Group Theory

Jacob Denson

October 26, 2014

Chapter 1

What is a group?

In 1761, Leonard Euler proved the following theorem. Take the totient function, a mapping that takes a positive integer, and returns the total number of integers less than the integer which are relative prime to that integer. A few examples are shown below.

$$\varphi(1) = 0$$

$$\varphi(10) = 3$$

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

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

Euler's proof involved mapping the set of numbers relatively prime to b to integer multiples of a and showing that, modulo b, these sets are the same. Since the ancient Greeks, mathematicians have been attempting to find integer solutions to Diophantine equations, integer values that result in zeroes to polynomial functions. In recent years this have given rise to the study of elliptic curves, an equation in the real plane of the form

$$y^2 = x^3 + Ax + B$$

Solutions to these equations interact in interesting ways which mean that the field is still going strong today. The most famous proof in recent history, Fermat's last theorem by Andrew Wiles, showed that these curves have correspondences to modular functions: shapes in the complex plane that have an inherent symmetrical property. Felix Klein in the 19th century attempted to answer questions about geometry by considering the inherent geometric properties of shapes. Everiste Galois had already used these techniques, but to explore the inherent symmetries in algebraic equations and determine the insolvability of the quintic equation. Though from separate fields, all of these problems are addressed by relating the objects to one another through different operations - whether they are symmetries or addition and multiplication of numbers. Though it was not known at the time of these discoveries, we know today the reason why these problems are connected is the objects studied follow the abstract concept of a group. This book attempts to introduce the reader to the concept of a group in order to explain the phenomena expressed above, as well as to show the reader the beauty of the abstract concepts developed.

Chapter 2

Operations and Groups

Groups are a way to generalize the notion of a set with a well behaved operation which creates a mechanism to combine objects in that set to form another. Operations on numbers are inherent, as are geometric symmetries. Since groups were defined in 1882 by Walter Dyck, the theory has arisen to become one of the largest fields in modern mathematics: abstract algera. 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. The longstanding questions shown above all lead to the group's definition. One of the tricks in every Mathematician's device is to solve a problem, and then to attempt to generalize the method used to solve the problem to apply to other techniques. It is only natural then that Group theory was to be generalized by these problems. It has proved to be one of the most useful ones to be created in the past two centuries.

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 are mathematical structures that we use almost everyday – for example, addition and multiplication of numbers. Thus, it is not surprising that generalizing the properties of everyday operations is quite difficult to see. 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 S. Note that in our definition we have inherently incorporated the property of **closure** – that the composition of any two objects in the set S lies inside the set S. If a and b are arguments to this function, the value mapped from a and b is most commonly denoted as ab, but also as $a \circ b$, or $a + b$, rather than $f(a, b)$, or some other common symbol commonly used for a function. This makes the notation less chunky, and keeps the intuitive definition of an operation as combinations of elements in our head while we work. A good notation like this can make working with the ideas of groups much simpler. While these symbols above are the most common for groups, any other symbol may be used for the operation.

We introduce a useful bit of notation to write down associative operations. Consider a finite sequence of elements

$$(x_1, x_2, \dots, x_n)$$

Then another notation convenience arises, the aptly named 'Pi' notation.

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

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

Make sure to note the similarity to the Σ notation used in sums of numbers.

Calculus is studied by narrowing down the functions it studies to narrower subsets with specific properties, the continuous and differentiable functions. The set of all functions is too general to be studied in totality. This is the same about studying human beings. We have vast differences, so very little can be said about the human race about the whole, but quite a lot can be said about certain groups of humans. As in these cases, the assignment property is too general for our study of groups, and thus we must specify subclasses of assignments which we will begin to study. The following characterizes how a group operates on the set. We will also show properties that elements of the set identify with the assignment.

Definition 2. *An assignment on a set is* **associative** *if for any three elements a, b, and c in that set, $a(bc) = (ab)c$.*

Altogether, this means any number of brackets written in an equation are irrelevant. The order in which elements are put together, however, is still important! In order to say this however, we must rigorously prove this. A formal specification of this conjecture is the following.

