of one element is the identity permutation. Now suppose for induction that this is true for any permutation of $n-1$ elements. Let

$$(x_1, \ldots, x_n)$$

be a sequence of elements, and π a permutation of the numbers in the range 1 to n. Let m be the number such that $\pi(n) = m$. The following calculation shows we can move x_m to the end of the product.

$$\prod_{k=1}^{n} x_k = \Big(\prod_{k=1}^{m-1} x_k\Big)\Big(x_m \prod_{k=m+1}^{n} x_k\Big) \tag{5}$$

$$= \Big(\Big(\prod_{k=1}^{m-1} x_k\Big)\Big(\prod_{k=m+1}^{n} x_k\Big)\Big)\, x_m \tag{6}$$

We transition from (2.5) to (2.6) by use of both the associativity and commutativity property of the assignment. Now define a new permutation φ on the numbers between 1 and n by the piecewise formula

$$\varphi(x) = \begin{cases} x & x < m \\ x-1 & x > m \\ n & x = m \end{cases}$$

What follows is that, via a change of notation,

$$\Big(\prod_{k=1}^{m-1} x_k \prod_{k=m+1}^{n} x_k\Big)\, x_m = \Big(\prod_{k=1}^{n-1} x_{\varphi(k)}\Big)\, x_{\pi(n)}$$

As φ and π are permutations, so is $\pi \circ \varphi^{-1}$. If we restrict $\pi \circ \varphi^{-1}$ to only the numbers between 1 and $n-1$, we still have a permutation, because n is fixed in the permutation.

$$(\pi \circ \varphi^{-1})(n) = \pi(\varphi^{-1}(n)) = \pi(m) = n$$

Hence we can consider $\pi \circ \varphi^{-1}$ as a permutation of the numbers between 1 and $n-1$. By induction and the fact that $(\pi \circ \varphi^{-1})(n) = m$, it follows that

$$\begin{aligned}\Big(\prod_{k=1}^{n-1} x_{\varphi(k)}\Big)\, x_{\pi(n)} &= \Big(\prod_{k=1}^{n-1} x_{(\pi\circ\varphi^{-1}\circ\varphi)(k)}\Big)\, x_{\pi(n)} \\ &= \Big(\prod_{k=1}^{n-1} x_{\pi(k)}\Big)\, x_{\pi(n)} \\ &= \prod_{k=1}^{n} x_{\pi(k)}\end{aligned}$$

Hence we can reorder elements arbitrarily. □

Commutativity is a rare quality of the operations encountered in group theory. It will never characterize an assignment unless it is explicitly stated, or if we use $+$ as the symbol for our assignment. It will still occur however, and thus we must mention it. Via commutatitity, an elementary theorem about the exponential arises. The same strategy to the previous proofs suffices, hence we leave the proof to the reader.

Lemma 2.5. *If a and b are elements such that $ab = ba$, then*

$$(ab)^n = a^n b^n$$

We have specified all of the 'global' properties of operations we will study in this report. The more subtle 'local' properties have not been mentioned. In particular, one property of addition and multiplication has not yet been mentioned. With addition, there is a number 0 such that, for any number x,

$$x + 0 = 0 + x = 0$$

Multiplication has a similar number with these properties, 1. Both numbers are **idempotent**; that is, when we combine this number with any other number, the composition of the numbers stays the same. We generalize this concept to arbitrary operations in the following way.

Definition 5. An **identity** of a set S is an element that is idempotent with all other elements in S. We commonly call this element e, and its properties can be written by the statement that for all elements a,

$$ae = ea = a$$

Definition 6. A **monoid** is a set with an associative operation that contains an inverse. We say a monoid is **commutative** or **abelian** if the operation defined on the monoid is commutative in addition to being associative.

Some easy examples are the positive integers under addition, the negative integers under addition, and so on. Extensive examples are not provided as the importance of the monoid structure is only to aid in the statement of theorems which relate monoids to groups.

A monoid can only have one such identity element, as we show below.

LEMMA 2.6. *A monoid has a unique identity*

Proof. Suppose a monoid has two identites, denoted e and e'. Then it is true that, because e is an identity,

$$ee' = e$$

Because e' is an identity, we obtain that

$$ee' = e'$$

and by transitivity, we get the equality required. □

If $\cdot$ is used as an operation's symbol, we write e as 1. If $+$ is used, we write the identity as 0. Though not used as numbers, The symbols 1 and 0 then become metaphors to help us think about the identity with more abstract operations.

For a monoid, we define, for any element a

$$a^0 = e$$

extending the definition of the exponential. The properties previously proved for exponentiation still hold. This follows as 0 is idempotent in addition,

$$a^0 a^n = ea^n = a^n = a^{n+0}$$

and zeroes out multiplication,

$$(a^0)^n = e^n = e = a^0 = a^{0n} \qquad (a^n)^0 = e = a^0 = a^{0n}$$

A further quality of addition and multiplication involves the interconnection between elements of the set. Given a number a, there is a number b such that $a + b = 0$. We typically use the symbol $-a$ for b, and write the operation more concisely as $a - a$. With multiplication, every non-zero element a has an number b such that $a \cdot b = 1$. We denote b as a^{-1} or $1/a$, and write the operation as a/a.

DEFINITION 7. Given a monoid, we say an element a is **invertible** if there is another element b such that

$$ab = ba = e$$

the element b is normally denoted a^{-1} and called the **inverse** of a.

Intuitively, invertibility means that the action of operating with any element of the group is reversible. If $+$ is used for the operation, $-a$ is used for the inverse of the element a, and if $\cdot$ is used, $1/a$ might be used. This continues the

metaphor established for identity elements. For an arbitrary assignment, the multiplicative notation a^{-1} is preferred.

Here are some properties common to all inverses in a monoid. Because of the monoid structure, we assume associativity, but not commutivity in the invertible operation. Proofs are abbreviated; we assume the reader can now handle the abstration presented and hence can 'fill in the blanks' of the arguments.

LEMMA 2.7. *Let l and r and a be arbitrary elements of a monoid. If $la = e$ and $ar = e$, then $l = r$, and thus a is invertible.*

Proof. For then it follows that $l = le = lar = er = r$. □

LEMMA 2.8. *For any element a in a monoid, a^{-1} is unique.*

Proof. Lemma (2.6) shows any two inverses are the same, for if x and y are two inverses, substitute x for l and y for r in the statement above. □

We are now ready to state the fundamental description which will concern us for the rest of this report.

DEFINITION 8. A **group** is a monoid such that every element has an inverse.

The simplicity of a group's definition is deceiving. We are still discovering new things about groups over one hundred years after the theory's conception.

If $n < 0$, define $a^n = (a^{-1})^n$. Again, the previous exponential properties proved hold, but we can now extend exponentiation to any integer.

Our first group theoretic properties result fairly easily from the definition. They all follow from the fact that the inverse of every element is unique. Let a and b be arbitrary elements of a group:

LEMMA 2.9. *The inverse of an inverse of an element a is a. That is,*

$$(a^{-1})^{-1} = a$$

Proof. The proof follows because of the calculation

$$aa^{-1} = a^{-1}a = e$$

hence a is an inverse of a^{-1}, and is unique by Lemma (2.7). □

LEMMA 2.10. $(ab)^{-1} = b^{-1}a^{-1}$.

Proof. $(ab)(b^{-1}a^{-1}) = a(bb^{-1})a = aea^{-1} = aa^{-1} = e$. □

LEMMA 2.11. *Every equation $ax = b$ has a unique solution x.*

Proof. $x = ba^{-1}$ is a trivial solution, and is the only solution, as if x' is any other solution such that $ax = ax'$, then $a^{-1}ax = a^{-1}ax'$, which when evaluated gives us $x = x'$. □

A reader may attempt to generalize a group, weakening the claim that elements only require left inverses or right inverses. A left inverse for an element a is an element b where it is true $ba = e$, but where it is not necessarily true that $ab = e$. We show that the theory of this new structure is no different to that of a group.

THEOREM 2.12. *Any monoid where every element has a left inverse contains arbitrary inverses for an element, so that the monoid is a group.*

Proof. Let G be a monoid with the properties above, and suppose a is an arbitrary element of G. Then there is $b \in G$ such that $ba = e$. b also has a left inverse c such that $cb = e$. But then b has both a left and right inverse, so $a = c$ from Lemma (2.6). But that also means a has a right inverse when substituted for c, so b is the inverse of a. Since a was arbitrary, all elements are invertible, so G is shown to be a group. □

There are also many examples of groups. We list some interesting ones below.Do not attempt to understand all of these examples until they are mentioned later on in the book. There are far too many to think of. Use this page as a reference for the groups we mention in later chapters. Right now just have a look at a choice few to get a feeling for groups:

- The set of integers, rational, real, and complex numbers under addition form the groups $\mathbf{Z}^+$, $\mathbf{Q}^+$, $\mathbf{R}^+$, and $\mathbf{C}^+$.
- The same sets with zero removed under the operation of multiplication form the groups $\mathbf{Z}^\times$, $\mathbf{Q}^\times$, $\mathbf{R}^\times$, and $\mathbf{C}^\times$.
- The set of bijective functions on a set X under composition form the symmetric group $S_{|X|}$. $S_{|X|}$ has a cardinality of $|X|!$, as this is precisely the number of bijective functions on the set.
- For a vector space V, the set of automorphisms under compositions form the general linear group $GL(V)$. An equivalent definition, if the vector space is dimension n in a field $\mathbf{F}$, is the set of invertible n by n matrices with entries in $\mathbf{F}$, which we denote $GL_n(\mathbf{F})$.
- Let S be a set, and G a group with operation $\cdot$. Then the set of functions from S to G form a group with operation $\circ$ defined by $(f \circ g)(x) = f(x) \cdot g(x)$. If f is in the group, f^{-1} is the set defined by $f^{-1}(x) = f(x)^{-1}$
- Consider an n-sided regular polyhedron (where regular means each side is equal). A symmetry on an n sided regular polyhedron is a distance preserving mapping between the edges of the shape. From this fact, each and every symmetry on the polyhedron can be considered a rotational symmmetry or a reflection symmetry. We have n of each kind of these symmetries, so the group formed by taking compositions of symmetries to form more symmetries forms the Dihedral group D_n of order $2n$. The transformational approach to geometry attempts to understand shapes in this group theoretic way.

- The quaternion group Q is equal to the set $\{\pm 1, \pm i, \pm j, \pm k\}$. One can work out all operations between eleemnts from the following sequence of equations.

$$\begin{aligned} ij &= k & ji &= -k \\ jk &= i & kj &= -i \\ ki &= j & ik &= -k \end{aligned}$$

$$ii = jj = kk = -1$$

  In a way, we have concisely presented the entire group to you. Presentations are a concept we will look at later in group theory. The presented group here is commonly used to represent three dimensional space in computer graphics.
- The Klein-4 Group or Viergruppe is a group with elements $\{a, b, c, e\}$, where for every element k in the group, $k^2 = e$, and such that $xy = z$, for any permutation (x, y, z) of (a, b, c).

Now we have introduced the basics of group theory, we can attempt to delve into some tools developed in the last two centuries, through which we can understand various groups.

