Group Theory

Jacob Denson

September 10, 2018

Table Of Contents

Chapter 1

What is Abstract Algebra?

In mathematics, one often focuses on particular properties of a certain family of objects, which lie in an ambient space. In order to understand this property, it is natural to consider transformations on this space which preserve the properties under consideration.

Example. *In Euclidean Geometry, rotations and translations preserve the angles between lines and distances between points. If we only care about such properties, we can translate and rotate any particular figure to a canonical form to simplify the situation; a triangle can always be rotated and translated to one side lies horizontally, and if we don't care about distances, we can dilate space so that one line of the triangle has length one.*

The technique of applying symmetries to simplify situations occurs over and over in mathematics, and so it is important to classify the general tools we can use when we meet new objects, and wish to understand their symmetries. The general study of symmetries in mathematics is known as *group theory*. Due to its utility, the theory provides a foundation to many rich mathematical theories.

Let us reconsider symmetries from a more abstract viewpoint. One of the basic objects of mathematics is the function, which transforms elements of some set A into elements of another set B. If the function is denoted f, which we often introduce using the abbreviated notation $f : A \to B$, then the b associated to an a is denoted $f(a)$. Given another map $g : B \to C$, we may consider the **composition map** $g \circ f : A \to C$, which maps a point $a \in A$ to $g(f(a))$ – that is, if a is mapped to b by f, and g maps b to c then $g \circ f$ maps a to c. A pleasant algebraic fact about

the composition is that it satisfies the **associative property**. Given a third map $h : C \to D$, we find that $h \circ (g \circ f) = (h \circ g) \circ f$, a relation taken for granted when we forget parenthesis and write $h \circ g \circ f$. The first idea leading to abstract algebra is that we can identify a *functional definition* of the identity map with an *algebraic definition* involving a series of algebraic relations with respect to composition. A key idea of group theory is that we can study the functional properties of symmetries by looking at the compositional properties of maps without losing essential information.

Example. *On each set B we have an* **identity map** $id_B : B \to B$*, such that $id(b) = b$ for each $b \in B$. For any $g : A \to B$ and $h : B \to C$, we find $id_B \circ g = g$, and $h \circ id_B = h$. If $f : B \to B$ is any map satisfying $f \circ g = g$ and $h \circ f$ for any g and h, then f is equal to the identity map, since $f = f \circ id_B = id_B$. Thus an 'identity map' is just an idempotent element with respect to composition.*

Example. *If a function $f : A \to B$ is bijective, then there is $f^{-1} : B \to A$, defined by mapping an element b to the unique element a with $f(a) = b$. We find that $f^{-1} \circ f = id_A$, and $f \circ f^{-1} = id_B$. If g is any function such that $g \circ f = id_A$ and $f \circ g = id_B$, then*

$$g = g \circ id_B = g \circ f \circ f^{-1} = id_A \circ f^{-1} = f^{-1}$$

Thus the inverse of a map is precisely one which composes with the map to give the identity map.

Again, we see that functions can be identified by algebraic relations with respect to the composition operator. Abstract algebra is the mathematical field whose goal is to study mathematical objects via an understanding of the algebraic relations of operations on that set, with the hope that less obvious properties of the object will be unveiled via the underlying algebraic properties. In the case of the theory of functions, the operator studied is composition. In the theory of classical algebra, the operators studied are addition, multiplication, subtraction, and division. The key realization of abstract algebra is that it is often more simple to discuss arbitrary, 'abstract operators' satisfying certain properties, for then we need not deal with the minutiae which occurs which studying the set theoretic aspects of functions. In these notes, we talk about a specific class of objects and operators, which generalize a set of invertible functions from a set to itself. These objects are known as groups.

Let us consider what properties the class of functions under composition should satisfy. Let X be a set, and let $\circ : X \times X \to X$ be an abstract 'composition function' on X. This means exactly that, given two objects $x, y \in X$, we may consider their composition $x \circ y \in X$. Now assume that $\circ$ satisfies the associative law $x \circ (y \circ z) = (x \circ y) \circ z$ for any three $x, y, z \in X$ (this fact is no longer for free, because our composition operation isn't normal function composition, since elements of X do not have to be functions). An 'identity' in X can then be defined to be an element $e \in X$ such that $e \circ x = x \circ e = x$ for all $x \in X$. We may then define an 'inverse' of an element $x \in X$ to be an element $y \in Y$ such that $x \circ y = y \circ x = e$. The element y is rarely denoted by anything other than x^{-1}, to parallel the set theoretic notation. Thus if $\circ$ is associative and the underlying set has an identity, then the resulting pair $(X, \circ)$ imitates a subset of functions from a set to itself, which is closed under composition. We call the pair $(X, \circ)$ a **monoid**. If every element of X is invertible, then $(X, \circ)$ imitates a set of invertible functions from a set to itself, closed under inversion, and we call this pair a **group**. Since symmetries can often be described as families of invertible functions, group theory describes the tools to understand these families. If the operation is obvious, we often abuse notation and just say that X is a group. To be even more brief, the symbol for the operation is often ignored as well, so we write xy for the composition $x \circ y$ of two elements.

Example. *Consider a topological space X. Then the set of all continuous functions from X to itself forms a monoid, and the set of all homeomorphisms from X to itself forms a group. This follows directly because if f and g are continuous, then $g \circ f$ is continuous, and the identity function is certainly continuous. If the space has a fixed metric, then the space of all isometries of the space forms a group as well.*

Example. *The set of all linear maps from a vector space V to itself forms a monoid. The set $\mathrm{GL}(V)$ of all invertible linear maps froms a group, known as the* **general linear group**. *If V has finite-dimension n, then we may essentially identify linear endomorphisms on V with the set of all $n \times n$ matrices $M_n(K)$ with entries in the scalar field K upon which V is defined, which can be*

viewed as a set with the abstract composition operation

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}$$
$$= \begin{pmatrix} a_{11}b_{11} + \cdots + a_{1n}b_{n1} & \dots & a_{11}b_{1n} + \cdots + a_{1n}b_{nn} \\ \vdots & \ddots & \vdots \\ a_{n1}b_{11} + \cdots + a_{nn}b_{n1} & \dots & a_{n1}b_{1n} + \cdots + a_{nn}b_{nn} \end{pmatrix}$$

Then the matrix I with ones on the diagonal operates as an identity, and $\mathrm{GL}(V)$ *can be identified with the subfamily of matrices*

$$\mathrm{GL}_n(K) = \{M \in M_n(K) : MN = NM = I \text{ for some } N \in M_n(K)\}$$

also called the general linear group.