Theorem 2.0.1 (Bracket Equivalence)**.** *Given an associative operation on a set, and a finite sequence $(x_1, x_2, \ldots, x_n)$ of elements in that set a sequence of elements in that set, for any positive integers l and m such that $l+m=n$,*

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

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

$$\Big(\prod_{k=1}^{1} x_k\Big)\Big(\prod_{k=1}^{1} x_{k+1}\Big) = x_1 x_2 = \prod_{k=1}^{2} x_k$$

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 $m \geqslant 2$, as we need to be able to extract an element from the secon product. The case where $m=1$ is similar (just remove the product that ranges from 1 to $n-1$ in the equation).

$$\begin{aligned}
\Big(\prod_{k=1}^{l} x_k\Big)\Big(\prod_{k=1}^{m} x_{l+k}\Big) &= \Big(\prod_{k=1}^{l} x_k\Big)\Big(\Big(\prod_{k=1}^{m-1} x_{l+k}\Big)\ x_m\Big) && (2.1)\\
&= \Big(\prod_{k=1}^{l} x_k \prod_{k=1}^{m-1} x_{l+k}\Big)\ x_m && (2.2)\\
&= \Big(\prod_{k=1}^{n-1} x_k\Big)\ x_m && (2.3)\\
&= \prod_{k=1}^{n} x_k && (2.4)
\end{aligned}$$

Here (2.1) follows by definition of Pi notation, (2.2) follows from the associativity law, (2.3) results from the inductive hypothesis, and (2.4) follows by definition. Thus by induction our conjecture in proven. □

From now on, whenever an associative operation is used, brackets will not be used in the equation under the understanding that any insertation of brackets results in the same outcome. We will substitute brackets to emphasize parts of an equation, but this is only to aid the reader, and not for any reasons of rigour.

We introduce some notation to ensure the elegance and conciseness of our writing. Given a positive integer n, we can describe the exponential of the assignment in the following way:

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

One can easily create a recursive definition to be more rigorous, defined by

$$a^1 = a$$

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

Two useful lemmas, though obvious, will be used inherently in the proofs that follow later on in the book.

Lemma 2.0.2. *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 arbitrary. We prove by induction on 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 $m = r$.

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

□

Lemma 2.0.3. *For any two 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 n. For $m = 1$,

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

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

$$(a^n)^r = (a^n)^{r-1}a^n = \underbrace{a^{n(r-1)}a^n = a^{n(r-1)+n}}_{\text{Following from Lemma 2.0.2}} = a^{nr}$$

□

Definition 3. *Another property of an assignment on a set is* **commutivity***: that for any elements a and b in the set, $ab = ba$.*

The power of commutativity is that, given an associative and commmutative operation, we can permute any elements in an equation. Rigorously, we are talking about the following theorem.

Theorem 2.0.4 (Commutative Permutation Equivalence)**.** *For any finite sequence of elements $(x_1, x_2, \ldots, x_n)$ from a set upon which an 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 in 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 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.

$$\prod_{k=1}^{n} x_k = \big(\prod_{k=1}^{m-1} x_k\big)\big(x_m \prod_{k=m+1}^{n} x_k\big) \tag{2.5}$$

$$= \Big(\big(\prod_{k=1}^{m-1} x_k\big)\big(\prod_{k=m+1}^{n} x_k\big)\Big)\, x_m \tag{2.6}$$

We transition from (1.5) to (1.6) by use of both the associativity and commutativity property of the assignment. Now define a new permutation φ on the numbers from 1 to n by

$$\varphi(x) = \begin{cases} x & x \leqslant m-1 \\ x-1 & x \geqslant m+1 \\ n & x = m \end{cases}$$

What follows is that

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

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

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

By induction and the fact that $(\pi \circ \varphi^{-1})(n) = m$, it follows that

$$\begin{aligned} \big(\prod_{k=1}^{n-1} x_{\varphi(k)}\big)\ x_m &= \big(\prod_{k=1}^{n-1} x_{(\pi\circ\varphi^{-1}\circ\varphi)(k)}\big)\ x_{\pi(n)} && (2.7) \\ &= \big(\prod_{k=1}^{n-1} x_{\pi(k)}\big)\ x_{\pi(n)} && (2.8) \\ &= \prod_{k=1}^{n} x_{\pi(k)} && (2.9) \end{aligned}$$

We have shown that arbitrary permutations of elements are equal. □

Commutivity does not always hold in all assignments we work with in this book. It is only if + is used as an operation symbol, it is assumed the operation is commutative. All other assignments will not have this property unless specifically stated.

A subtle 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, the number 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 can generalize this concept to arbitrary elements in the following way.