Exercise 1. *If G is a group such that, $x^2 = e$ for every element x, then G is abelian.*

Proof. Let a and b be arbitrary elements in G. Then $(ab)^2 = e$, hence $(ab)(ab) = e$, hence $(ab)^{-1} = ab$. But $(ab)(ba) = e$ also, hence we obtain by Lemma (2.8) that $ab = ba$. □

Exercise 2. *Let G be a finite abelian group written additively, with elements*

$$\{x_1, x_2, \ldots, x_n\}$$

such that, for all elements $x \neq e$, $2x \neq e$. What is the value of

$$\left(\sum_{k=1}^{n} x_k \right)$$

Proof. Consider the collection of sets A defined by

$$\{\{x_i, x_j\} \in G^2 : x_i + x_j = 0\}$$

No set in A is a singleton, as if $\{x\}$ is in A, we know that $2x = 0$. Furthermore, these sets partition G because inverses are unique (which shows in addition that there are an even number of elements in the group). Since A is finite, we may order the sets in A in the form

$$(x_{a_1}, x_{a_2}, x_{a_3}, x_{a_4}, \ldots, x_{a_{m-1}}, x_{a_m})$$

such that for any odd number k between 1 and m, $\{x_k, x_{k+1}\} \in A$. Now this sequence is just a permutation of the sequence which we sum above, hence by

Theorem (2.4),

$$\left(\sum_{k=1}^{n} x_k\right) = \left(\sum_{k=1}^{n} x_{a_k}\right)$$

and by Theorem (2.1),

$$\left(\sum_{k=1}^{n} x_{a_k}\right) = \left(\sum_{k=1}^{n/2-1} \left(x_{2k-1} + x_{2k}\right)\right) = \sum_{k=1}^{n/2-1} (0) = 0$$

Hence the sum above must be zero. □

Exercise 3 (Wilson's Theorem). *If p is a prime number, prove that*

$$(p-1)! \equiv -1 \mod p$$

Proof. As p is prime, all non-zero numbers modulo p form a group under multiplication. Suppose that, for some integer x,

$$x^2 \equiv 1 \mod p$$

then we know that

$$x^2 - 1 \equiv (x+1)(x-1) \equiv 0 \mod p$$

Hence either $(x+1) \equiv p$ or $x-1 \equiv p$. We conclude that $x-1 \equiv 0$, or $x+1 \equiv p$, hence $x \equiv 1$, or $x \equiv p-1$ modulo p. By using the technique of the exercise above, pairing elements that are not of order 2 (of which $p-1$ and 1 are the only exception), we conclude that

$$(p-1)! = \prod_{k=1}^{n} k \equiv (p-1) \equiv -1 \mod p$$

□

One can prove the converse of Wilson's theorem, but we leave this for the aspiring number theorist.

Exercise 4. *A Latin Square is an $n \times n$ array such that any row and column is a permutation of a fixed set of n elements*

$$\{x_1, x_2, \ldots, x_n\}$$

Given a finite group G of cardinality n, order elements of G by $(g_1, g_2, \ldots, g_n)$ and define an $n \times n$ array M by $M_{ij} = g_i g_j$. Prove that this is a latin square, and conversely, show that any latin square defines a finite group.

CHAPTER 3

Subgroups, Generators, Cosets, and Normality

We can understand a machine by the various components from which it is constructed. Likewise, we can understand a group by the various components that it contains. Most of this book will be attempting to define a group's components – they are not so evident as the gears of a car or the wheel that makes it turn as they are formed inherently in the process of defining a group. The first such component of a group is a subgroup.

Definition 9. A **subgroup** is a subset of a group that contains the identity, is closed under the operation which defines the group, and contains inverses for any element in the subgroup. Consisely, a subgroup is a set such that, if a and b are any elements in the subgroup, then so are the elements (ab), (a^{-1}), and (b^{-1}).

A subgroup is basically a group that is a subset of a bigger group. Examples of subgroups are below:

- Given the general linear group $GL_n(\mathbf{F})$, define the special linear group $SL_n(\mathbf{F})$ to be the set of matrices in the general linear group with determinant one. The subgroup property follows as the determinant operation has the multiplicative property, so that if
$$\det(X) = \det(Y) = 1$$
it follows that
$$\det(XY) = \det(X)\det(Y) = 1$$
as well as
$$\det(X^{-1}) = \det(X)^{-1} = 1$$
- Let M be a set, and N a subset. Then the set of bijective functions on M that leave elements in N fixed is a subgroup of $S_{|M|}$. In some sense, this set of functions is equivalent to $S_{|M|-|N|}$.
- Any group is a subgroup of itself, as is the set containing the identity. We call these trivial subgroups for self evident reasons.

Unexpectedly, we can verify subgroups based on a single statement, specified in the following lemma. We leave it to the reader to verify – the proof is just an unravelling of definitions.

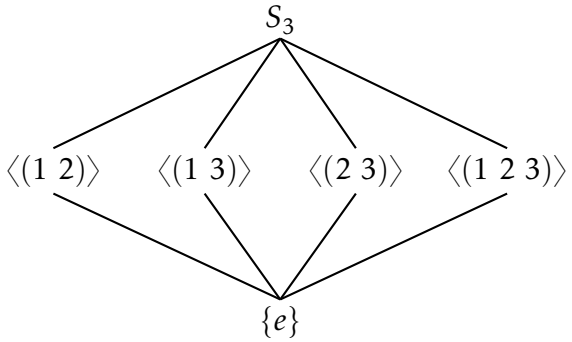
LEMMA 3.1. *A non-empty subset H of a group G is a subgroup if and only if, for any elements a and b in H, ab^{-1} is in H.*

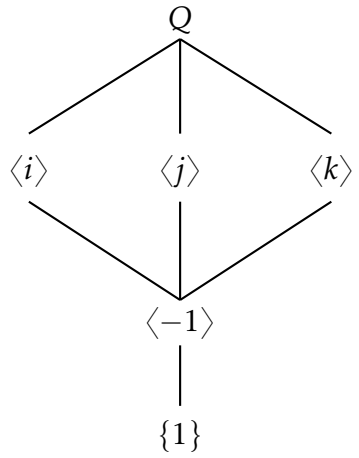
In our definition of a subgroup H of a group G we also assume that the identity is in a subgroup. However, it could be true that there is a different element e' such that e' acts idempotently on all elements in H, but not necessarily on all of G, and thus becomes a second identity! We show this cannot occur.

THEOREM 3.2. *Any group whose elements are a subset of another group must have equal identity elements.*

Proof. Let H be a group with identity e', and suppose H is a subset of G with identity e. Let h be an arbitrary element in H. Then, since H is a subset of G, $eh = h$, but also $e'h = h$, so $e'h = eh$. By multiplying both sides on the right by h^{-1}, we obtain that $e = e'$. □

An interesting view of subgroups results from the following mechanic. Given a group, take the set of all its subgroups ordered by the subset relation. Then the set of subgroups form a lattice : every two subgroups has a smallest subgroup that contains the two subgroups. This becomes very important in the context of Galois theory. We draw the lattices for the symmetric group S_3 and quaternion group Q below.





Here is an interesting discovery on collections of subgroups.

LEMMA 3.3. *Let G be a group, and $(H_j)_{j\in\mathcal{J}}$ a family of subgroups. Then it follows that*

$$\bigcap_{j\in\mathcal{J}} H_j$$

is also a subgroup of G.

Proof. If a and b are in $\bigcap_{j\in\mathcal{J}} H_j$, then they are in every group H_j, which means that (ab^{-1}) is in H_j for every H_j, hence (ab^{-1}) must be in the intersection of these groups, so the intersection is a subgroup. □

Now let G be a group, and S a subset of that group. Take the set $\mathcal{M}$ to be the set of all subgroups of G which contain S. Of course, $\mathcal{M}$ is non-empty, as G is a subgroup which contains S. Suppose we can index $\mathcal{M}$ completely by an index set $\mathcal{J}$. We make the following definition.

DEFINITION 10. Given a subset S of a group G, we define the subgroup generated by S to be the smallest subgroup that contains S, alternatively the intersection of all subgroups that contain S. We denote the subgroup $\langle S\rangle$. We call S the **generator** of the group $\langle S\rangle$. If S is a finite group $\{s_1, s_2, \ldots, s_n\}$, we also write $\langle S\rangle$ as $\langle s_1\ s_2\ \ldots\ s_n\rangle$.

Equivalently, the generated subgroup is the set of all elements of the form $x_1x_2\ldots x_n$ where either x_i or x_i^{-1} is in S. This is beacuse this forms a subgroup of G, and also every subgroup that contains S must contain these elements. In this way, generators work for groups analogously to how bases work in vector spaces.

A simple example is taken from linear algebra. One standard theorem proven is that every invertible matrix is the product of elementary matrices. This means that $GL_n(\mathbf{F})$ is generated by the set of all elementary n by n matrices in $\mathbf{F}$.

Definition 11. If a group is generated by a single element, then the group is called **cyclic** . Of course, this means every element is a power of that element.

One example is $\mathbf{Z}^+$, which is generated by both 1 and -1.

Let g be an element of a group G, and suppose that the cardinality of $\langle g \rangle$ is a non-negative integer c. Then the following properties hold for g:

Lemma 3.4. *$\{g, g^2, \ldots, g^c\}$ are all distinct elements of g.*

Proof. Suppose $g^i = g^j$, for $i \neq j$, and such that $0 \leqslant j < i < c$. Then $g^{i-j} = e$, for $i - j \neq 0$. Take any element g^m in $\langle g \rangle$. Then, by the euclidean division algorithm,

$$m = (i - j)q + r$$

for some integers q and r, where $0 < r < i - j$. Then

$$g^m = (g^{i-j})^q g^r = g^r$$

hence the size of $\langle g \rangle$, which we have denoted c, is less than or equal to $i - j$, for every element in the set is g^r for some r between 0 and $n-1$. But $i - j < c$, which leads us to our contradiction. Hence $g^i \neq g^j$ for numbers i and j in the range $0 < i < j < c$. □

Corollary 3.5. *For $0 < k < c$, $g^k \neq e$.*

Corollary 3.6. *If $\langle g \rangle$ is infinite, then $g^i \neq g^j$ if $i \neq j$.*

Proof. If $g^i = g^j$ for some $i > j$, then $g^{i-j} = e$, showing the cyclic group is at most order $i - j$. □

Corollary 3.7. *$g^c = e$.*

Proof. g^c cannot be equal to any element between g and g^{c-1}, so it must be the element of the group that is different from the other elements before it. Thus $g^c = e$, as no other element before g^c is e, and this is the only such element. □

Lemma 3.8. *$g^k = e$ if and only if $c \mid k$*

Proof. We leave this our argument to the reader. It is a simple application of euclidean division. □

Given an element g in an arbitrary group G, we define the order of g to be the cardinality of the group $\langle g \rangle$. Of course, if $\langle g \rangle$ is finite, this is exactly the least positive integer a such that $g^a = e$.

Lemma 3.9. *The order of an element (ab) is the same as the order of an element (ba).*

Proof. Consider the group $\langle ab \rangle$. We know that $(ba)^{-1} = a^{-1}b^{-1}$. Suppose the order of (ab) is finite, of order k. Then

$$(ab)^k = e$$

which means

$$b(ab)^k = b$$

and as $b(ab)^k = (ba)^k b$,

$$(ba)^k b = b$$

We conclude $(ba)^k = e$. Thus the order of (ba) is less than or equal to the order of (ab). This process can be done backwards to determine that the order of (ab) is less than or equal to the order of (ba), so the two must be equal. □

Now for any cyclic group $\langle g \rangle$, and for any integer a, one can verify $\langle g^a \rangle$ is a subgroup of $\langle g \rangle$. What is surprising is that any subgroup is of this form.