Example. *Certain vector fields $X : U \to \mathbf{R}^n$ on open subsets of Euclidean space induce 'one parameter groups' $\phi : U \times \mathbf{R} \to U$ (where the image of (x,t) is denoted $\phi_t(x)$, which we can view as a 'parameterized' family of maps), which satisfy the differential equation*

$$\frac{d\phi_t(x)}{dt} = X_{\phi_t(x)}$$

and also satisfy $\phi_t \circ \phi_s = \phi_{t+s}$, and $\phi_0 = id_U$, so this set of functions really is a group. The study of differential equations is really just the study of the relationship between smooth vector fields and the one-parameter groups of diffeomorphisms they generate.

Example. *Combinatorics and algebra intertwine when we study finite groups. The classical finite group is the class of all bijective maps from a set containing n elements to itself. This is the* **symmetric group** *of order n, denoted S_n. The group contains $n!$ elements, because an arbitrary permutation $\pi : [n] \to [n]$ can be obtained by first choosing $\pi(1) \in [n]$ (for which we have n choices), then choosing $\pi(2)$ (of which we have $n-1$ choices, since we cannot have $\pi(2) = \pi(1)$), and so on and so forth. More generally, if X is a set, we can consider the group of bijections on X, denoted $Sym(X)$. Using cycle notation, we denote the permutation π satisfying $\pi(a_1) = a_2$, $\pi(a_2) = a_3, \dots, \pi(a_m) = a_1$, and fixing all other elements by $(a_1\ a_2\ \dots\ a_m)$. We shall find that all permutations on a finite set have a unique cycle decomposition, when we discuss the symmetric group in more detail later.*

Example. *We can often form groups by abstractly defining relations between objects in a set. For instance, consider the set of 8 objects* $Q = \{\pm 1, \pm i, \pm j, \pm k\}$. *We then try and find a composition operation* $Q \times Q \to Q$ *which makes* Q *into a group, satisfying certain equations. In this case, we want* -1 *to act as we would expect by the notation, so that* $(-1)i = -i$, $(-1)(-i) = 1$, *and so on and so forth. Less trivially, we also want* $i^2 = j^2 = k^2 = ijk = -1$. *We obtain the multiplication table, whose row labelled* x *and column labelled* y *give the element* xy.

	i	j	k
i	-1	k	$-j$
j	$-k$	-1	i
k	j	$-i$	-1

One can check that the induced operation is associative. The group Q *is known as the* **quaternion** *group. It is a particular subset of the quaternions, which are expressions of the form* $a + bi + cj + dk$ *with* $a, b, c, d \in \mathbf{R}$, *and model the set of all rotations in three dimensional space. Thus, even though abstractly defined,* Q *can still be interpreted in a meaningful way as a symmetry on a group.*

Chapter 2

Basic Properties of Groups

Now we shall start the general theory of groups, starting with the theory of 'manipulating equations', of which every student of compulsory education should be very familiar. Consider a finite sequence of elements $x_1, \ldots, x_n$ in a monoid X. Then the 'pi' notation for multiplication is introduced, defined recursively by

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

It is a convention that if $k > i$, then $\prod_{j=k}^{i} x_j = e$. The similarity to Σ notation used in arithmetical sums is intentional, and the two definitions correspond in the monoid $(\mathbf{Z}, +)$. Ultimately, the property of associativity means brackets in an equation are irrelevant. That is, we need not introduce brackets to $x \circ y \circ z$ to determine the overall result. We prove this rigorously, and then dodge the use of brackets in the rest of these notes, except in emphasizing components of equations.

Theorem 2.1. *Let there be given an associative operation on S, and a finite sequence $(x_1, \ldots, x_n)$ of elements in S. Then, for any integer $1 \leqslant l < n$,*

$$\left(\prod_{k=1}^{l} x_k \right) \left(\prod_{k=l+1}^{n} x_k \right) = \prod_{k=1}^{n} x_k$$

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 obvious by definition.

We now proceed inductively. If we are now given n elements $(x_1,\ldots,x_n)$, and an integer $1 \leqslant l < n$, then

$$\left(\prod_{k=1}^{l} x_k\right)\left(\prod_{k=l+1}^{n} x_k\right) = \left(\prod_{k=1}^{l} x_k\right)\left(\left(\prod_{k=l+1}^{n-1} x_k\right) x_n\right)$$
$$= \left(\prod_{k=1}^{l} x_k \prod_{k=l+1}^{n-1} x_k\right) x_m = \left(\prod_{k=1}^{n-1} x_k\right) x_m = \prod_{k=1}^{n} x_k$$

By induction, this statement holds for all values of n. □

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 2.2. *For any finite sequence of elements $(x_1,\ldots,x_n)$ from a set upon which an associative and commutative assignment is defined, and for any permutation $\pi \in S_n$,*

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

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, which changes nothing. Now suppose, by induction that the statement is true for an arbitrary 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.

$$\prod_{k=1}^{n} x_k = \left(\prod_{k=1}^{m-1} x_k\right)\left(x_m \prod_{k=m+1}^{n} x_k\right) = \left(\left(\prod_{k=1}^{m-1} x_k\right)\left(\prod_{k=m+1}^{n} x_k\right)\right) x_m$$

The permutation which swaps the remaining $n-1$ elements really does only swap $n-1$ elements, hence by induction the equality is obtained, and we find that it is possible to reorder finite sequences of elements arbitrarily. □

So now we have seen proofs of facts intuitively obvious from a elementary school education. Before we start our real work, we should establish that groups are not that much more general than sets of functions. Arthur Cayley is credited with noticing that the synthetic definition really is the same as the intutive one, so that our algebraic relations uniquely model the theory of bijective functions.