Definition 4. *An* **identity** *of a set with an assignment is an element that is idempotent with all other elements in the group. We commonly call this element e, and its properties can be written by the statement that $ae = ea = a$ for all elements a.*

Lemma 2.0.5. *If an assignment has an identity, it is unique*

Proof. Suppose an assignment has two identites, denoted e and e'. Then $ee' = e$ by definition of the identity on e, but also $ee' = e'$ by the definition of identity on e'. By transitivity of equality, $e = e'$. □

If $\cdot$ is used as the operation's symbol, we may write e as 1, and if $+$ is used, we may write the identity as 0, even though the element is not always a number – the symbol 1 and 0 just becomes a metaphor to help us think about the identity with more abstract operations.

If a set has an identity, we define, for any element a, $a^0 = e$, extending the definition of the exponential. The properties previously proved still hold. this is because 0 is idempotent in addition,

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

And zeroes out multiplication,

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

Another subtle 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. Let us extend this notion to an arbitrary operation.

Definition 5. *Given a set with an identity e, we say an element a is* **invertible** *if there is another element b such that $ab = ba = e$. b is normally denoted a^{-1} and called the* **inverse** *of a.*

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.

Here are some properties common to all inverses. We assume associativity, but not commutivity in the operation used.

Lemma 2.0.6 (Length Inverse Equivalency Lemma)**.** *Let l and r and a be arbitrary elements of a set. If $la = e$ and $ar = e$, then $l = r$, and a is invertible*

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

Lemma 2.0.7 (Uniqueness of inverses)**.** *a^{-1} is unique*

Proof. The above property shows any two inverses are the same, for if x and y are two inverses, substitute x for l and y for r in Lemma (1.1.3). □

We now understand operations enough to describe the class of objects we will study for the rest of the book, a group.

Definition 6. *A* **group** *is a set with an associative operation that contains a unit element (such that the group is non-empty), and such that every element has an inverse. We say the group is abelian if its operation is commutative as well as associative.*

Definition 7. *The* **order** *of a group is the cardinality of the set that the operation operates on.*

If $n < 0$, define $a^n = (a^{-1})^n$. Again, the properties proved hold, but we can now extend exponentiation to any integer. Some easy 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 2.0.8. *The inverse of an inverse of an element a is a. That is, $(a^{-1})^{-1} = a$.*

Proof. The proof follows as $aa^{-1} = a^{-1}a = e$ and (2.0.7). □

Lemma 2.0.9. *$(ab)^{-1} = b^{-1}a^{-1}$.*

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

Lemma 2.0.10. *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 $x = x'$. □

We require inverses for a group, but we can weaken this claim only requiring left inverses, that is, for an element a, an element b where it is true $ba = e$, but it is not necessarily true that $ab = e$.

Theorem 2.0.11. *Any set with an associative operation and an identity such that every element has a left inverse contains arbitrary inverses for an elemenent, such that the set is a group.*

Proof. Let G be a set 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 1.1.3. But that also means a has a right inverse when substituted for c, so b is the inverse of a, and since a was arbitrary, all elements are invertible and thus G is a group. □

There are also many examples of groups. We list some interesting ones below. Note you do not need to understand all of these examples until they are mentioned later on in the book. You can just use this page as a reference for the groups we mention. 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 $S_{|X|}$. Note that the order of $S_{|X|}$ is $|X|!$ 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 operations $\circ$ 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). Each symmetry on the polyhedron, can be considered a rotational symmmetry or a reflection symmetry. We have n of each kinds of these symmetries, so the group formed by taking compositions of symmetries to form more symmetries forms the Dihedral group D_n of order $2n$. Dihedral groups were what Felix Klein was studying in transformational geometry.

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

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

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

  Quaternions are commonly used to represent three dimensional space in computer graphics.

- The Klein-4 Group or Viergruppe is a group of order 4 (denote elements $1, a, b, c$ and d) with the multiplication table:

	1	a	b	c
1	1	a	b	c
a	a	1	c	b
b	b	c	1	a
c	c	b	a	1

Here is an interesting group that may help with understanding the definitions we have created.

Theorem 2.0.12. *The homophonic group is defined as the group consisting of sequences of characters, where two sequences are equal if they are pronounced the same. What does the homophonic group look like? For instance, since sea = see, we can multiply on the left by $(se)^{-1}$ to obtain that $a = e$.*

Chapter 3

Subgroups

We can understand a machine by the various components in contains, thus understand the whole from the parts that make up the whole. Likewise, we can understand a group by the various components that it contains. Most of this book will be attempting to define these components – they are no so evident as the gears of a car or the wheel that makes it turn.