THEOREM 3.10. *G is a subgroup of a cyclic group $\langle g \rangle$ if and only if G is of the form $\langle g^a \rangle$ for some integer a. In short, the only subgroups of a cyclic group are cyclic.*

Proof. Let G be a subgroup of $\langle g \rangle$. If $G = \{e\}$, then $G = \langle g^0 \rangle$. In any other case, G has some non-zero element g^a. Thus G contains an element with positive exponent, as if a is negative, $-a$ is positive, and g^{-a} must be an element of the group by the closure property of a subgroup. By the well-ordering principle, G contains an element with smallest positive exponent g^b. Using euclidean division, every element $g^c \in G$ is of the form g^{mb+n}, where $0 < n < b$. Now $g^n \in G$, as $g^n = g^c g^{-mb}$, so we must conclude $n = 0$, as it cannot be a smaller positive exponent than b. Thus every exponent in G is divisible by b, and every number divisible by b is in G, so we conclude $G = \langle g^b \rangle$. □

We have built a complicated tower of definitions for us to comprehend without extensive use of examples. We will consider the additive integer group $\mathbf{Z}^+$ as a concrete example, the most basic cyclic group. Before we begin, some more notation is useful. For a group with an operation $\circ$ with two subsets S and M, define $S \circ M = \{s \circ m : s \in S, m \in M\}$. For a single element a, define $a \circ M = \{a\} \circ M$. The theorem above has some interesting repercussions in number theory:

- For any numbers $a, b \in \mathbf{Z}^+$, $a\mathbf{Z}^+ + b\mathbf{Z}^+$ is a group. so it is equal to some cyclic group $c\mathbf{Z}^+$ for an integer c. It turns out c is the greatest common denominator of a and b, denoted $\gcd(a, b)$.
- Given $a, b \in \mathbf{Z}^+$, $a\mathbf{Z}^+ \cap b\mathbf{Z}^+$ is a subgroup of $\mathbf{Z}^+$, so it too is $c\mathbf{Z}^+$, and c is the lowest common multiple of the two elements, denoted $\operatorname{lcm}(a, b)$.

THEOREM 3.11. *Consider a group G, with two elements g and h such that g is of order n and h is of order m. Then, if g and h commute (if $gh = hg$), and there order is relatively prime, then the order of (gh) is mn.*

Proof. Consider elements described above, and let the order of (gh) be p. $(gh)^{mn} = g^m h^n = e$, hence $p \mid mn$. We know that

$$(gh)^p = g^p h^p = e$$

hence, by multiplying both sides by n,

$$g^{mp}h^{mp} = g^{mp} = e$$

so that $n|mp$. As $\gcd(m,n) = 1$, $n|p$. □

We have another interesting number theoretic theorem before we finish our talk of cyclic groups. The proof is not trivial.

THEOREM 3.12. *For any prime p, $(\mathbf{Z}/p\mathbf{Z})^\times$ (consisting of all numbers that are invertible modulo p) is a cyclic group.*

Proof. We will use the fact that for any $r \geqslant 1$, $a^r \equiv 1 \mod p$ has no more than r solutions for a in $(\mathbf{Z}/p\mathbf{Z})^\times$. This follows that fact that the group is also a field, and thus its polynomials decompose into linear factors. Let n be the maximal order of elements in $(\mathbf{Z}/p\mathbf{Z})^\times$. Of course, we know $n \leqslant p-1$. We know by our initial claim that the equation $x^n = 1$ has at most n solutions. But for every $g \in (\mathbf{Z}/p\mathbf{Z})^\times$, $g^n = 1$, hence we have $p-1$ solutions. Thus $p-1 \leqslant n$, and by combining the two inequalities, $p-1 = n$. Thus there is an element g such that $\langle g \rangle$ has the same order as $(\mathbf{Z}/p\mathbf{Z})^\times$, so the two must be equal. □

Now we have specified the concept of a subgroup, we can attempt to gain a deeper understanding of the group via the subgroups it contains. We do this through the tool of cosets. We were previously introduced to subgroups as an attempt to understand the group as a sum of its parts. Cosets expand upon this method to understand the group properties.

DEFINITION 12. Let H be a subgroup of a group G. Define an equivalence relation $\sim$ by $x \sim y$ if $x \in yH$. The equivalence classes formed by the relation are denoted G/H and pronounced as 'G modulo (mod) H'. Each equivalence class is called a **left coset**.

Think of cosets are subgroups that are translated around by an element in a group.

LEMMA 3.13. *Every left coset is of the form gH for some element g that is in the equivalence class.*

Proof. Let C be an arbitrary equivalence class in G/H. Then C is non-empty; there is some element g in the class. We know $gH \subseteq C$, as for any element $h \in H$, $g \sim gh$. But also $C \subseteq gH$, as if $g \sim c$ for some $c \in C$, $c \in gH$. □

Right cosets are defined equivalently, by the equivalence relation $g \sim k$ if $g \in Hk$. Like left cosets, all right cosets can be written Hg for some g. Like left cosets, we define the set of right cosets by $H\backslash G$. Whether we use left cosets or right coses does not matter, theorems can be proved in equal power for each. The theorem below shows that there is actually a close connection between the two coset types.

LEMMA 3.14. *There is a one to one correspondence between left cosets and right cosets of any group.*

Proof. Let G be a group, and H a subgroup that generates G/H. Consider the mapping from left cosets to right cosets defined by $gH \mapsto Hg^{-1}$. We claim this mapping is a function. Suppose for two elements g and g' in G, $gH = g'H$. Then $gh = g'h'$ for some elements h and h' in H. But then, it follows that $(gh)^{-1} = (g'h')^{-1}$, which when evaluated gives us the equation $h^{-1}g^{-1} = h'^{-1}g'^{-1}$. We rearrange to get that $g^{-1} = hh'^{-1}g'^{-1}$. By the property of closure in a group, $hh'^{-1} \in H$, so that $g^{-1} = h''g'^{-1}$ for $h'' = hh'$. This means precisely that $g^{-1} \in Hg'^{-1}$, but also $g^{-1} \in Hg^{-1}$ (simply take $e \in H$). As cosets partition the group, we must conclude that $Hg^{-1} = Hg'^{-1}$. The two are equal, as was desired. In addition to this, the map is a bijection, with an inverse function defined by $Hg \mapsto g^{-1}H$. Thus we have a one-to-one correspondence, as was required. □

DEFINITION 13. The number of cosets (whether left or right) in G/H is denoted $(G : H)$, and is called the **index** of H in G.

We now come to one of the most important theorems in basic group theory, named after one of the pioneers of group theory, the french mathematician Joseph-Louis-Lagrange. It gives a useful characteristic of all subgroups of a finite group. Though the statement is formidable, the mechanics we have built up make the proof relatively simple – our definitions were the hard part to understand.

THEOREM 3.15 (Lagrange's Theorem). *The order of a subgroup of a finite group divides the order of the entire group.*

Proof. Let G be a finite group, and H a subgroup. Let g and g' be arbitrary elements of G. Define a function from elements of gH to elements of $g'H$ defined by the mapping $a \mapsto g'g^{-1}a$. This mapping is bijective, as it has an inverse function defined by the mapping $b \mapsto gg'^{-1}b$. Thus the order of one coset is equal to the other coset. We know that the cardinality of G is the sum of its partitions. That is, if G is partitioned into $\{g_1H, g_2H, \ldots, g_nH\}$, then

$$|G| = \sum_{k=1}^{n} |g_kH|$$

But we have proved that the order of any two of these cosets are equal, hence, for any coset gH

$$|G| = \sum_{k=1}^{n} |gH| = n|gH|$$

and as n is the index of the subgroup H, we obtain the following correspondence: for any coset gH, $|G| = |gH|(G : H)$. What this means that if any two of

the three elements is finite, so is the third, and the equation holds. By noting that H is a coset (simply take the coset of e), we obtain Lagrange's magnificant theorem as a corollary,

$$|G| = |H|(G : H)$$

hence $|H| \mid |G|$. □

Lagrange did not completely prove the theorem, showing it only for subgroups of the symmetric groups. The first complete theorem was published by Gauss in 1801. The following shows the power of Lagrange's theorem.

Corollary 3.16. *Any group of prime order is cyclic.*

Proof. Let G be a group of prime order. Take a non-zero element $g \in G$ (which is possible since $|G| > 1$), and consider $\langle g \rangle$. This is a subgroup, and thus the order of the group must divide G. But the only numbers that divide G are 1 and the order of G, as the number is prime, and $\langle g \rangle$ definitely contains more than one element. Thus the order of $\langle g \rangle$ is the same as the order of G, so $G = \langle g \rangle$. □

Corollary 3.17 (The Multiplicative Property). *Let G be a group, H a subgroup of G, and M a subgroup of H. Then $(G : M) = (G : H)(H : M)$.*

Proof. If M is a subgroup of H, Lagrange's theorem tells us that

$$|H| = |M|(H : M)$$

By further application of (1.4.2), It then follows that

$$|G| = |H|(G : H) = |M|(G : H)(H : M)$$

Noticing that M is also a subgroup of G,

$$|G| = |M|(G : M)$$

Thus we conclude

$$|M|(G : M) = |M|(G : H)(H : M)$$

By dividing by $|M|$ (which is non-zero as M is non-empty), we obtain the fact that $(G : M) = (G : H)(H : M)$. □

We now have the power to prove one of the theorems that introduced group theory at the beginning of the book. Euler used the methods of Lagrange to prove his totient function theorem. To show the power of group theory, we know prove his theorem. Let us be reminded of the theorem's statement.

Corollary 3.18 (Euler's Theorem). *For any two relatively prime integers a and b,*

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

where $\varphi(b)$ counts the number of integers less than b which are relatively prime.

Proof. Consider the group $\mathbf{N}^\times/b\mathbf{N}$, which consists of all invertible elements of $\mathbf{N}$ modulo b. We claim the size of this group is $\varphi(b)$, by showing that a necessary and sufficient property for inclusion in the group is being a relatively prime number of b less than b. Let x be an element of $\mathbf{N}^\times/b\mathbf{N}^\times$. Then there is some number y such that

$$xy \equiv 1 \pmod{b}$$

which means exactly that $xy + mb = 1$, for some integer m. Then it obviously follows that $\gcd(x, b) = 1$, hence x is relatively prime to b. Now suppose for some integer x, $\gcd(x, b) = 1$. Then there are two integers m and n such that

$$xm + bn = 1$$

so that $xm \equiv 1 \pmod{b}$, and thus x is an element of the group. Now let x be an arbitrary positive integer relatively prime to b. Then $\langle x \bmod b \rangle$ forms a subgroup of $\mathbf{N}^\times/\mathbf{N}$. Since this group is order $\varphi(b)$, we have by Lagrange's theorem that the order of $x \bmod b$ divides $\varphi(b)$. But then

$$(x \bmod b)^{\varphi(b)} \equiv 1 \pmod{b}$$

leading us to the final conclusion that

$$x^{\varphi(b)} \equiv 1 \pmod{b}$$

we obtain Euler's theorem simply from the theory of groups. □

Corollary 3.19 (Fermat's Little Theorem). *If p is a prime, and a is a number that does not divide p, Then*

$$a^{p-1} \equiv 1 \mod p$$

There are some very special subgroups that have interesting properties.