Theorem 2.3 (Cayley's Theorem)**.** *Any synthetic group is equivalent to a group of functions.*

Proof. Let G be a synthetic group. For each $g \in G$, consider the function $g_* : G \to G$, defined by $g_*(h) = gh$. Then g_* is a bijective function, for it has an inverse $(g^{-1})_*$. The transformation $g \mapsto g_*$ 'preserves' the operation of the group, for $(gh)_* = g_* \circ h_*$ so $G_* = \{g_* : g \in G\}$ really is a group of functions, which is essentially the same group as G, for the operations are pretty much the same. We will later make these notions precise by saying G and G_* are *isomorphic* by this map. □

Thus we are back where we started. We have the abstract group theory at our tool belt, but when all is said and done, we really are just discussion groups of transformations. The formalism gives us abstract insight into these groups, but we aren't abstracting for an arbitrary reason, since every group can be considered as a concrete set of functions.

Corollary 2.4. *The identity of a group is unique, and the inverse of any element, is uniquely determined.*

Proof. We have showed these results hold for groups of functions. □

Theorem 2.5. *Let G be a semigroup, with an element e such that $eg = g$ for all $g \in G$, and suppose that every g has a left inverse h such that $hg = e$. Then G is a group, and e is the identity.*

Proof. Give g and h as in the theorem, we calculate $hgh = eh = h$, and if we find k such that $kh = e$, we find $gh = egh = khgh = keh = kh = e$. Thus h is also the right identity for g, hence a normal inverse. But this means that $ge = gee = geg^{-1}g = gg^{-1}g = g$, so e is a right inverse as well. □

2.1 Subgroups and Cosets

Often in math the 'symmetries' to choose from are not completely obvious, and as we range our symmetries to preserve an increasingly strict set of properties, the number of symmetries we have reduces to a smaller and smaller family. Thus from our group of symmetries we obtain a **subgroup**, a subset of a group whose elements also form a group. Even if we really do care about the entire group, the subgroups of the group enable us to understand what parts of the group are 'self contained', which enables us to understand the entire group by the components it contains.

Example. *Define the special linear group $SL_n(K)$ to be the subset of matrices in the general linear group $GL_n(K)$ with determinant one. The determinant operation satisfies*

$$\det(XY) = \det(X)\det(Y) \qquad \det(X^{-1}) = \det(X)^{-1}$$

which enables us to easily show $SL_n(K)$ is closed under composition and inversion. We will later see that this is a special case of forming a subgroup from the kernel of a homomorphism, which in this case is the map $\det : GL_n(K) \to K^*$.

Example. *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 $S_{|M|-|N|}$ as the elements that are in N can be ignored in the definition of the function.*

Example. *Any group is a subgroup of itself, as is the set $\{e\}$ solely containing the identity. We call these the trivial subgroups of a group.*

If we fix a group, then the subgroups of a group have a nice lattice structure obtained by considering $G < H$ if $G \subset H$, that we now elaborate. The next proposition shows that we can find a greatest lower bound to any set of subgroups of a group.

Proposition 2.6. *If $\{H_\alpha\}$ are subgroups of G, then $\bigcap H_\alpha$ is a subgroup of G.*

Proof. If a and b are in $\bigcap H_\alpha$, then they are in every group H_α, which means that ab and a^{-1} is in H_α for every H_α, hence they must also be in the intersection of these groups. □

Conversely, let G be a group, and S a subset of elements, we can consider the set $\mathcal{M}$ of all subgroups of G which contain S. Of course, $\mathcal{M}$ is non-empty, as G is a subgroup which contains S. If we take $\bigcap \mathcal{M}$, then we obtain a group containing S, which is contained in every group which contains S. This 'smallest' group is called the group **generated by** S, denoted $\langle S \rangle$. Equivalently, the generated subgroup is the set of all elements of the form $x_1 x_2 \dots x_n$ where either x_i or x_i^{-1} is in S. This is because this set forms a subgroup of G, and also every subgroup that contains S conversely must contain these elements. In this way, generators work for groups analogously to how bases work in vector spaces, which are formed by arbitrary sums of the generators.

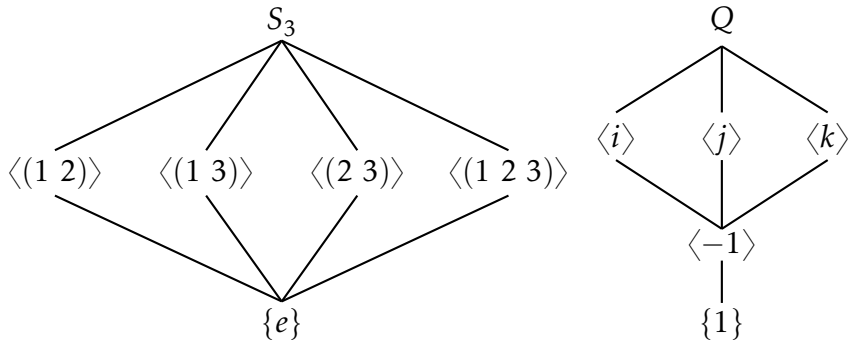
Example. *Gaussian elemination shows that every invertible matrix is the product of elementary matrices, so $GL_n(K)$ is generated by the elementary matrices.*

Example. *The integers form a group under addition, which is generated by* 1*, because every positive integer $n > 0$ can be written as*

$$n = 1 + 1 + \cdots + 1$$

and is thus in the group generated by 1*, and thus $-n$ is also in the group generated by* 1 *as well, hence all integers. A group generated by a single element is known as* **cyclic**.

The greatest lower bound of a set of subgroups of a group is the group generated by the union of the elements of a group. We have a maximal subgroup, which is the entire group, and a minimal subgroup, which is the trivial group $\{e\}$. We can often gain insight into groups by drawing the subgroup lattice, and these lattices are essential in the application of group theory to Galois theory. Here are the lattices for S_3, and the group of quaternions.



We can gain a deeper understanding of the relations between elements of a group, because a subgroup neatly contains all possible algebraic structure between the elements trapped in the subgroup. A natural question is how much we can obtain about the relations obtained by composing elements outside of a subgroup with elements inside the subgroup. We cannot hope to understand this question by studying the subgroup as an isolated object, so we must see the subgroup as part of the overall group. One tool for understanding this containment is by studying cosets, which break apart the group by its relations with elements of a subgroup.