Definition 8. *A* **subgroup** *is a subset of a group that contains the identity, is closed under the operation which defines that group, and contains inverses for any element in the subgroup – in other words, a subgroup is a group that is a subset of a bigger group. If a and b are any elements in the subgroup, ab is in the subgroup as well. as well as a^{-1} and b^{-1}.*

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. This follows as the determinant operation has the multiplicative property, so that if $\det(X) = \det(Y) = 1$, it follows that $\det(XY) = \det(X)\det(Y) = 1$

- 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 $S_{|M|}$, and is equivalent to $S_{|M|-|N|}$.

- Given a group G, G and the set containing the identity are both subgroups. We call these trivial subgroups for self evident reasons, and say that any other group is non-trivial.

- The intersection of a family of subgroups of some group is also a subgroup.

It may be unexpected, but we can verify subgroups based on a single statement, specified in the following lemma. We leave it to the reader to verify the statement.

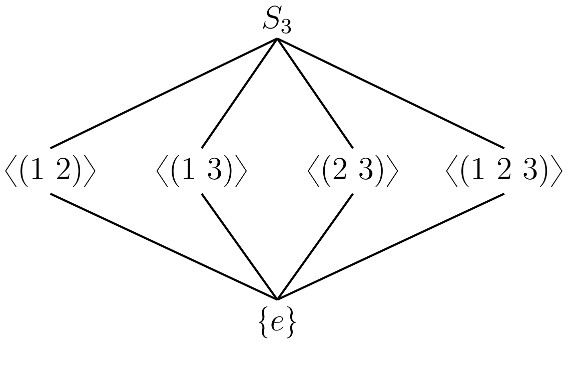
Lemma 3.0.13. *A non-empty subset H of a group G is a subgroup if and only if, for any elements a and b in H, ab^{-1} is in H.*

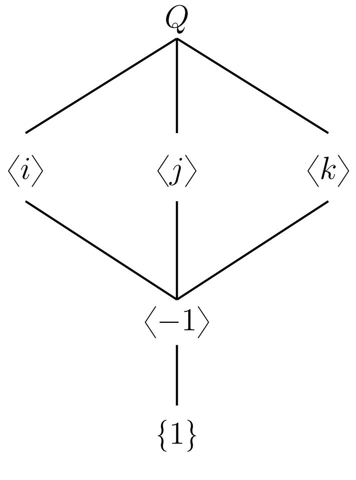
In our definition of a subgroup H of a group G we also assume that the identity is in a subgroup. However, it could be true that there is a different element e' such that e' acts idempotently on all elements in H, but not necessarily on all of G, and thus becomes a second identity. We show this cannot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

Theorem 3.0.14. *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 least subgroup that contains the two subgroups. This becomes very important in the context of Galois theory. We draw the lattices for S_3 and Q below.





Chapter 4

An Excursion: Subgroups of $\mathbf{Z}^+$

We have built a complicated tower of definitions for us to comprehend without extensive use of examples. This aside is intended show the power of the concepts developed and to provide concrete usage of the abstractions developed. We will consider the additive integer group $\mathbf{Z}^+$. 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 $a \circ M = \{a\} \circ M$.

Now for any integer a, one can verify that $a\mathbf{Z}^+$ forms a subgroup of the integers. What is suprising is the following.

Theorem 4.0.15. *G is a subgroup of $\mathbf{Z}^+$ if and only if G is of the form $a\mathbf{Z}^+$ for some integer a.*

Proof. Let G be a subgroup of $\mathbf{Z}^+$. If $G = \{0\}$, then $G = 0\mathbf{Z}^+$. Otherwise, G has some other non-zero element a. Thus G contains a positive element, for a is negative or positive, and if it is negative, then its inverse must be positive and contained in the subgroup by the closure property of a subgroup. By the well-ordering principle, G contains a smallest positive element b. Using euclidean division, every element $c \in G$ is of the form $mb+n$, where $0 < n < b$. Now $n \in G$, as $n = c - mb$, so we must conclude $n = 0$, as it cannot be a smaller positive integer than b. Thus every integer in G is divisible by b, and every number divisible by b is in G, so we conclude $G = b\mathbf{Z}^+$. □

Here are some common uses of these results 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 $c\mathbf{Z}^+$ for some 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 $\operatorname{lcm}(a, b)$.