Theorem 3.20. *Let H be a subgroup of a group G. The following statements are equivalent, and if any hold, we say H is normal in G and write $H \triangleleft G$:*

(1) $gHg^{-1} \subseteq H$ for all g
(2) $gHg^{-1} = H$ for all g
(3) $gH = Hg$ for all g
(4) For all g, there is g' such that $gH = Hg'$

Proof. First we show (1) is equivalent to (2). Suppose $ghg^{-1} \subseteq H$ for all g. Then $gH \subseteq Hg$ (multiply both sides of the relation on the right by g. But also $g^{-1}Hg \subseteq H$, such that $Hg \subseteq gH$, so that $Hg = gH$. The reverse is obvious. We obtain (3) from (2) by multiplying both sides of the equation on the right by g, and the reverse by multiplying on the right by g^{-1}. The implication from (3) to (4) is obvious. From (4), note if $gH = Hg'$, $ge = g \in Hg'$, so that $Hg' = Hg$ as cosets are equal or disjoint. Thus all statements are shown to be equivalent. □

The trivial group is always normal in any group is lies in, because given any element g,

$$g^{-1}eg = g^{-1}g = e \in \{e\}$$

Furthermore, for any group G, $G \lhd G$. Thus no group posesses the characteristic that no subgroup is normal. We must take this into account in defining the property of sparcity of normal subgroups.

Definition 14. A group is **simple** if it contains no non-trivial normal subgroups, that is, if the only normal subgroups are $\{0\}$ and the group itself.

Some examples of normal subgroups are the following. Verification of normality is left as an exercise:

- If G is abelian, and H is a subgroup, $H \lhd G$.
- $SL_n(\mathbf{F}) \lhd GL_n(\mathbf{F})$
- If H is a subgroup of G of index two, $H \lhd G$
- If a group G is normal, and H is a cyclic subgroup, for any subgroup I in H, $I \lhd G$.
- Given a group G and a subset S, consider the subgroup

  $$N_G(S) = \{x \in G : xSx^{-1} = S\}$$

  This is the normalizer of S , and if S is a group, $N_G(S)$ is the largest group that S is normal in. That is, if K is a group containing S, and $S \lhd K$, then $K \subset N_G(S)$. We leave this as an exercise.
- Given the last normal group, consider a group G and subset S, and a resultant subgroup

  $$C_G(S) = \{x \in G : (\forall s \in S)(xsx^{-1} = s)\}$$

  called the centralizer subgroup of S , which is obviously normal in G.

The following propositions are an easy test of knowledge about normality.

Theorem 3.21. *If K is a subgroup of $N_G(H)$, then KH is a group, and $H \lhd KH$.*

Proof. First we prove KH is a subgroup. If k_1h_1 and k_2h_2 are in KH, then $k_1h_1(k_2h_2)^{-1}$ is in KH by the following calculation, which shows that KH is a subgroup by Lemma (3.1):

$$\begin{aligned} k_1h_1(k_2h_2)^{-1} &= k_1h_1h_2^{-1}k_2^{-1} \\ &= k_1(k_2^{-1}k_2)h_1h_2^{-1}k_2^{-1} \\ &= (k_1k_2^{-1})\left[k_2(h_1h_2^{-1})k_2^{-1}\right] \end{aligned}$$

As $k_2 \in K$, $k_2 \in N_G(H)$, hence the value enclosed in square brackets above is an element of H. $k_1k_2^{-1}$ is in K as K is a subgroup, hence the entire equation is in KH. Thus we obtain that KH is a group. Now Consider an arbitrary element

$h \in H$, and the equation $h^{-1}kh'h$ for some other arbitrary elements $k \in K$ and $h' \in H$. Using the same tricks as above,

$$h^{-1}kh'h = k\left[k^{-1}h^{-1}kh'h\right]$$

and the square brackets are contained in H. Thus $H \lhd KH$ □

We leave the proof of the next proposition to the reader.

THEOREM 3.22. *Let G be a group, and H and K subgroups of G such that $K \subset N_G(H)$. Then $H \cap K \lhd H$.*

CHAPTER 4

Homomorphisms and Isomorphism Theorems

Another way we attempt to dissect objects that we don't understand is by connecting metaphors to those objects that we do understand. We can formalize this in a group with the concept of a homomorphism.

Definition 15. Let G be a group with operation $\circ$ and H a group with operation $\cdot$. A **homomorphism** between G and H is a function f such that for any elements x and y. $f(x \circ y) = f(x) \cdot f(y)$. We say that G and H are homomorphic. If a homomorphism is injective, we say that G can be embedded in H . If $G = H$, we call a homomorphism an endomorphism .

What a homomorphism means intuitively is that information about the group G can be implanted into a subgroup of H. Some elements may become one element, but the information is still there. The following outlines the specific elements that are merged into one by a homomorphic transformation.

Definition 16. The **kernel** of a homomorphism φ, denoted $\ker(\varphi)$ is the set of elements in the domain that are mapped to the identity element in the range.

The following properties hold for any homomorphism φ:

Lemma 4.1. $\varphi(e) = e$

Proof. $\varphi(e) = \varphi(ee) = \varphi(e)\varphi(e)$, hence $\varphi(e)$ is idempotent. □

Lemma 4.2. $\varphi(a^{-1}) = \varphi(a)^{-1}$.

Proof. $\varphi(a^{-1})\varphi(a) = \varphi(aa^{-1}) = \varphi(e) = e$. □

We use normal subgroups along with homomorphisms to connect groups. To tease this fact, we show the following theorem.

Lemma 4.3. *The kernel of a homomorphism is a normal subgroup of the domain of the homomorphism.*

Proof. Let G and H be groups, and φ a homomorphism between G and H. If $\varphi(j) = e$, $\varphi(gjg^{-1}) = \varphi(g)\varphi(j)\varphi(g)^{-1} = \varphi(g)\varphi(g)^{-1} = e$. Thus gjg^{-1} is in the kernel for any element g in G, and we have shown normality. □

Lemma 4.4. *The image of a homomorphism is a subgroup*

Proof. If a is a group element such that $\varphi(x) = a$ for some elements a and x, then $\varphi(x^{-1}) = a^{-1}$, and if b is an element such that $\varphi(y) = b$ for some element y, then $\varphi(x^{-1}y) = a^{-1}b$.by Lemma (3.1). □

LEMMA 4.5. *A homomorphism is injective if and only if $\varphi(a) = e$ implies $a = e$.*

Proof. We prove both implications. Suppose a homomorphism is injective, and if $\varphi(a) = e$. Then $a = e$ as $\varphi(e) = e$. Instead, to prove the converse, if $\varphi(a) = e$ implies $a = e$, then if $\varphi(a) = \varphi(b)$, then $\varphi(ab^{-1}) = e$, so that $ab^{-1} = e$ and thus $a = b$. □

Some examples of homomorphisms are the following:

- The determinant function from $GL_n(\mathbf{F}) \to \mathbf{F}^\times$
- The exponentiation map $x \mapsto e^x$
- For any element a in G, the map from $\mathbf{Z}^+$ defined by $x \mapsto a^x$.
- The absolute value map from $\mathbf{C}^\times$ to $\mathbf{R}^\times$

DEFINITION 17. An **isomorphism** is a bijective homomorphism. In this case the inverse of the homomorphism is also a homomorphism. A bijective endomorphism is also called an automorphism . If G is isomorphic to H, we write $G \cong H$.

An isomorphism states that all algebraic information about G holds in H – effectively, they are the same group with different names for the operations and elements of the group. An automorphism basically states that various objects in the same group behave in the same way when permuted according to the function.

Let us consider an automorphism on $\mathbf{C}^\times$. Take the map

$$a + bi \mapsto a - bi$$

that swaps every complex number with it's complex conjugate. When i was introduced to the real number system, what this automorphism means is we could have introduced $-i$ as the basic element and the complex number system would behave exactly the same.

For a group, the set of automorphisms on the group, taken with the operation of composition of functions, form a group . Given an element g in G, the set of automorphisms $h \mapsto ghg^{-1}$ defines the set of inner automorphisms, a subgroup of the set of automorphisms. The map that sends g to its inner automorphism is a homomorphism. The kernel of this homomorphism is the center group

$$Z(G) = \{g \in G : \forall h : gh = hg\}$$

it is obviously normal, which can be proved either by the fact it is a kernel of the mapping from G to $S_{|X|}$ by conjugation, or by direct analysis.

A useful theorem, though obvious, is very useful. We state it without proof.

THEOREM 4.6. *Let G be a group, and S a subset such that $G = \langle S \rangle$. Suppose $f : S \to H$ is a map to another group H. If there is a homomorphism from G to H whose restriction to S is f, then this is the only homomorphism with this property.*

The theorem is no different from the fact that two linear transformations which are equal when restricted to the basis elements of a vector space are equal in full.

We can finally use coset constructions to prove something meaningful. Let G be a group and H a normal subgroup. For two cosets M and N in G/H, define an operation on the cosets by $M \circ N = MN$. As $M = gH$ and $N = g'H$ for some $g, g' \in H$,

$$MN = gHg'H = gg'HH = gg'H$$

This follows by the normality of H. Thus the operation we have constructed is closed in G/H, and G/H forms another group: the factor or quotient group . H is the identity in this group. The map $g \mapsto gH$ is the canonical map or projection π from G to G/H, and is a surjective homomorphism. The kernel is the normal subgroup, hence every normal subgroup is the kernel of some homomorphism; some people take this as the primary definition of a normal subgroup. The projection of G onto G/H has a property that we prove in a more general form.

THEOREM 4.7 (The First Isomorphism Theorem). *Let φ be a homomorphism between two groups G and H, and let N be a normal subgroup of the kernel of φ. Then there is a homomorphism $\overline{\varphi}$ from G/H to H such that $\overline{\varphi} \circ \pi = \varphi$, where π is the canonical map. If N is the kernel, the map $\overline{\varphi}$ is an isomorphism to* $\mathrm{im}(\varphi)$.

Proof. For every $n \in N$, we have $\varphi(n) = e$ as N is a subgroup of the kernel. Thus if $gN = hN$ for $g, h \in G$, $\varphi(g) = \varphi(h)$. The map $\overline{\varphi} : gN \mapsto \varphi(g)$ then becomes well defined. It is a homomorphism as $gHhH = ghH$, so ghH is mapped to $\varphi(gh) = \varphi(g)\varphi(h)$. We then obtain that $\overline{\varphi} \circ \pi = \varphi$ by construction. Because π is surjective, the map is unique.

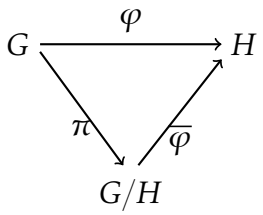
Now if N is the kernel, the homomorphism is injective. $\varphi(a) = \varphi(b)$ implies $\varphi(ab^{-1}) = e$. Then $ab^{-1} \in N$, and $ab^{-1}N = N$, but as N is normal, it is also true that $ab^{-1}N = aNb^{-1}$, so that $aNb^{-1} = N$, and thus $aN = Nb = bN$. What this says is that, if $\overline{\varphi}(aN) = \overline{\varphi}(bN)$, then $aN = bN$, so the map is injective. The map is of course surjective onto its image, so the map is an isomorphism. □

It is convenient here to introduce the concept of a commutative diagram. A commutative diagram is a directed graph where vertices are sets and edges are functions between the sets it connects, with the following property. If there are two paths

$$S \xrightarrow{f_1} A_1 \xrightarrow{f_2} \dots \xrightarrow{f_{n-1}} A_n \xrightarrow{f_n} E$$

$$S \xrightarrow{g_1} B_1 \xrightarrow{g_2} \dots \xrightarrow{g_{m-1}} B_m \xrightarrow{g_m} E$$

from S to E, then $f_n \circ \cdots \circ f_1 = g_m \circ \cdots \circ f_1$. An example diagram is to the upper right, representing the functions in the first isomorphism theorem.