Let $H < G$. Define an equivalence relation $\sim$ on G by $x \sim y$ if $x \in yH$. The equivalence classes formed by the relation are denoted G/H, pronounced as 'G modulo (mod) H'. Each equivalence class is called a **left coset**. Every coset can be expressed $gH = \{gh : h \in H\}$, for some $g \in G$. Think of cosets are subgroups that are translated around by an element in a group, like subspaces in a vector space shifted by a vector.

Remark. Right cosets can be 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. We denote the set of right cosets by $H\backslash G$. It doesn't really matter whether we talk about left or right cosets, because we have a natural map from one family to the other, mapping the coset gH to the coset Hg^{-1}. We choose to use left cosets.

The number of cosets in G/H or $H\backslash G$ is denoted $(G : H)$, and is called the **index** of H in G. We now come to one of the most fruitful 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.

Theorem 2.7 (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 k be arbitrary elements of G. Define a map from gH to kH, defined by $gh \mapsto kh$. This map is certainly surjective, and is also injective, for if $kh = kh'$, $h = h'$, so $gh = gh'$. It follows all cosets have the same cardinality. 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, $|G| = |gH|(G : H)$. By noting that H is a coset (simply take the coset of e), we obtain that $|G| = |H|(G : H)$, hence $|H|$ divides $|G|$. □

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.

Corollary 2.8. *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 2.9 (The Multiplicative Property)**.** *Let G be a group, H a subgroup of G, and L a subgroup of H. Then $(G : L) = (G : H)(H : L)$.*

Proof. If L is a subgroup of H, then $|H| = |L|(H : L)$, and thus $|G| = |H|(G : H) = |L|(G : H)(H : L)$. Noticing that L is also a subgroup of G, we have $|G| = |L|(G : L)$. Thus we conclude $|L|(G : L) = |L|(G : H)(H : L)$. By dividing out $|L|$, we obtain the equation. □

We now have the power to prove another interesting number theoretic statement, known as Euler's theorem. Consider the totient function φ, which takes an integer n and gives us the number of integers relatively prime to n, which are less than n. The theorem is simple with the power of the methods we now possess.

Corollary 2.10. *For any relatively prime n, m, $n^{\varphi(m)} \equiv 1 \pmod{m}$.*

Proof. Consider the group $(\mathbf{Z}/m\mathbf{Z})^\times$, which consists of all invertible elements of $\mathbf{Z}$ modulo m. We claim the size of this group is $\varphi(m)$, by showing that a necessary and sufficient property for inclusion in the group is being congruent modulo m to a relatively prime positive integer less than or equal to m. To show that such a number must be relatively prime, we use Bezout's theorem. If x is invertible in the group, this means there exists y such that $xy \equiv 1 \pmod{m}$, for some y, which means we can write $xy + km = 1$, for some integer k. But this means $\gcd(x, m) = 1$. Thus $(\mathbf{Z}/m\mathbf{Z})^\times$ has order $\varphi(m)$. Now if n is an arbitrary integer relatively prime to m, then we can form a cyclic subgroup $\langle n \rangle < (\mathbf{Z}/m\mathbf{Z})^\times$, and so, by Lagrange's theorem, the order of $\langle n \rangle$ divides $\varphi(m)$, and in particular this implies that $n^{\varphi(m)} \equiv 1 \pmod{m}$. □

One corollary is Fermat's Little Theorem.

Corollary 2.11. *If p is a prime, and $p \nmid n$, then $n^{p-1} \equiv 1 \mod p$.*

2.2 Normal Subgroups

We previously emphasized that subgroups can be thought of as a subset of symmetries which preserve an additional structure to a space. But there is an important distinction to make in the families of these subgroups. For instance, the space $O_n(\mathbf{R})$ orthogonal matrices forms a subgroup of $GL_n(\mathbf{R})$. These matrices operate as the transformation of rotation. A linear change of coordinates can be represented as an invertible matrix N, and if M is a rotation matrix, it's representation in a new coordinate system is NMN^{-1}. Thus, the reason why rotation matrices are not rotation matrices in every coordinate system is because the conjugation of an orthogonal matrix by an invertible matrix need not be orthogonal. On the other hand, if $M \in SL_n(\mathbf{R})$, so M preserves volume, then M still preserves volume in a new coordinate system, i.e. $NMN^{-1} \in SL_n(\mathbf{R})$ if $M \in SL_n(\mathbf{R})$. The subgroups of a group which remain in the subgroup under a change of coordinates are known as **normal**.

Theorem 2.12. *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 statement from (2) to (1) 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. □

Example. *If* G *is abelian, and* H *is a subgroup, then* $H \triangleleft G$, *because*

$$xHx^{-1} = (xx^{-1})H = H$$

Normality only becomes an issue when a group is non-abelian.

Example. *The special linear group* $SL_n(K)$ *is a normal subgroup of the general linear group* $GL_n(K)$. *This is because if* M *has determinant one, and* N *has non-zero determinant, then*

$$\det(NMN^{-1}) = \det(N)\det(M)\det(N)^{-1} = \det(M) = 1$$

The determinant map is an example of a homomorphism, a mapping between groups which will become very useful later.

Example. *Given a group* G *and* $S \subset G$, *consider the* **normalizer subgroup**

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

If S *is a subgroup of* G, *then* $S \triangleleft N_G(S)$. *In fact,* $N_G(S)$ *is the largest subgroup of* G *in which* S *is normal, for if* $S \triangleleft H$, *then* $H \subset N_G(S)$. *We can also define the* **centralizer subgroup**

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

Then $C_G(S) \triangleleft N_G(S)$. *If* $S = G$, *then we call* $C_G(S)$ *the* **center** *of* G, *also denoted as* $Z(G)$.