Though we have focused only on $\mathbf{Z}^+$ in this aside, we note that the proofs can be applied to all cyclic groups, a concept we will define in the next chapter.

Chapter 5

Generators

We begin with an interesting discovery on collections of subgroups.

Lemma 5.0.16. *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

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 9. *Given a subset S of a group G, we define the subgroup generated by S to be the smallest subgroup that contains S, and we denote the subgroup $\langle S \rangle$. We call S the* **generator** *of the group $\langle S \rangle$. In the notation defined above this definition,*

$$\langle S \rangle = \bigcap_{j \in \mathcal{J}} \mathcal{M}$$

Equivalently, the generated subgroup is the set of all elements of the form $x_1 x_2 \ldots x_n$ where x_i or x_i^{-1} is in S. Any subgroup which contains S must contain these elements, and these elements themselves from a subgroup of G. We write this subgroup as $\langle S \rangle$, and if S is a finite group ordered by the sequence $\{x_1, x_2, \ldots, x_n\}$, we also write the subgroup as $\langle x_1, x_2, \ldots, x_n \rangle$.

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. It can be shown that this can be reduced to the set of all elementary matrices of the first and third kind.

If a group is generated by a single element, then the group is called cyclic. One example is $\mathbf{Z}^+$. Let g be an element of a group G, and suppose that $\langle g \rangle$ is order c for some natural number c. Then the following properties hold for g:

Lemma 5.0.17. *$\{g, g^2, \ldots, g^c\}$ are all distinct elements of g.*

Proof. If $g^i = g^j$, for $0 < i < j \leqslant c$, then $g^{j-i} = e$. Then the smallest subgroup that contains g is $\{g, g^2, \ldots g^{i-j}\}$, hence the order of g is $j - i < c$. By contradiction, $g^i \neq g^j$. □

Corallary 5.0.18. *For $k < c$, $g^k \neq e$*

Proof. For then $g^{k+1} = g$. □

Corallary 5.0.19. *$g^c = e$.*

Proof. g^c is not any number from g to g^{c-1}, so it must be the element of the group generated by g that is different from the other elements before it. Thus $g^c = e$ by corallary (2.2.3), as no other element before g^c is e. □

Lemma 5.0.20. *$g^k = e$ if, and only if, $c|k$*

Proof. This is left to the reader. It is a simple application of euclidean division used previously in our discussion of $\mathbf{Z}^+$. □

Lemma 5.0.21. *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 $i - j$. □

We have also shown that in $\mathbf{Z}^+$ that every subgroup is cyclic, but this proof can be easily extended to the following: every subgroup of a cyclic group is cyclic. We leave it to the reader to use the techniques of the excursion to establish this.

Given an element g in an arbitrary group G, we define the order of g to be the order 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 5.0.22. *The order of an element ab is the same as the order of an element ba.*

Proof. Consider the group $\langle ab\rangle$. $(ba)^{-1} = a^{-1}b^{-1}$. Suppose the order of ab is finite, of order k. Then $(ab)^k = e$, which means that $b(ab)^k = b$, and $b(ab)^k = (ba)^k b$, so that $(ba)^k b = b$, and thus $(ba)^k = e$. 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 are equal. □

Chapter 6

Homomorphisms

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

Definition 10. *Let G be a group with operation $\circ$ and H a group with operation $*$. A* **homomorphism** *between G and H is a function f such that for any elements x and y. $f(x \circ y) = f(x) * f(y)$. We say that G and H are homomorphic. If a homomorphism is bijective, we call it an isomorphism. If $G = H$, we call a homomorphism an endomorphism, and an isomorphism an automorphism.*

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. 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.

Definition 11. *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 6.0.23. $\varphi(e) = e$

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

Lemma 6.0.24. $\varphi(a^{-1}) = \varphi(a)^{-1}$.

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

Lemma 6.0.25. *The kernel of a homomorphism is a subgroup*

Proof. If $\varphi(a) = e$ and $\varphi(b) = e$, $\varphi(a^{-1}) = \varphi(a)^{-1} = e^{-1} = e$ so that $\varphi(a^{-1}b) = e$. □

Lemma 6.0.26. *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 $\varphi(x^{-1}y) = a^{-1}b$. □

Lemma 6.0.27. *A homomorphism is injective if and only if $\varphi(a) = e$ implies $a = e$.*

Proof. We prove both implications.
($\Rightarrow$) If a homomorphism is injective, and if $\varphi(a) = e$, then $a = e$ as $\varphi(e) = e$.
($\Leftarrow$) 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 $\mathbf{R}^\times$

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

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

Chapter 7

Cosets

We were previously introduced to subgroups as an attempt to understand the group as a sum of its parts. Cosets are another way to do this.