Another notation that is more lateral is to consider sequences of groups

$$G_1 \xrightarrow{f_1} G_2 \xrightarrow{f_2} \dots \xrightarrow{f_n} G_{n+1}$$

with arrows representing homomorphisms. This sequence is **exact** whenever $\text{im}(f_i) = \ker(f_{i+1})$ for any i from 1 to $n-1$. To test your knowledge of this, note that the sequence

$$\{e\} \to G \xrightarrow{f} H$$

being exact states exactly that f is an injective homomorphism, as there is only one homomorphism from $\{e\}$ to any group. Likewise,

$$G \xrightarrow{f} H \to \{e\}$$

states that f is surjective.

A simple application of the first isomorphism theorem is a way classify the cyclic groups. Specifically, a classification is a set of equivalence classes defined by the relation on groups $x \sim y$ if $x \cong y$. In algebra, we refer to these as the isomorphism classes, and finding a classification of some segment of groups is a primary goal of group theory. One can only really say they 'know' a group if, given another group, one can intuitively say whether the two groups are isomorphic or not.

THEOREM 4.8 (The Classification of Cyclic Groups). *Every cyclic group is isomorphic to either $\mathbf{Z}^+$ or $\mathbf{Z}^+/n\mathbf{Z}$ for some integer n.*

Proof. Let $\langle g \rangle$ be a cyclic group. Define a surjective homomorphism from $\mathbf{Z}^+$ to $\langle g \rangle$ by the mapping $r \mapsto g^r$. If $\langle g \rangle$ is order n, $n\mathbf{Z}^+$ is the kernel of the map. Then $\langle g \rangle \cong \mathbf{Z}^+/n\mathbf{Z}^+$ by the first isomorphism theorem. If $\langle g \rangle$ is infinite, the kernel of the map is $\{e\}$, and $\mathbf{Z}^+/0\mathbf{Z}^+ \cong \mathbf{Z}^+$, so $\langle g \rangle \cong \mathbf{Z}^+$. □

The first isomorphism is the catalyst to many important isomorphism theorems.

THEOREM 4.9 (The Second Isomorphism Theorem). *Let G be a group, and K and H subgroups such that $K \subset N_G(H)$. Then we have that*

$$H/(K \cap H) \cong HK/K$$

Proof. We have already justified in our discussion of cosets that $K \cap H$ will be normal in H, and K normal in HK. Define an assignment map from H to HK/K by

$$h \mapsto hK$$

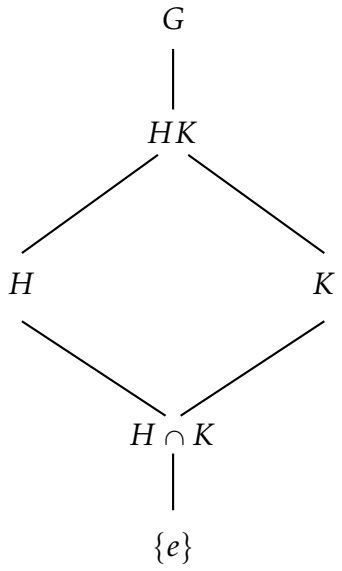
This is a surjective homomorphism, as any coset hkK can be written as a coset hK. If, for some element $h \in H$, $hK = K$, then it is sufficient and necessary for

$h \in K$ as well. Thus the kernel of the mapping defined above is $H \cap K$, and by the first isomorphism theorem, we obtain that

$$H/(H \cap K) \cong HK/K$$

□

The second isomorphism theorem is known as the diamond isomorphism theorem because of the lattice of subgroups below.



The final isomorphism theorem is the following, with commutative diagram inlaid.

FINISH THIS DIAGRAM.

$$\{e\} \longrightarrow M \longrightarrow G \longrightarrow G/M \longrightarrow \{e\}$$

$$\{e\} \longrightarrow M/N \longrightarrow G/N \longrightarrow G/M \longrightarrow \{e\}$$

THEOREM 4.10 (The Third Isomorphism Theorem). *Let M and N are normal subgroups of a group G, where N is also a normal subgroup of M. Then M/N is a normal subgroup of G/N, and $(G/N)/(M/N) \cong G/M$.*

Proof. Define an assignment from G/N to G/M by $gN \mapsto gM$. It is a surjective homomorphism, well defined as N is a subgroup of M, so that $gN \subseteq gM$ for any g. The kernel of this map are all sets of elements gN such that $gM = M$, which is precisely the elements g that are elements of M. Then the kernel is M/N (a normal subgroup), so by the first isomorphism theorem, we obtain that $(G/N)/(M/N) \cong G/M$. □

THEOREM 4.11 (The Lattice/Fourth Isomorphism Theorem). *Let G be a group, and N a normal subgroup. Then there is a bijection f from subgroups of G which contain N to subgroups of G/N. The expression $f(H)$ will be denoted $\overline{H}$. The bijection has the following properties for any two subgroups H and K:*

- *$H \subset K$ if and only if $\overline{H} \subset \overline{K}$.*
- *If $H \subset K$, $(H : K) = (\overline{H} : \overline{K})$.*
- *$\overline{\langle H, K \rangle} = \langle \overline{H}, \overline{K} \rangle$*
- *$\overline{A \cap B} = \overline{A} \cap \overline{B}$*
- *$H \lhd K$ if and only if $\overline{H} \lhd \overline{K}$*

Proof. Given a subgroup H of G which contains N, define a mapping by $h \mapsto hN$. The properties above can (and should be) be checked by the reader. □

We prove a tricky theorem now which will become very important when we investigate the theory of solvable groups. It is called the butterfly lemma because of the diagram below.

INSERT PRETTY BUTTERFLY DIAGRAM (OH NO THIS IS GOING TO TAKE A WHILE TO GET RIGHT)

THEOREM 4.12 (The Butterfly Lemma: Zassenhaus' Lemma). *Let U and V be subgroups of a group G, and let U', V' be two other subgroups such that $U' \lhd U$, $V' \lhd V$. Then*

$$U'(U \cap V') \lhd U(U \cap V)$$
$$(U \cap V)V' \lhd (U \cap V)V'$$

and the factor groups are isomorphic:

$$\frac{U'(U \cap V)}{U'(U \cap V')} \cong \frac{(U \cap V)}{(U' \cap V)(U \cap V')} \cong \frac{V'(V \cap U)}{V'(V \cap U')}$$

Proof. Since the problem is symmetric in U and V, we need only prove the isomorphism from first to second, as then by swapping V to U we get third to second for free. Since $U \cap V$ is a subgroup of U, and U' is a normal subgroup of U, $U'(U \cap V)$ is a subgroup of U. Since U' is normal in U, U' is normal in all subgroups of U. For this proof specifically, it is needed that U' is normal in $U'(U \cap V)$. Thus we obtain a factor group $U'(U \cap V)/U'$. The second isomorphism tells us that, if K is a group, S and N are subgroups, and $N \lhd K$, then

$$\frac{S}{S \cap N} \cong \frac{SN}{N}$$

Setting $S = U \cap V$, and $N = U'$, $K = G$ we obtain that

$$\frac{U \cap V}{(U \cap V) \cap U'} \cong \frac{U'(U \cap V)}{U'}$$

Of course, $(U \cap V)/(U \cap V \cap U') = (U \cap V)/(U' \cap V)$, hence

$$\frac{U \cap V}{U' \cap V} \cong \frac{U'(U \cap V)}{U'}$$

The group $(U \cap V)/((U' \cap V)(U \cap V'))$ is a factor group of $(U \cap V)/(U' \cap V)$; that is,

$$\frac{U \cap V}{(U' \cap V)(U \cap V')} = \frac{(U \cap V)/(U' \cap V)}{U \cap V'}$$

This follows as, in general, for any groups S, T, and K for which this statement makes sense,

$$\frac{G}{HK} = \frac{G/H}{K}$$

Because of this factor group, and the isomorphism obtained from the second theorem, we obtain a surjective homomorphism from $U'(U \cap V)/U'$ to $(U \cap V)/((U' \cap V)(U \cap V'))$,

$$sU' \mapsto s(U' \cap V)(U \cap V')$$

The kernel of this homomorphism is exactly $(U'(U \cap V'))/U'$. From the first isomorphism theorem,

$$\frac{U'(U \cap V)/U'}{U'(U \cap V')/U'} \cong \frac{U \cap V}{(U' \cap V)(U \cap V')}$$

And from the third isomorphism theorem, we can replace the left hand side to obtain the needed isomorphism.

$$\frac{U'(U \cap V)}{U'(U \cap V')} \cong \frac{U \cap V}{(U' \cap V)(U \cap V')}$$

Now, swapping U and V,

$$\frac{V'(V \cap U)}{V'(V \cap U')} \cong \frac{V \cap U}{(V' \cap U)(V \cap U')}$$

Note that both hand sides of the equations above are the same, hence we obtain the butterfly isomorphism. □

The proof of this theorem shows that isomorphism theorems are powerful, but because of this power, the proof is very obtuse. If you aren't able to understand the proof, just try and understand the isomorphism formed.

Exercise 5. *Let a be an element of a group of finite order, and f a homomorphism. Show that the order of $f(a)$ divides the order of a.*

Proof. By the first isomorphism theorem,

$$\langle a \rangle / \ker(f) \cong f(\langle a \rangle)$$

And we know the order of $\langle a \rangle / \ker(f)$ divides the order of $\langle a \rangle$. □

EXERCISE 6. *This exercise has two parts.*

(1) *If S and T are subgroups of a group G, then a $(S-T)$ double coset is a subset of G of the form SgT, where $g \in G$. Prove that the set of all $(S-T)$ double cosets partitions the group.*
(2) *Let S and T be subgroups of a finite group G, and suppose for some sequence $(g_1, g_2, \ldots, g_n)$ such that the double cosets Sg_kT are disjoint, we have that*

$$G = \bigcup Sg_kT$$

Prove that

$$G = \sum_{k=0}^{n} (S : S \cap g_k T g_k^{-1})$$

Note this is a generalization of Lagrange's theorem, which results when $T = \{e\}$

EXERCISE 7. *A normal subgroup N of G is maximal if there is no normal subgroup that contains N properly. Show that N is maximal if and only if G/N is simple.*

Proof. The lattice isomorphism theorem shows that a group is normal in G/N if and only if there is a corresponding normal subgroup of G containing N. □

This exercise has important properties in the theory of solvable groups, a theory which we will study later.

CHAPTER 5

Group Actions and The Symmetric Group

The symmetric group was previously defined as the set of permutations on a set. In the context of an example, this group seems trivial, but this is not so. One reason why the group is generally interesting is Cayley's theorem, which relates the set of groups to all other groups.

THEOREM 5.1 (Cayley's Theorem). *Every group is isomorphic to a subgroup of a symmetric group:*

Proof. Let G be a group. For each $g \in G$, define a permutation π_g on the group defined by the map $h \mapsto gh$. The function is a permutation as it is bijective – there is an inverse function $h \mapsto g^{-1}h$. The map from the group to its permutation is a homomorphism as for any two elements g and g' $\pi_g \circ \pi_{g'} = \pi_{gg'}$. Furthermore, the homomorphism is injective, as if $\pi_g = \mathrm{id}$, then $gh = h$ for all elements h, and for any specific one, we obtain that $g = e$. Thus G is isomorphic to the image of the permutation map, which is a subgroup of $S_{|G|}$. □

Intuitively, what Cayley's theorem states is that every every element of a group can be considered a symmetry of some set of objects. For instance $\mathbf{Z}$, the number n can really be considered the symmetry of adding n to every number in $\mathbf{Z}$, shifting all numbers to the right by n such that the resultant object is symmetric to the original. In general, a homomorphism onto a familiar group is known as a representation.