Theorem 2.13. *If K is a subgroup of $N_G(H)$, then KH is a group, and $H \triangleleft KH$.*

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 because

$$k_1h_1(k_2h_2)^{-1} = k_1h_1h_2^{-1}k_2^{-1} = k_1(k_2^{-1}k_2)h_1h_2^{-1}k_2^{-1} = (k_1k_2^{-1})\big[k_2(h_1h_2^{-1})k_2^{-1}\big]$$

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,

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

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

The trivial group is always normal in any group is lies in. Furthermore, for any group G, $G \triangleleft G$, resulting from the properties of closure in the group operation. It follows that every group has normal subgroups. We must take this into account in defining the property of sparcity of normal subgroups. We say 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. Think of simple groups as the equivalent of prime numbers for groups (they cannot be broken up into simple groups). If we can characterize all simple groups, then intuition tells us that there must be some way to put them together to characterize all groups. This is why the Hölder program of mathematics attempts to classify all finite simple groups, and characterize all finite groups in the process. In 2008, over one hundred years since the program started, we succeeded in characterizing all finite simple groups. Each one belongs to one of 18 infinite families of groups, and 26 'sporadic' groups which we cannot organize into families. The proof of this has taken over tens of thousands of journal articles, and for obvious reasons we will not replicate it.

Exercise 2.1. *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! Show that this is not possible.*

2.3 Homomorphisms and Quotients

Another way to understand groups is to understand how they are interrelated to one another. In group theory, the interrelations between different groups are formalized as 'homomorphisms'. If G and H are groups, a **homomorphism** between G and H is a function f such that $f(xy) = f(x)f(y)$. If a homomorphism is injective, we say that G can be **embedded** in H. If $G = H$, we call a homomorphism an **endomorphism**. Intuitively, a homomorphism is map which preserves the group structure of G. For instance, if $\varphi : G \to H$ is a homomorphism, then one easily verifies that $\varphi(e) = e$, $\varphi(a^{-1}) = \varphi(a)^{-1}$, so the identity and inversion is preserved. Intuitively, a homomorphisms is a map which implants the information of G into a subgroup of H, in such a way that certain elements may be identified. The **kernel** of a homomorphism φ, denoted $\ker(\varphi)$, is the set of elements in the domain of the homomorphism which map to the identity.

Lemma 2.14. *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 $\varphi(h) = e$, $\varphi(ghg^{-1}) = \varphi(g)\varphi(h)\varphi(g)^{-1} = \varphi(g)\varphi(g)^{-1} = e$. Thus ghg^{-1} is in the kernel for any element g in G, and we have shown normality. □

One verifies easily that a homomorphism is injective if and only if the kernel of the homomorphism is trivial. The image of a homomorphism is a subgroup of the range, but is not necessarily normal.

Corollary 2.15. *Let G be a group, and H and K subgroups of G such that $K \subset N_G(H)$. Then $H \cap K \lhd H$.*

Proof. If $K \subset N_G(H)$, then KH is a group. Consider the cannonical mapping from H to H/K by the mapping $h \mapsto hK$. $hK = K$ when $h \in K$, hence the kernel of the mapping is $H \cap K$, and thus $H \cap K \lhd H$. □

Example. *If M and N are matrices, the fact that* $\det(MN) = \det(M)\det(N)$ *implies that the map* $\det : GL_n(K) \to K^\times$ *is a homomorphism from the group of invertible matrices to the multiplicative group of non-zero elements in K.*

Example. *If $g \in G$, the map $n \mapsto g^n$ is a homomorphism from the additive group $\mathbf{Z}$ to G. In a similar fashion, we can extend this map to the exponential $x \mapsto e^x$, from $\mathbf{C}$ to the multiplicative group of complex numbers.*

Example. *The absolute value map on the multiplicative group of nonzero numbers is a homomorphism into the multiplicative group of nonzero real numbers, since* $|zw| = |z||w|$.

An **isomorphism** is a bijective homomorphism. If there exists an isomorphism between two groups, G and H, we denote this by writing $G \cong H$. The existence of an isomorphism means that all algebraic information about the domain is preserved in the image, and conversely, all the information in the range is contained in the domain. An **automorphism** is a bijective homomorphism from a group to itself. Thus an automorphism says that various objects in a group behave the same way. Note that the set of all automorphisms on a group G is a set of invertible functions on a space preserving some structure, and thus forms a group, denoted $\mathrm{Aut}(G)$.

Example. *The conjugation map* $z \mapsto \overline{z}$ *is an automorphism of both the multiplicative and additive group of complex numbers. Thus the number* i*, introduced to the real numbers to form the complex numbers, operates exactly the same as* $-i$.

Example. *Given an element* $g \in G$*, we have an automorphism* $h \mapsto g^{-1}hg = h^g$*, known as an* **inner automorphism** *of* G*. The map that sends* g *to it's inner automorphism is a homomorphism, since*

$$h^{g_0 g_1} = (g_0 g_1)^{-1} h (g_0 g_1) = g_1^{-1} g_0^{-1} h g_0 g_1 = (h^{g_0})^{g_1}$$

It's kernel is the center group $Z(G)$.

The next theorem is essentially 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.

Theorem 2.16. *Let* G *be a group generated by a subset* S*. Given any other group* H*, there is at most one homomorphism extending any given function* $f : S \to H$.

Proof. Given two homomorphisms f_0 and f_1 extending f, the set of elements of g upon which $f_0(g) = f_1(g)$ forms a subgroup containing S, hence equal to G. □

We can use the coset constructions of H and G to do something meaningful. Let G be a group and H a normal subgroup. For two cosets M and

N in G/H, define an operation $\circ$ on the cosets by $M \circ N = MN$. As $M = gH$ and $N = g'H$ for some $g, g' \in H$,

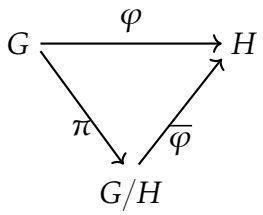
$$MN = gHg'H = gg'HH = gg'H$$

This follows by the normality of H. Thus the operation we have constructed is closed in G/H. The operation has an identity H, and the inverse of gH is $g^{-1}H$. With these properties, 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 H, hence we have shown that every normal subgroup is the kernel of some homomorphism.