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**.

Lemma 7.0.28. *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 in the obvious way, by the equivalence relation $g\ k$ if $g \in Hk$. Like left cosets, all right cosets can be written Hg for some 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 7.0.29. *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. □

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 formidible, the mechanics we have built up make the proof relatively simple – our definitions were the hard part to understand.

Theorem 7.0.30 (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 gH to $g'H$ defined by $a \mapsto g'g^{-1}a$. 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 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 theorem as a corallary, $|G| = |H|(G : H)$, so $|H| \mid |G|$. □

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.

Corallary 7.0.31 (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, $|G| = |M|(G : M)$. 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 harken back to the beginning of the book. We said that Euler used the ideas of Lagrange to prove his theorem of the totient function (Look back to the first chapter if you don't remember). We now have the power to prove this as a corallary.

Corallary 7.0.32 (Euler's Theorem)**.** *For any two relatively prime integers a and b,*

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

Proof. Consider the set P of relatively prime integers to b, with the operation of multiplication. We show that, modulo b, this set forms a group. Let x be a number relatively prime to b. Then there exists integers m and n such that $mb + nx = 1$, so that $mb \equiv 1 \mod b$. Furthermore, if for some number x there exists a number y such that $xy \equiv 1 \mod b$, then $xy + mb = 1$ for some number m. It follows that the greatest common denominator of x and b must be 1, so the two are relatively prime. Finally, if x and y are relatively prime to b, so is xy, because the two numbers share no prime factors. Thus we conclude that the elements in P modulo b, denoted $P/b\mathbf{Z}^\times$ (cosets of the multiplicative group $b\mathbf{Z}^\times$) form a group under multiplication. a modulo b is a member of this group. Take the subgroup $\langle a \rangle$, which must be a finite order, which we denote by k. Since the set contains $\varphi(b)$ elements, we have that $k \mid \varphi(b)$ by Lagrange's theorem. But then $a^{\varphi(b)} \equiv 1 \mod b$ by (5.0.20). □

Chapter 8

Normal Subgroups

Theorem 8.0.33 (Characterization of Normal Subgroups)**.** *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 $H \lhd G$:*

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. □

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

Theorem 8.0.34. *The kernel of a homomorphism is a normal subgroup*

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 (1) in the definition. □

The trivial group is always normal, because given any element g, $g^{-1}eg = g^{-1}g = e$. Furthermore, for any group G, $G \lhd G$ (we leave this to the reader to verify). Thus no group posesses that characteristic that it contains no normal subgroups. Thus we must define a property along these lines in a slightly more precise way. 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, $I \lhd G$.

Chapter 9

Isomorphism Theorems

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 $M \circ N = MN$. As $M = gH$ and $N = g'H$ for some $g, g' \in H$, $MN = gHg'H = gg'HH = gg'H$. Thus the operation is closed in G/H, and G/H forms another group: the product or factor 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 (note this shows that any normal subgroup is the kernel of some homomorphism). The projection of G onto G/H has a property that we prove in a more general form.

Theorem 9.0.35 (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. □

$$\begin{tikzcd}
G \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & H \\
& G/H \arrow[ur, "\overline{\varphi}"'] &
\end{tikzcd}$$

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.

A simple application of the first isomorphism theorem is a way to completely understand the cyclic groups, a classification. The relation on groups $x \sim y$ if $x \equiv y$. In algebra, we refer to a classification of groups as a set of isomorphism classes for those groups. It is the ultimate achievement as it allows us to show that these are the only way the operations can perform.