From Cayley's theorem, it follows that anything algebraically we prove about the subgroups of symmetric groups follows for all groups by the isomorphism property. Thus we will spend the rest of this chapter focusing on the components that form the symmetric group.

Through Cayley's theorem, all groups can be considered subgroups of the symmetric group, hence all groups can be considered a symmetric action on some set. These actions provide another way to understand the structure of a group. We now describe these actions in detail.

DEFINITION 18. A **group action** on a set G and set X is a homomorphism π from G to the symmetry group on $|X|$ characters. To be concise, we write gs for the permutation $(\pi(g))(s)$ associated with g acting on s. We call X a **G-set** .

It is simple to show that, for any group action G on a G-set X, we have that, for all g and h in G and x in X, $g(hx) = (gh)x$, and $ex = x$. These properties follow directly from the definition of a homomorphism; another way of saying the first statement is that, if φ is the homomorphism defining the action,

$$\varphi(g) \circ \varphi(h) = \varphi(gh)$$

The second statement says

$$\varphi(e) = \mathbf{1}$$

where $\mathbf{1}$ is the identity transformation. This is just a definition of a homomorphism from a group to the symmetric group, hence any map satisfying these properties is a group action.

A basic example of a group action is to consider G a G-set on itself by conjugation. That is, our group action is defined by

$$g(x) \mapsto g^{-1}xg$$

It is trivial to verify the properties above. What's more, the permutation associated with any g in G is an automorphism of G. This does not always hold when the G-set is a group. We call any automorphism of this form an **inner** automorphism.

Definition 19. Given a group G, and a G-set S, for $s \in S$, let the **orbit** of s be Gs, the set of all gs for $g \in G$. Let the set of all orbits be denoted X/G.

The relation on a G-set defined by $x \sim y$ if $Gx = Gy$ is an equivalence relation and partitions the set into orbits of S. What this means is that the group acts independently on each of a G-set's orbits.

Definition 20. A G-set X is **transitive** if it has just one orbit. This just means that for any two elements x and y in X, there is some g in G such that $gx = y$.

Definition 21. An action is **faithful** if the homomorphism defining it is injective, which means that no group element other than the identity acts idempotently to the G-set associated with the group action.

Definition 22. A map α from a G-set X to a G-set Y is a G-morphism if $\alpha(gx) = g\alpha(x)$ for all $g \in G$ and $x \in X$. α is a G-isomorphism if it is bijective.

Like homomorphisms between groups, G-morphisms and isomorphisms embed the algebraic structure of one set into another. The only algebraic structured assumed on a G-sets is its relation to G, so we must use the group action to define the isomorphism.

Definition 23. An element x in a G-set X is a **fixed point** if $gx = x$ for every $g \in G$. This means exactly that the permutation associated with g is $\mathbf{1}$ in the symmetric group; g is in the kernel of the group action. The set of all fixed points is denoted X^G.

Definition 24. Given any $x \in X$, the set G_x defined as

$$\{g \in G : gx = x\}$$

is a subgroup called the **isotropy subgroup** or **stabilizer** of x in G, and is normal in G.

As an example, let G act on itself by conjugation. The isotopy subgroups are called centralizers $C_G(h) = \{g \in G : gh = hg\}$. A fixed point is called a center, and the set of all centers is denoted $Z(G)$, which we have previously shown as the center group. In general, for any group G and G-set X,

$$X^G = \bigcap_{x \in X} G_x$$

As another example, consider conjugation from G on its subgroups defined by the mapping

$$gH \mapsto gHg^{-1}$$

Then the isotropy group of a subgroup H is the normalizer $N_G(H)$. The fixed points of this transformation are precisely the normal subgroups.

As a final more complicated example, consider the group $SL_n(\mathbf{R})$ acting on the upper half of the complex plane, the set

$$\{z \in \mathbf{C} : \mathrm{im}(z) > 0\}$$

by the mobius transform

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az + b}{cz + d}$$

This defines a transitive action. The isotropy subgroup of the imaginary number i is the special orthogonal group $SO(2)$, the set of matrices with orthonormal columns. A meromorphic function on H invariant under $SO(2)$ is called a modular function, and is essential to the study of number theory, string theory, and the study of monstrous moonshine.

We now give a theorem which establishes an intricate connection between G and its G-sets.

Theorem 5.2 (Orbit Stabilizer Lemma). *Let X be a G-set. Then, for every x in X, there exists a G-isomorphism from G/G_x to Gx. It follows that*

$$|Gx| = (G : G_x)$$

Proof. Define a mapping by

$$gG_x \mapsto gx$$

We leave the reader to verify this is a well defined function. The reasoning is similar to the verification of the function created in the first isomorphism theorem. This mapping is surjective by construction, and furthermore, the map

is injective. If $gx = hx$, then $(h^{-1}g)x = x$, hence $(h^{-1}g) \in G_x$, so $gG_x = hG_x$. The mapping is also a G-isomorphism, hence we have constructed the required isomorphism. □

Corollary 5.3 (The Orbit Decomposition Formula). *Given a G-set X, with a finite number of orbits $(X_1, X_2, \ldots, X_n)$. From each orbit, pick a representative x_i. Then we have*

$$|X| = \sum_{k=1}^{n} (G : G_{x_i})$$

which we call the orbit decomposition formula. In particular, if we let $X_n = |X^G|$

$$|X| = |X^G| + \sum_{k=1}^{n-1} (G : G_{x_i})$$

Proof. X is the disjoint union of its orbits. Hence

$$|X| = \sum_{k=1}^{n} |Gx_i|$$

But we have contructed an isomorphism from Gx_i to G/G_{x_i} above, hence

$$|Gx_i| = |G/G_{x_i}|$$

and we obtain the final formula by Lagrange's theorem. □

The next theorem is just a restatement of the previous.

Corollary 5.4 (The Class Equation). *Consider the group action of conjugation from a group G onto itself. Then*

$$|G| = |Z(G)| + \sum_{i=1}^{n-1} (G : C_G(x_i))$$

This theorem will be very useful for our next topic of study, Sylow theory. Before we get into this theory, let us consider an example to show the power of the class equation. Consider a group of order 55 acting on a set of order 39. We claim there is at least one fixed point in the group action. The orbit decomposition formula entails that we have

$$|X| = 39 = |X^G| + \sum_{k=1}^{n} (G : G_{x_i})$$

Each G_{x_i} forms a subgroup of G, hence by Lagrange's theorem, $|G_{x_i}| \mid 55$, so $|G_{x_i}|$ is either 1, 5, 11, or 55. If $|G_{x_i}| = k$, then $(G : G_{x_i}) = 55/k$, so if we let m_j denote the number of orbits whose isotropy subgroups are order j. Then

$$39 = 55m_1 + 11m_5 + 5m_{11} + m_{55}$$

Showing that there is at least one fixed point is the same as showing there is an isotropy group of order 55, for this means that some element in X is fixed by every point in G, and hence a fixed point. By considering all possible solutions to the equations above, we obtain that $m_{55} \geqslant 1$ and hence the theorem.

Lemma 5.5 (Burnside's Lemma). *If X is a finite G-set, then*

$$|X/G||G| = \sum_{g \in G} |X^g|$$

Proof. By a simple calculation,

$$\sum_{g \in G} |X^g| = |\{(g,x) : gx = x\}| = \sum_{x \in X} |G_x|$$

Combining this calculation with the orbit stabilizer lemma, we obtain that

$$\sum_{x \in X} |G_x| = \sum_{x \in X} |G|(G : G_x)^{-1} = |G| \sum_{x \in X} (G : G_x)^{-1}$$

Now $(G : G_x) = |Gx|$, hence

$$|G| \sum_{x \in X} (G : G_x)^{-1} = |G| \sum_{x \in X} |Gx|^{-1}$$

Now partion X into its orbit X/G. For each x and y in a particular orbit, it is obvious that $|Gx| = |Gy|$. Hence, if we have a partition $(X_1, X_2, \ldots, X_{|X/G|})$, and we pick representatives from each x_i from each X_i, we have that

$$|G| \sum_{x \in X} |Gx|^{-1} = |G| \sum_{k=1}^{|X/G|} |X_k||Gx_k|^{-1}$$

Now for each $|X_k|$, we have that $|Gx_i| = |X_k|$ by definition, so finally, we obtain that

$$|G| \sum_{k=1}^{|X/G|} |X_k||Gx_k|^{-1} = |G| \sum_{k=1}^{|X/G|} |Gx_i|/|Gx_i| = |G| \sum_{k=1}^{|X/G|} 1 = |G||X/G|$$

and by transitivity, our proof is complete. □

Before we get into Sylow theory, let us establish some interesting facts about the symmetric group. First, of course, we must define some facts.

Definition 25. Given a set M and a permutation π on M, the **support** of π, denoted $\sup(\pi)$, is defined as the set

$$\{m \in M : \pi(m) \neq m\}$$

A **cycle** of length k is a permutation π such that $|\sup(\pi)| = k$, and we can order $\sup(\pi)$ to be $(x_0, x_1, \ldots, x_{k-1})$ in a way that $\pi(x_n) = x_{n+1 \bmod k}$. A cycle of length two is called a transposition .

We denote a cycle like π as $(x_1, x_2, \ldots, x_k)$.

If σ and τ are two permutations, such that $\sup(\sigma) \cap \sup(\tau) = \varnothing$, $\sigma \circ \tau = \tau \circ \sigma$. This is because the two act independently on the set they permute.

THEOREM 5.6. *Every permutation on a finite non-empty set which is not the identity can be written as the product of cycles with disjoint support. This is unique up to reordering:*

Proof. Let σ be an arbitrary element of the symmetric group S_n, and consider the cyclic group generated by σ. Consider the set $\{1, 2, \ldots, n\}$, with $\langle\sigma\rangle$ acting on the set by the mapping

$$\pi k = \pi(k)$$

in the obvious manner. We obtain disjoint partitions of orbits from this action. We claim that π when restricted to this orbit is a cycle, and thus π consists of products of cycles from each orbit. Consider an orbit $(\langle\pi\rangle k)$ for some number k between one and n. Every integer in k's orbit can be written $\pi^m(k)$ for some integer m. For each integer l in the range, associate it with the smallest positive integer m such that $\pi^m(k) = l$. We obtain an ordering

$$(\pi^0(k), \pi^1(k), \pi^2(k), \ldots, \pi^n(k))$$

such that $\pi^{n+1}(k) = k$. This generates a cycle, and we have shown what was needed. □

If a permutation π is equal to the disjoint composition of cycles $\sigma_1, \sigma_2, \ldots, \sigma_n$, then we write $\pi = \sigma_1\sigma_2 \ldots \sigma_n$. Every permutation on a finite set can be written in this way.

We would like to specify a specific set of permutations having the property of 'evenness', like the integers. Specifically, we would like the following properties:

(1) The composition of two even permutations is even.
(2) The composition of two odd (not even) permutations is even.
(3) The composition of an odd and even permutation is odd.