Theorem 2.17 (The First Isomorphism Theorem)**.** *Let φ be a homomorphism between two groups G and H. Then $G/\ker(\varphi) \cong \mathrm{im}(\varphi)$.*

Proof. Let K be the kernel of φ. If $gK = hK$ for $g, h \in G$, $\varphi(g) = \varphi(h)$. Hence the mapping $\overline{\varphi}$ defined by $gK \mapsto \varphi(g)$ is well defined. It is a homomorphism as $gKhK = ghK$, 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. We also obtain that $\overline{\varphi}$ is injective, because $\varphi(a) = \varphi(b)$ implies $\varphi(ab^{-1}) = e$; hence $ab^{-1} \in K$, and $ab^{-1}K = K$. As K is normal, it is also true that $ab^{-1}K = aKb^{-1}$, so that $aKb^{-1} = K$, and thus $aK = Kb = bK$. The map is of course surjective onto its image, hence $\overline{\varphi}$ is an isomorphism, and we have proven what was needed. □

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 represents the functions in the first isomorphism theorem. Another notation, 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 $0 \to G \to H$ is exact precisely when the homomorphism between G and H is injective. Likewise, $G \to H \to 0$ is exact precisely when the map between G and H 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 2.18. *Every cyclic group is isomorphic to* $\mathbf{Z}$ *or* $\mathbf{Z}/n\mathbf{Z}$.

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

Thus we need only look at $\mathbf{Z}$ and $\mathbf{Z}/m\mathbf{Z}$ to prove things about general cyclic groups.

Theorem 2.19. *An infinite cyclic group has exactly two generators*

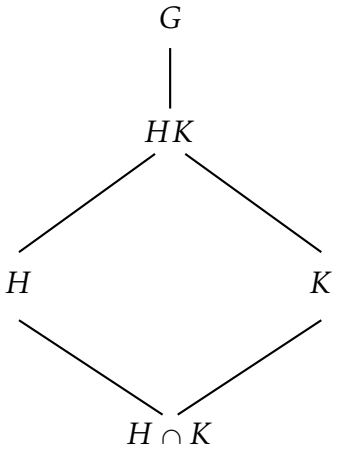
Proof. Any infinite cyclic group is isomorphic to $\mathbf{Z}$. Distinct generators of these cyclic groups are mapped to distinct generators in $\mathbf{Z}$, hence if we prove that $\mathbf{Z}$ has only two generators, then every infinite cyclic group has this property. If n is a generator for $\mathbf{Z}$, then there is m such that $mn = 1$, implying $n = 1/m$. Since n is an integer, this implies $n = \pm 1$. □

The first isomorphism is the catalyst for many other important isomorphism theorems, which enables us to construct canonical isomorphisms between objects.

Theorem 2.20 (The Second Isomorphism Theorem)**.** *Let G be a group, and K and H subgroups such that $K \subset N_G(H)$. Then $H/(H \cap K) \cong HK/K$.*

Proof. We have already justified in our discussion of cosets that $K \cap H$ will be normal in H, and K normal in HK. The chain of canonical homomorphisms $H \to HK \to HK/K$ has kernel $H \cap K$, and since the homomorphism is surjective from H onto HK/K, we obtain the required isomorphism. □

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



Theorem 2.21 (The Third Isomorphism Theorem)**.** *Suppose 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 we know that $(G/N)/(M/N) \cong G/M$.*

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 $(G/N)/(M/N) \cong G/M$. □

All we have done is shown $0 \to M/N \to G/N \to G/M$ is exact, and then the theorem follows naturally from the first isomorphism theorem.

Theorem 2.22 (The Lattice/Fourth Isomorphism Theorem)**.** *Let G be a group, and N a normal subgroup. Then for every subgroup H of G, the set $\{hN : h \in H\}$, which we denote $(G/N)_H$, is a subgroup of G/N and also, for every subgroup $(G/N)_H$ of G/H the underlying set H is a subgroup of G. This correspondence satisfies the following properties:*

- *$H \subset K$ if and only if $(G/N)_H \subset (G/N)_K$.*
- *If $K \subset H$, $(H : K) = ((G/N)_H : (G/N)_K)$.*
- *$(A \cap B)N = (G/N)_A \cap (G/N)_B$*
- *$H \lhd K$ if and only if $(G/N)_H \lhd (G/N)_K$*
- *$\langle (G/H)_S, (G/H)_T \rangle = G/H_{\langle S,T \rangle}$*

Proof. Given a subgroup H of G which contains N, define a mapping from H to G/N by $h \mapsto hN$. This is a homomorphism, and thus its range, which we have already defined as $(G/N)_H$, forms a subgroup of G/N. We have

thus proved this in general that $(G/N)_H$ is a subgroup of G/H for every subgroup H of G. Now suppose $(G/N)_S$ is a subgroup of G/N for some set S. Suppose two elements a and b are in S, so that aN and bN are in $(G/N)_S$. Then $ab^{-1}N$ is in $(G/N)_S$ as it is a subgroup, hence ab^{-1} is in S, so S is a subgroup of G. We leave it to the reader to show the rest of the properties of this correspondence. □

The list of properties above is not exhaustive. Almost all properties of subgroups are preserved by the mapping, so stop a while and think whether you can think of more.

Theorem 2.23. *If H and K are subgroups of a group G, then we define a (H,K) coset to be a subset of G of the form HgK, where $g \in G$. Then the set of (H,K) cosets partitions the group, and if G is the union of $Hg_1K,\ldots,Hg_nK$,*

$$|G| = \sum (H : H \cap g_n K g_n^{-1})$$

When K is trivial, we obtain Lagrange's theorem.

Proof. If $h_0 g_0 k_0 = h_1 g_1 k_1$, then $g_0 = h_0^{-1} h_1 g_1 k_1 k_0^{-1}$, and so

$$Hg_0K = Hh_0^{-1}h_1g_1k_1k_0^{-1}K = Hg_1K$$

so the (H,K) cosets form a partition. If $Hg_1K,\ldots,Hg_nK$ partition G, then

$$|G| = \sum |Hg_nK|$$

The subgroup H operates on the left on the set G/K, and $|Hg_nK|$ is the union of a particular orbit class of this action. One element of this orbit is g_nK, and $hg_nK = g_nK$ if and only if $h \in g_nKg_n^{-1} \cap H$, so the orbit stabilizer formula says

$$|Hg_nK| = (H : H \cap g_nKg_n^{-1})$$

This completes the proof. □

Theorem 2.24. *If $\varphi : G \to H$ is surjective, and $H' \lhd H$, if we define $G' = \varphi^{-1}(H')$, then $G' \lhd G$, and $G/G' \cong H/H'$ by the map $g \mapsto \varphi(g)H'$.*

Proof. The homomorphism $G \mapsto H \mapsto H/H'$ is surjective and has kernel G', so $G/G' \cong H/H'$. □

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

2.4 Cyclic Groups

The easiest groups to classify are those generated by a single element, the cyclic groups.