Theorem 9.0.36 (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 $\{0\}$, and $\mathbf{Z}^+/0\mathbf{Z}^+ \cong \mathbf{Z}^+$, so $\langle g \rangle \cong \mathbf{Z}^+$. □

We have a trilogy of isomorphism theorems to show. Here is the second.

$$\begin{tikzcd}
 & G \arrow[d] & \\
 & HN \arrow[dl] \arrow[dr] & \\
H \arrow[dr] & & N \arrow[dl] \\
 & H \cap N &
\end{tikzcd}$$

Theorem 9.0.37 (The Second Isomorphism Theorem)**.** *Let G be a group, and N and H subgroups such that $N \lhd G$. The NH is a subgroup of G, and $N \cap H \lhd G$. The assignment map $h(N \cap H) \mapsto hN$ is an isomorphism, and so $H/N \cap H \cong NH/H$:*

Proof. First we prove NH is a subgroup. If n_1h_1 and n_2h_2 are in NH, then $n_1h_1(n_2h_2)^{-1}$ is in NH by the following calculation:

$$\begin{aligned} n_1h_1(n_2h_2)^{-1} &= n_1h_1h_2^{-1}n_2^{-1} \\ &= n_1((h_1h_2^{-1})n_2^{-1}(h_1h_2^{-1})^{-1})h_1h_2^{-1} \end{aligned}$$

The equation above is in NH as N is normal, so that $(h_1h_2^{-1})n_2^{-1}(h_1h_2^{-1})^{-1}$ is in N. The map $h \mapsto hN$ is a surjective homomorphism from H to NH/N, and the kernel is $N \cap H$ (so $N \cap H$ is normal), and $H/N \cap H \cong NH/N$ by the first isomorphism theorem. □

The final isomorphism theorem is the following.

Theorem 9.0.38 (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 $(G/N)/(M/N) \cong G/M$. □

Surprise, turns out we have another isomorphism theorem.

Theorem 9.0.39 (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 such that $f(H)$ is 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 $H \mapsto H/N$. The properties above can (and should be) be checked by the reader. □

Chapter 10

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 just a trivial example. 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 10.0.40 (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 $S_{|G|}$. □

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.

Definition 13. *Given a set M and a permutation π on M, the* **support** *of π, denoted $\sup(\pi)$, is equal to $\{m \in M : \pi(m) \neq m\}$. A cycle of length k is a permutation π such that $|\sup(\pi)| = k$, and such that we can order $\sup(\pi)$ to be $(x_1, x_2, \ldots, x_k)$ in a way that $\pi(x_n) = x_{n+1 \mod k+1}$. A cycle of length two is called a transposition.*

We write the function π as $(x_1, \ldots, x_k)$. 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 10.0.41. *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. We prove this theorem by induction. For a one element set, the proof is vacously true. Now suppose that permutations on sets of less than n elements can be written as disjoint cycles. Consider a set X of n elements. Take an arbitrary element x_1. Then form a sequence $(x_1, \pi(x_1), \pi^2(x_1), \ldots, \pi^n(x_1)$. Such that π^{n+1} is the least function for which $\pi^{n+1}(x_1) = x_1$. This is always possible in a finite set by an easy application of the pidgeonhole principle. Denote these elements as the sequence $(x_1, x_2, \ldots, x_n)$. Then $\pi|_{\{x_1,\ldots,x_n\}}$ is a cycle of length n. Now $X - \{x_1, \ldots, x_n\}$ is a set of less than n elements, so that we can decompose it $\pi|\{x_1, \ldots, x_n\}$ into a product of disjoint cycles. Then, since the support of each is disjoint, $\pi|_{\{x_1,\ldots,x_n\}} \circ \pi|\{x_1, \ldots, x_n\} = \pi$, and π is a product of disjoint cycles. □

Lemma 10.0.42. *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 $\pi(x) \notin \sup(\pi\sigma\pi^{-1})$. □

Lemma 10.0.43. *Any symmetric group is generated by the set of transpositions in that group.*

Proof. Each cycle can be decomposed into transpositions. Concretely, this is the fact that

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

We have shown that S_X is generated by the set of cycles, hence S_X is generated by the set of transpositions. □

The parity or signum of a permutation π is based on the number of transpositions that is is composed of. If the number is even the value is 1, else it is -1. We show that it this definition is well defined.

Lemma 10.0.44. *Any two compositions of transpositions that are equal either both have an even number of transpositions or both have an odd number.*

Proof. Let $(\sigma_1\ \sigma_2 \dots \sigma_n) = (\pi_1\ \pi_2 \dots \pi_m)$, where each composition is a sequence of transpositions. Every transposition can be written as a product of odd number of transpositions of adjacent elements, i.e

$$(n\ m) = (n\ n+1)(n+1\ n+2)\dots(m-1\ m)(m-1\ m-2)\dots(n+1\ n)$$

Now decompose each transposition into a product of adjacent elements like above. We get a new sequence, that $(\sigma'_1\ \sigma'_2 \dots \sigma'_{n'}) = (\pi'_1\ \pi'_2 \dots \pi'_{m'})$. Now if we multiply from left to right the new sequence of the left with the old sequence, we obtain the fact that $n - n'$ and $m - m'$ are even. If we do this on the left of the right sequence with the left, we get that $m' - n'$ is even. Hence $m' - n$ is even, and so $m - m'$ is even, which is what we wanted to show. □

The parity operation defines a homomorphism from S_n into $\mathbf{Z}^\times$. The kernel of this is A_n, the alternating group, a normal subgroup of S_n. Here are some properties of A_n.