With the properties above, we can consider the property of 'evenness' to be a homomorphism from S_n to the multiplicative group consisting of ± 1. If $f(\pi) = 1$, then π is even. Thus our task is to characterize a homomorphism with this property. From elementary properties of homomorphisms, we know that $\mathbf{1}$ must be even (it is in the kernel). In addition,

(1) The inverse of an even permutation is even.
(2) The inverse of an odd permutation is odd.

Let us add the additional characteristic that any transposition must be odd. Then it follows that there is only one homomorphism with the properties above. The next few lemmas will establish this claim.

LEMMA 5.7. *S_n is generated by transpositions.*

Proof. We have proved that S_n is generated by cycles, hence we need only prove that each cycle can be decomposed into transpositions. The calculation below shows that this is true.

$$(x_1, x_2, \ldots, x_n) = (x_1\ x_n)(x_1\ x_{n-1})\ldots(x_1\ x_2)$$

□

Now of course, by Theorem (4.6) we may conclude that if there exists a homomorphism with the properties above, then it must be unique. Thus we need only establish that there exists a homomorphism with the properties above. We state the theorem in full.

THEOREM 5.8. *There is a unique homomorphism from S_n to $\{\pm 1\}$ such that the mapping from any transposition is -1.*

Proof. Consider a polynomial P defined for any tuple of natural numbers by

$$P(x_1, x_2, \ldots, x_n) = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i)$$

Define the map sgn from S_n to $\{\pm 1\}$ by

$$sgn(\pi) = \frac{P(\pi(1), \pi(2), \ldots, \pi(n))}{P(1, 2, \ldots, n)}$$

For any factor $(x_i - x_j)$ in $P(1, 2, \ldots, n)$, we either have the factor $(x_j - x_i)$ or the factor $(x_i - x_j)$ in $P(\pi(1), \pi(2), \ldots, n)$ (π just permutes the orders of the elements, hence the numerator and denominator only differ by sign, and the value of sgn is always positive or negative one. We have that

$$\frac{\pi(i) - \pi(j)}{i - j} = \frac{\pi(j) - \pi(i)}{j - i}$$

Therefore it does not matter whether $i < j$ as much as we do not add the same fraction twice. We conclude, for two permutations π and σ, that

$$\begin{aligned} sgn(\pi \circ \sigma) &= \prod_{1 \leqslant i < j \leqslant n} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{i - j} \\ &= \prod_{1 \leqslant i < j \leqslant n} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{\sigma(i) - \sigma(j)} \prod_{1 \leqslant i < j \leqslant n} \frac{\sigma(i) - \sigma(j)}{i - j} \\ &= \prod_{1 \leqslant \sigma(i) < \sigma(j) \leqslant n} \frac{\pi(\sigma(i)) - \pi(\sigma(j))}{\sigma(i) - \sigma(j)} sgn(\sigma) \\ &= \prod_{1 \leqslant i < jn} \frac{\pi(i) - \pi(j)}{i - j} sgn(\sigma) \\ &= sgn(\pi) \cdot sgn(\sigma) \end{aligned}$$

Here the third and fourth equality works because σ is a permutation of the numbers from 1 to n. From this calculation, we conclude sgn is a homomorphism; all that is left to prove is that for any transposition (x_1, x_2), $sgm((x_1, x_2)) = -1$.

$$\begin{aligned} sgn((x_1, x_2)) &= \left(\prod_{\substack{1 \leqslant i < j \leqslant n \\ (i,j) \neq (x_1, x_2)}} \frac{i - j}{i - j} \right) \frac{x_2 - x_1}{x_1 - x_2} \\ &= -\frac{x_2 - x_1}{x_2 - x_1} \\ &= -1 \end{aligned}$$

Thus we have constructed the isomorphism that we wanted. □

The sgn map was constructed only to satisfy the proof. Here is a simpler way to think of the map. We know that any permutation π in S_n can be decomposed into the product of a finite number of transpositions. If we let k denote the number of transpositions, then we have that

$$sgn(\pi) = (-1)^k$$

which follows exactly from the homomorphic properties of the sign function.

Exercise 8. *If X is a G set that is not a singleton, then there is an element x in X with no fixed points.*

Lemma 5.9. *Let $\pi \in S_X$ and $\sigma = (x_1\ x_2\ \ldots\ x_n)$. Then it follows that*

$$\pi\sigma\pi^{-1} = (\pi(x_1)\ \pi(x_2)\ \ldots\ \pi(x_n))$$

Proof. This follows as $\pi\sigma\pi^{-1}(\pi(x_i)) = \pi\sigma(x_i) = \pi(x_{i+1})$. If $x \notin \sup(\sigma)$, then $\pi\sigma\pi^{-1}(\pi(x)) = (\pi\sigma)(x) = \pi(x)$, so that $\pi(x) \notin \sup(\pi\sigma\pi^{-1})$. □

The kernel of this is A_n, the alternating group , a normal subgroup of S_n. Here are some properties of A_n.

LEMMA 5.10. *If τ is a transposition, $S_n = A_n \cup \tau A_n$. Thus $A_n = n!/2$.*

LEMMA 5.11. *A_n is generated by the set of all three cycles*

Proof. We need only prove that the product of two arbitrary transpositions $(a\ b)$ and $(c\ d)$ is generated by three cycles. If $\sup(a\ b) \cap \sup(c\ d) = \varnothing$, $(a\ b)(c\ d) = (a\ c\ b)(a\ c\ d)$. Otherwise without loss of generality, we may consider $a = c$. If $b = d$, then $(a\ b)(c\ d) = \mathbf{1}$. If $b \neq d$, $(a\ b)(c\ d) = (a\ b)(c\ d) = (c\ b\ a)$. Since A_n is generated by pairs of transpositions, it is generated by three cycles. □

Saving the hardest for last:

LEMMA 5.12. *A_n is simple when $n \neq 4$*

Proof. This proof is about a page long. I will write it later. □

The fact that A_4 is not simple results in far reaching ramifications in Galois theory, where it implies that there is no formula for finding the roots of quintic polynomial roots.

CHAPTER 6

Sylow Theorems

In 1872, Norwegian Ludwig Sylow proved a collection of theorems, called the Sylow theorems , which give detailed information about the number of subgroups of a certain size that a group contains. Unlike the majority of chapters in this book, we begin with a theorem, rather than a definition.

THEOREM 6.1. *For every finite abelian group, and every prime number which divides the order of the group, there is an element whose order is that prime number.*

Proof. Let G be an abelian group, and p a prime number such that $p \mid |G|$. We prove this statement by induction on $|G|$. When $|G| = 1$, the statement holds vacously. Now suppose this theorem holds for all group sizes less than the order of another group G. Take $g \neq e$ in G. If the order of g is divisible by that prime number, that is, if there is some integer m such that $g^{mp} = e$, then g^m is order p, and we are done. Thus we may assume without any loss of generality that g is not divisible by that prime number. Since G is abelian, $\langle g \rangle$ is normal, hence we can form the group $G/\langle g \rangle$. We have the fact that $|G| = |G/\langle g \rangle||\langle g \rangle|$, and since $\langle g \rangle$ has an order greater than one, since $g \neq e$, then the order of $G/\langle g \rangle$ is less than the order of G, and since $p \nmid \langle g \rangle$, we know $p \mid G/\langle g \rangle$. We conclude there is some element $h\langle g \rangle$ that is order p. Thus $(h\langle g \rangle)^p = h^p\langle g \rangle = \langle g \rangle$. Then we conclude that $h^p = e$, and thus p divides the order of h. Using the case proved earlier, we must have that some power of h is order p. □

A theorem of Cauchy generalizes this to arbitrary groups.

THEOREM 6.2 (Cauchy's theorem). *Given any group whose order divides a prime, there is an element whose order is that prime.*

Proof. We prove this theorem by induction again. We prove no base case as a group of any size less than 6 is abelian and thus we can apply theorem (12.1). Now suppose the theorem holds for all groups of order less than a group G. Let p be a prime, and suppose $p \mid |G|$. If G contains a proper subgroup whose order is divisible by p, then we can apply induction rather easily to show that this theorem holds for G. The hard part is when G contains no proper subgroup whose order is divisible by p. Consider G acting on itself by conjugation. For every element g, the centralizer $C_G(g)$ is a subgroup of G. By Lagrange's theorem,

$$|G| = |C_G(g)|(G : C_G(g))$$

The class equation also gives us that

$$|G| = |Z(g)| + \sum_{k=1}^{n-1} (G : C_G(x_i))$$

If g is not in $Z(g)$, $C_G(g)$ is a proper subgroup of G, so by our assumption $p \nmid |C_G(g)|$, so by the equation created by Lagrange's theorem, we obtain that $p \mid (G : C_G(g))$. But then by rearranging the class equation, we obtain that $p \mid |Z(g)|$, hence $Z(g)$ cannot be a proper subgroup, and so $G = Z(g)$. Thus G is abelian, and we can apply (12.1) again. □

Definition 26. Let p be a prime number. A group G is called a **p-group** if the groups order is a power of p.

By Cauchy's theorem, a group is a p-group if and only if every element has order a power of a prime.

Lemma 6.3. *Let G be a p-group. If G acts on a finite set X, then the fixed points X^G satisfies*

$$|X^G| \equiv |X| \pmod{p}$$

Proof. It was previously proven that $|X| = |X^G| + \sum_{k=1}^{n-1} (G : G_{x_i})$, the class equation. For each G_{x_i}, we have that $p \mid (G : G_{x_i})$ by an easy application of Lagrange's theorem. This shows exactly the equation we were attempting to prove. □

Lemma 6.4. *Let $G \neq \{e\}$ be a p-group. Then the center $Z(G) \neq \{e\}$.*

Proof. Let G act on itself by conjugation. Then by Lemma (12.3), we have the $|Z(G)| \equiv |G| \pmod{p}$, so $|Z(G)| \equiv 0 \pmod{p}$ since $p \mid G$. We obtain that there are at least p elements that are fixed points, since we cannot have 0 elements. □

Corollary 6.5. *Let p be a prime. Every group of order p^2 is abelian.*

Proof. Let G be a group of order p^2. According to Lemma (12.4), the center $Z(G)$ of G is non-trivial. Since $Z(G)$ is a subgroup, it thus must be order p or p^2 by Lagrange's theorem. Suppose that $Z(G)$ is order p, and let h be an element such that $h \notin Z(G)$. Also consider conjugation acting from G to itself. Then G_x is a group larger than $Z(G)$ since $x \in G_x$ and $x \in Z(G)$, so we conclude that G_x must be order p^2. But then, of course, x commutes with every element, so $x \in Z(G)$, a contradiction. Hence $Z(G)$ is order p^2, and it follows that G is abelian. □

Now, to the real meat of the chapter!

Definition 27. Let G be a group of order $p^m q$, where p is a prime and q and p are relatively prime. Then a subgroup is called a **p-Sylow Subgroup** if the order of the subgroup is a power of p^m.

In the next few proofs, let G be a group with the same notation of the definition of a Sylow subgroup.

LEMMA 6.6. *For every k such that $1 \leqslant k \leqslant m$, there is a subgroup of G of order p^k.*

Proof. We prove by induction on the size of m. Observe if $m = 0$, the theorem holds vacously. Consider the conjugation of G acting on itself. We know that

$$|G| = |Z(G)| + \sum_{i=1}^{n-1} (G : C_G(x_i))$$

We consider two cases to our proof. One where p divides the center group, and one where it does not.