Theorem 2.25. *Let G be a finite cyclic group of order n, generated by an element g. Then g^m is a generator for G if and only $(n,m) = 1$.*

Theorem 2.26. *If G is a cyclic group with two generators x and y, then there exists a unique automorphism mapping x onto y.*

Theorem 2.27. *For every finite cyclic group G of period n, and for any integer d which divides n, there exists a unique subgroup of order d.*

Lemma 2.28. *Let g be an element of a group G, and suppose that the cardinality of $\langle g \rangle$ is a non-negative integer n. Then $g, g^2, \ldots, g^n$ 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 2.29. *For $0 < k < c$, $g^k \neq e$.*

Corollary 2.30. *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 2.31. $g^c = e$.

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 2.32. *$g^k = e$ 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 $g^a = e$. We also call this number the period of a. If this is infinite, we say a has infinite period.

Lemma 2.33. *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 2.34. *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 $G = \langle g^b \rangle$. □

Theorem (3.10) has some interesting repercussions in number theory. First, some notation is needed. For a group with two subsets S and M, define

$$SM = \{sm : s \in S, m \in M\}$$

For a single element a, define $aM = \{a\}M$, and Ma equivalently.

- 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 2.35. *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 m and n are 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.

Theorem 2.36. *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 $r \geqslant 1$, the equation $x^r \equiv 1 \mod p$ has no more than r solutions for x in $(\mathbf{Z}/p\mathbf{Z})^\times$ where p is prime. This follows that fact that the group is a field, and thus roots of a polynomial decompose the polynomial into linear factors. Let n be the maximal order of the cyclic subgroups $\langle m \rangle$, for $m \in (\mathbf{Z}/p\mathbf{Z})^\times$, generated by an integer g. Consider the polynomial $X^n - 1$. For any $m \in (\mathbf{Z}/p\mathbf{Z})^\times$, the order of m divides n, since the order of gm is the lowest common multiple of m and n, and must be less than n, and is hence equal. Thus $X^n - 1$ has $p - 1$ different solutions, but this implies $n \geqslant p - 1$. Of course, $n \leqslant p - 1$, so equality is obtained, and thus the group is cyclic. □

A generator of this group is known as a **primitive root**, and has many applications in number theory and cryptography. The problem with the above proof is that it gives us no method to find a generating element for the multiplicative group. This is an open problem that is incredibly important to cryptography, where multiplicative groups of the form above are used to construct encodings. Finding the primitive root for a really large prime is very difficult, which makes then very useful for cryptography.

Chapter 3

Group Actions and Symmetries

The symmetric group was previously defined as the set of permutations on a set. In the context of an example, this group seems trivial. This chapter will show why 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 3.1 (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, in the group $\mathbf{Z}^+$, the number n can really be considered the symmetry of adding n to every number in $\mathbf{Z}$, shifting all numbers to the right such that the resultant object is symmetric to the original. This is exactly the symmetry to which n corresponds with in Cayley's proof.

Through Cayley's theorem, all groups can be considered subgroups of the symmetric group, hence all groups can be considered a symmetric ac-

tion on some set. These actions provide another way to understand the structure of a group. Let us describe this in detail.

Definition. A **group action** on a group G and set X is a homomorphism π from G to the symmetry group on $|X|$ characters, inducing a symmetry on X for each group element in G. 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 two properties:

1. For all elements g and h in G and x in X, $g(hx) = (gh)x$
2. For the identity e in G, and elements x in X, $ex = x$

These properties are just restatements of the definition of a homomorphism; another way of saying the first 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 the definition of a homomorphism, hence these properties are just an equivalent way of defining a group action.

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

$$gx \mapsto g^{-1}xg$$

It is trivial to verify the group action properties. More interesting than this verification, we find that the permutation associated with any g in G is an automorphism of G. This does not always hold when the set G operates on is a group. Any automorphism which can be considered a conjugation of elements is an **inner** automorphism.

Definition. 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. We call the above object an orbit because intuitively, the group acts independently on each of a G-set's orbits, just like various planets orbit independently around the sun.

Definition. 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. 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.

Cayley's theorem asserts that for every group there exists an action that is faithful.

Definition. 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. 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. 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.

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 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 unit 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 3.2 (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 3.3 (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, for every orbit which is a singleton $\{x\}$, $G_x = G$, hence $(G : G_x) = 1$; thus, if we collect all these orbits, and remove them from the list we have, we obtain that

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

where $\{x_1, \ldots, x_n\}$ is the new set of orbit representatives where the orbit is not one.

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 following corollary is just a specialization of the previous theorem, though is just as useful.

Corollary 3.4 (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 3.5 (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. 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 3.6. *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 **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 3.7. *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 3.8. *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

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

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 homomorphism 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. In fact, one way of creating the sign homomorphism is to define it precisely in this way. The only problem with this intuitive definition is proving that if a permutation is the product of two different sets of cycles, then the value of the function is the same. Of course, our homomorphism above proves this, but only by constructing the homomorphism with the properties that we need to get the sign homomorphism.

The kernel of the homomorphism we have created is called the alternating group A_n, a normal subgroup of S_n. It of course consists of all permutations that are products of an even number of transpositions. Let us flesh out the theory of A_n.

We state the first lemma without proof due to its triviality, though it is useful in characterizing the order of the alternating group.

Lemma 3.9. *If τ is any transposition, S_n is the disjoint union of A_n and τA_n. Thus $A_n = n!/2$.*

And now we will begin a chain of lemmas, leading up to the statement that A_n is simple for $n \neq 4$.