Lemma 10.0.45. *If τ is a transposition, $S_n = A_n \cup \tau A_n$. Thus $A_n = n!/2$.*

Lemma 10.0.46. *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 10.0.47. *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 11

Group Actions

Automorphisms are symmetries on a group, which, when taken together, form a group as a whole. The symmetry group acts as a more general notion of symmetries of a set. Through these actions, we obtain large amounts of information about both the group and the set. These are specific notions of a more general structure that we now describe, a group action.

Definition 14. *A* **group action** *on a set G and set X is a homomorphism from G to $S_{|X|}$. As each $g \in G$ has an associated permutation, for $x \in X$ we write gs for the permutation associated with g acting on s. We call X a* **G-set***.*

It is simple to show $g(hx) = (gh)x$ and $ex = x$, for $g, h, e \in G$ and $x \in X$. Another way of saying the first part of the statement is that, if φ is the homomorphism defining the action, $\varphi(g) \circ \varphi(h) = \varphi(gh)$, and the section is saying that $\varphi(e) = \mathbf{1}$. This is just a definition of a homomorphism from a group to the symmetric group, so is van equivalent way of defining a group action.

Definition 15. *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 $g \in G$.*

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

Definition 16. *A G-set is transitive if it has just one orbit.*

This just means that for any two elements x and y in a G-set X, there is some $g \in G$ such that $gx = y$.

Definition 17. *An action is faithful if the homomorphism defining it is injective.*

Definition 18. *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 an isomorphism if it is bijective.*

To relate this concept back to concepts previously talked about, a G-morphism is sort of like an isomorphism between G-sets. G-sets have no operation defined to compose them, however, so we must use the group action to define the isomorphism.

Definition 19. *An element x in a G-set X is a* **fixed point** *if $gx = x$ for every $g \in G$. The set of all fixed points is denoted X^G.*

Definition 20. *Given any $x \in X$, $G_x = \{g \in G : gx = x\}$ is a subgroup called the isotropy subgroup of x in G.*

As an example, let G act on itself by conjugation, that is, $g(h) = ghg^{-1}$. 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.

As another example, consider conjugation from G on its subgroups. Then the isotropy group of a subgroup H is the normalizer $N_G(H)$, which is the set $\{g \in G : gHg^{-1} = H\}$. The fixed points of this transformation are precisely the normal subgroups.

Consider the group $SL_n(\mathbf{R})$ acting on the upper half of the complex plane, that is, the set $\{z \in \mathbf{C} : \mathrm{im}(z) > 0\}$ by the mobius transform below:

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

The isotropy subgroup of i is the special orthogonal group $SO(2)$, the set of matrices with orthonormal columns. The mobius transform is transitive. 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.

Chapter 12

Direct Products, Semiproducts, and Abelian Groups

Let I be an index set, and $\{G_i\}_{i \in I}$ a family of groups. Then the direct product of $\{G_i\}$, denoted $\bigtimes_{i \in I} G_i$, is a group with operation $\bigtimes_{i \in I} g_i \circ \bigtimes_{i \in I} h_i = \bigtimes_{i \in I} g_i h_i$. The group is called the product group.

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 $y^s = e$. This follows as $(x, y)^{rs} = (x^{rs}, y^{rs}) = (e, e)$. If $(x, y)^m = (x^m, y^m) = (e, e)$, $r|m$ and $s|m$, so $rs|m$. □

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 $(h, k) \mapsto hk$. The map is bijective, as $HK = G$, and if $hk = e$, $h = k^{-1}$, so $k^{-1} \in H$, so $k = h = e$. $hkh^{-1} \in K$, as K is normal, but it is also in N as N is normal, hence $hkh^{-1}k^{-1} = e$, so $hk = kh$, and thus the map is a homomorphism as $h_1 k_1 h_2 k_2 = h_1 h_2 k_1 k_2$. □