Suppose that $p \nmid |Z(G)|$. This implies that there is at least one x_i such that $p \nmid (G : C_G(x_i))$, as otherwise we could move the indexes to the left hand side of the equation and conclude that $p \mid |Z(G)|$. By Lagrange's theorem, $|G| = (G : C_G(x_i))|C_G(x_i)|$, and hence $p \mid |C_G(x_i)|$. We know that $|C_G(x_i)| = p^m q'$, as the index takes no powers of p away, and $q' < q$, as the index cannot be 1. Hence we can use induction to show there is a subgroup of order p^k in the $C_G(x_i)$ and hence in G.

On the other size, suppose $p \mid |Z(G)|$. By Cauchy's theorem, we conclude there is some element g of order p. Since $Z(G)$ commutes with elements of G, every subgroup of $Z(G)$ is normal in G. Thus $\langle g \rangle \triangleleft G$. $G/\langle g \rangle$ is thus a group of order $p^{m-1}q$, so by induction there is a subgroup H of $G/\langle g \rangle$ such that $|H| = p^{k-1}$. H can be written as $V/\langle g \rangle$ for some subgroup V of G, and by Lagrange's theorem, $|V| = |H||\langle g \rangle| = p^{k-1}p = p^k$. □

LEMMA 6.7. *If H is a p-subgroup of G, and S is a p-Sylow Subgroup of G, then there is an element $g \in G$ such that $H \subset gSg^{-1}$.*

Proof. H acts of G/S by the mapping $h(gS) \mapsto hgS$. By lemma (12.3),

$$|X^H| \equiv |X| \pmod{p} = q \bmod p$$

q is relatively prime to p, and thus in particular, $q \not\equiv 0 \pmod{p}$. Thus the set of fixed point $|X^H|$ contains at least one element g such that $hgS = gS$ for all elements $h \in H$, but this means precisely that for some elements $s, s' \in S$, $hgs = gs'$, and thus $h = gs's^{-1}g^{-1}$, so that $h \in gSg^{-1}$. We conclude $H \subseteq gSg^{-1}$ □

THEOREM 6.8. *Let s be the number of p-Sylow Subgroups of G. Then $s|q$.*

Proof. Let S be a p-Sylow subgroup of G of order p^k, and let X be the set of all p-sylow subgroups of G. Since all p-Sylow subgroups are conjugate to each other, the action of conjugation from G on X is transitive. Consider the normalizer $N_G(S)$. We obtain the class equation

$$|X| = (G : N_G(S))$$

hence $(G : N_G(S)) = s$. By the multiplicative property of indices,

$$(G : S) = (G : N_G(S))(N_G(S) : S)$$

By Lagrange's Theorem, we get that $(G : S) = |G|/|S| = p^m q/p^m = q$, hence $s \mid q$. □

THEOREM 6.9. *Keeping the notation the same as in the last theorem,* $s \equiv 1 \pmod{p}$

Proof. S acts on X via conjugation. We claim that S is the only fixed point in this action. To prove this, we first require a second claim; if a p-group H is contained in $N_G(S')$, then H is contained in S'. This follows as S' is a subgroup (that is also normal) of $N_G(S')$, hence HS' is a subgroup of $N_G(S')$. By the third isomorphism theorem, $HS'/S' \cong H/H \cap S'$. $H/H \cap S'$ is a p-group, as it is the factor group of a p-group. By Lagrange's theorem, we have that $|HS'| \mid |G|$, and thus $|HS'|/|S'| \mid |G|/|S'|$, which states precisely that $(HS' : S') \mid (G : S')$. As there is a bijection from HS'/S' to $H/H \cap S'$, resulting from the third isomorphism theorem, it follows that $(H : H \cap S')$, which is equal to $(HS' : S')$, divides $(G : S')$. However, $(G : S') = q$, and q is coprime to p, so we must have that $|H|/|H \cap S'| \mid q$. But $|H|$ and $|H \cap S'|$ cannot possibly divide q, as they are both p-subgroups, so we must have that $|H| = |H \cap S'|$, and thus $H = H \cap S' = S'$. Hence $H \subset S'$, as we wanted.

Now we prove a final claim, that if S' is a fixed point, then $sS's^{-1} = S'$. Take $H = S$ above. Then $S \subset S'$, but both have the same cardinality as both are p-Sylow subgroups, hence $S = S'$, and we have proved our theorem, as then $|X| \equiv |X^S| \pmod{p}$, by lemma (12.3), and $|X^S| = |\{S\}| = 1$. □

THEOREM 6.10. *Let p and q be prime numbers such that $q < p$, and $p \nmid q$. Then every group of order pq is cyclic.*

Proof. Let S be a p-Sylow subgroup of G, and U a q-Sylow subgroup of G. Then the order of S is p and the order of U is q, and the groups are cylic. As the two are not equal, $S \cap U = \{e\}$, as this is a subgroup and thus must divide both primes. Let s be the number of p-Sylow subgroups, and r the number of q-Sylow subgroups. Then we know from theorem (12.9) that

$$r \equiv 1 \pmod{q} \qquad s \equiv 1 \pmod{p} \qquad s \mid q$$

As $s \mid q$, we know that $s = 1$ or $s = q$. If $s = q$, the $q \equiv 1 \pmod{p}$, hence $q - 1 \equiv 0 \pmod{p}$, and thus $p \mid q - 1$, a contradiction. Hence $s = 1$, and thus S is normal. It follows that SU is a subgroup of G □

CHAPTER 7

Solvability

Definition 28. Let G be a group. A **series** or **tower** is a finite sequence of groups

$$G = G_0 \supset G_1 \supset G_2 \supset \cdots \supset G_m$$

In this chapter we will focus on a very specific type of tower.

Definition 29. A tower of a group is called a **normal series** if for any group G_i that is not G_m, $G_{i+1} \lhd G_i$. A normal series is **abelian** if G_i/G_{i+1} is abelian for all groups G_{i+1}, and cyclic if every one of these quotient groups are cyclic.

The following theorem follows as, for a homomorphism $f : G \to H$, with another group H' such that $H' \lhd H$, then $f^{-1}(H') \lhd f^{-1}(H)$.

Theorem 7.1. *Consider a normal tower*

$$H_m \lhd H_{m-1} \lhd \cdots \lhd H_0 = H$$

and a homomorphism $f : G \to H$. Define a tower on G by $G_i = f^{-1}(H_i)$. Then this tower is normal.

This theorem, along with the following theorem, are left to be filled in by the reader. To prove the next one, note that the mapping $gG_{i+1} \mapsto f(g)H_{i+1}$ is an injective homomorphism.

Theorem 7.2. *Consider the same towers are in the previous theorem. If the tower based from H is abelian, then so is the tower on G. Furthermore, if H's tower is cyclic, so is G's.*

Every group has a normal tower. For any group G, simply take the tower

$$G \supset \{e\}$$

we want to take a tower which is maximalized in some way, to strain out some properties from the groups formed.

Definition 30. A **refinement** of a tower is a new tower obtained by inserting finitely more subgroups into the original tower.

Definition 31. We say two normal series S and T are **equivalent** if they have the same length and such that there is a permutation φ such that, for any group S_i in S but the terminating element,

$$S_i/S_{i+1} \cong T_{\varphi(i)}/T_{\varphi(i)+1}$$

Theorem 7.3. *Two normal series in a group G ending with the trivial group have refinements that are equivalent.*

Proof. Consider two normal towers

$$G = G_0 \supset G_1 \supset \cdots \supset G_n = \{e\}$$

$$G = H_0 \supset H_1 \supset \cdots \supset H_m = \{e\}$$

Define $G_{i,j} = G_{i+1}(H_j \cap G_i)$ for i between 0 and $n-1$ and j between 0 and m. Then we have the tower

$$\begin{aligned} G = G_1(G) &= G_1(H_0 \cap G_0) \\ &= G_{0,0} \supset G_{0,1} \supset \cdots \supset G_{0,m} \supset G_{1,0} \supset \ldots \supset G_{n-1,m} \\ &= G_n(H_m \cap G_{n-1}) = \{e\} \end{aligned}$$

Similarily, if we define $H_{ij} = H_{i+1}(G_j \cap H_i)$, with a tower of H_j generated in a similar fashion. By the butterfly lemma, with $U = G_{i+1}$, $U' = G_i$, $V = H_{j+1}$, and $V' = H_j$, we obtain that the above towers are normal, and equivalent, as

$$G_{i,j}/G_{i,j+1} \cong H_{i,j}/H_{i,j+1}$$

and each tower is a refinement of the original towers, as

$$G_{k,1} = G_k(H_1 \cap G_{k-1}) = G_k(G \cap G_{k-1}) = G_k G_{k-1} = G_k$$

and similarily for $H_{k,1}$. □

Theorem 7.4. *From any abelian tower we can construct a cyclic tower.*

Proof. We prove this theorem by induction on the order of the group. Consider a group G of order n where an abelian tower of any smaller group can be constructed into a cyclic tower. Suppose we have an abelian tower

$$G = G_0 \supset G_1 \supset \cdots \supset G_m$$

Consider a non-zero group element g in G, and the quotient group $G/\langle g \rangle$. We still have an abelian tower

$$G_0/\langle g \rangle \supset G_1/\langle g \rangle \supset \cdots \supset G_m/\langle g \rangle$$

Because by the third isomorphism theorem, the quotient groups are isomorphic to the original abelian tower's quotient groups. By induction, we can construct refine this tower into a cyclic tower. We have the canonical homomorphism from G to $G/\langle g \rangle$, hence the inverse image is a cyclic tower in G. □

Definition 32. A group is **solvable** if it has an abelian tower whose last element is the trivial subgroup $\{e\}$.

Theorem 7.5. *A group G is solvable if and only if, for any normal subgroup H, H and G/H are solvable.*

Proof. Consider the tower that makes G solvable.

$$G = G_0 \supset G_1 \supset \cdots \supset G_n = \{e\}$$

From this tower, construct a new tower

$$H = H \cap G_0 \supset G_1 \cap H \supset \cdots \supset G_n \cap H = \{e\}$$

Now H_i/H_{i+1} □

CHAPTER 8

Direct Products, Semiproducts, and Abelian Groups

Definition 33. Let I be an index set, and $\{G_i\}_{i\in I}$ a family of groups. Then the direct product of $\{G_i\}$, denoted $\times_{i\in I} G_i$, is a group with an operation defined by

$$\bigtimes_{i\in I} g_i \circ \bigtimes_{i\in I} h_i = \bigtimes_{i\in I} g_i h_i$$

The group is called the product group.

Theorem 8.1. *Let r and s be two relatively prime integers. Suppose G is a cyclic group of order rs. Then G is isomorphic to the direct product of cyclic groups R and S, where R is order r and S is order s.*

Proof. $R \times S$ is a cyclic group generated by (x, y), where $x^r = e$, and $y^s = e$, and the order of this cyclic group is rs. Since the cyclic groups are of the same order, they must be isomorphic. □

Theorem 8.2. *Let H and K be normal subgroups of a group G, such that $H \cap K = \{e\}$, and $HK = G$. Then $H \times K \cong G$.*

Proof. Define a map from $H \times K$ to G by $(h, k) \mapsto hk$. The map is bijective, as $HK = G$, and if $hk = e$, $h = k^{-1}$, so k^{-1} is in H, so $k = h = e$. We know that hkh^{-1} is contained in K, as K is normal, but it is also contained in N as N is normal, hence $hkh^{-1}k^{-1} = e$, so $hk = kh$, and thus the map is a homomorphism as $h_1k_1h_2k_2 = h_1h_2k_1k_2$, and we have shown the map is an isomorphism. □

Index