Lemma 3.10. *A_n is generated by the set of three cycles.*

Proof. A_n is of course generated by the set of all compositions of two transpositions, hence we need only prove that each pair of transpositions can be represented as a three cycle, and vice versa. Let $(i,j)(m,n)$ be an arbitrary pair of two cycles ($i \neq j$, $m \neq n$)). Then one of three cases apply:

1. $i = m$, $j = n$: In this case $(i,j)(m,n) = \mathbf{1} = (123)^3$.
2. $i = m$, $j \neq n$: Then $(i,j)(m,n) = (mnj)$
3. $i \neq m$, $j \neq n$: Then $(i,j)(m,n) = (i,m,j)(i,m,n)$

We have covered all cases, hence any set of two pairs is generated by a three cycle, and thus any element in the alternating group. □

Lemma 3.11. *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 $\pi(x) \notin \sup(\pi\sigma\pi^{-1})$. Thus we have shown the value of $\pi\sigma\pi^{-1}$. □

Corollary 3.12. *All n cycles are conjugate.*

Proof. Let $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ be two three cycles. Consider a permutation φ such that $\phi(x_k) = y_k$. Then the lemma follows right from Lemma (5.11). □

Now we are ready to prove the most strenuous proof of the chapter, the simplicity of the alternating group.

Lemma 3.13. *A_n is simple when $n > 4$*

Proof. We show that if that if there is a normal subgroup of A_n that is not $\{e\}$, then it is precisely A_n. To do this, we need only show it contains all three cycles by Lemma (5.10). The normality of A_n in tandem with Corollary (5.12) means we need only show one three cycle is in A_n. Let G be a normal subgroup of A_n. Let $\varphi \neq \mathbf{1}$ be the permutation in G that contains the maximum number of fixed points. Partition φ into disjoint cycles. We must have at least one cycle with more than one element (else $\varphi = \mathbf{1}$). We want at least one cycle to be of length more than three so we can extract a three cycle. Suppose every cycle is of size two. Then, since φ is even, we must have at least two cycles. Denote two of these cycles (i, j) and (r, s). Pick some element k in the n characters acted on by A_n that is not i, j, r, or s (possible if the n in A_n is 5 or more), and let $\tau = (r, s, k)$. We know that, since G is normal, $\tau\varphi\tau^{-1}\varphi^{-1}$ is in G. This permutation leaves i and j fixed, as well as all other elements that φ fixes. In addition, it is not the identity; for example, s is not a fixed point. Thus we obtain an element with more fixed points that φ, a contadiction.

We conclude there is at least one cycle in φ with three or more characters, We pick three of these characters, and denote them i, j, and k. Suppose φ is not a three cycle. Then φ acts on at least five characters that

are not fixed; it must then include at least one more than three characters else it is a three cycle, and if it only moves four, it is an odd permutation. Let two other elements in $\sup(\varphi)$ be denoted r and s. Let $\tau = (k, r, s)$, and, as before, consider $\tau\varphi\tau^{-1}\varphi^{-1}$. Then this new permutation leaves i fixed, and is not the identity, a contradiction. Our assumption was that φ was not a three cycle, thus it must be. □

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.

Exercise 3.1. *If X is a G set that is not a singleton, then there is an element x in X with no fixed points.*

Exercise 3.2. *Suppose we have n prisoner's in jail, sentenced to death. The executioner's offer the prisoners a way to escape their judgement. They place n boxes in a room, each with a number from 1 to n in it, and a separate number from 1 to n inscribed on it (not related to the number inside the box in any way). They give each prisoner a unique number in the same manner of the boxes, and give each an opportunity. Each prisoner can open $n/2$ boxes, and if he finds inside a box a number sharing his or her own, then he succeeds his task. If every prisoner accomplishes this task, no-one will be executed. The naive method of solving this problem accomplishes this with a probability of less than 1%. Show, using the methods of permutations and cycles, that the prisoner's can design a strategy that leads to a 30% chance of success.*

Chapter 4

Sylow Theory

In 1872, Norwegian mathematician Ludwig Sylow proved a collection of theorems, called the Sylow theorems, which give detailed information about subgroups of a certain size within a group. Unlike the majority of chapters in this book, we begin with a theorem, rather than a definition. A strange methodology used in this proof will be used throughout the chapter: we induct on the size of the group.

Theorem 4.1. *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 an element g in G that is not the identity. If the order of g is pm, then g^m is order p. Instead, assume that g's order is not divisible by p. Since G is abelian, $\langle g \rangle$ is normal, hence we can form the group $G/\langle g \rangle$. We know that $|G| = |G/\langle g \rangle||\langle g \rangle|$. We know that $|\langle g \rangle|$ does not divide p, hence p must divide $|G/\langle g \rangle|$. As g is not the identity, we know the factor group is smaller than G, hence by induction, there is some element h in G such that $h\langle g \rangle$ is order p. Let n be the order of h. Then of course, since $h^n = e$, $p|n$. Using the same technique as before, we can obtain an element of order p from powers of h. □

A theorem of Cauchy generalizes this idea to arbitrary groups.

Theorem 4.2 (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 need no base case, as a group of any size less than 6 is abelian and thus we can apply Theorem (6.1). 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 $p \nmid |C_G(g)|$, and 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. By case to cases analysis we obtain the truth of the statement. □

Definition. 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, we obtain an interesting corollary: a group is a p-group if and only if every element has order a power of a prime.

Lemma 4.3. *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 4.4. *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 there is at least one element that is in the group, the identity. □

Corollary 4.5. *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_h is a group larger than $Z(G)$, since h itself is in G_h and h is in $Z(G)$, so we conclude that G_h must be order p^2 since it too is a subgroup of G. This means of course that every element commutes with h, so h is 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 – the proper Sylow Theorems!

Definition. 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 – the maximum order of a p subgroup in G.

In the next few proofs, let G be a group of cardinality $p^m q$.

Lemma 4.6. *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 trivially; simply consider the trivial subgroup. Now suppose

by induction that for all groups of smaller cardinality than G the theorem holds. Consider the group action of conjugation of G acting on itself. We know by the class equation 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 otherwise $C_G(x_i) = G$, and hence $Z(G)$ is empty. Hence we can use induction to show there is a subgroup of order p^k for each in $C_G(x_i)$ and hence in G for each k that we want. 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 4.7. *Let H be a p-subgroup of G, and P a p-Sylow subgroup. If $H \subset N_G(P)$, then $H \subset P$.*

Proof. We know that HP is also contained in the normalizer, and $P \triangleleft N_G(P)$. But by the second isomorphism theorem, we know that

$$(HP : P) = (H : H \cap P)$$

Hence by Lagrange's theorem, $HP = |H|/(|H \cap P||P|)$, and since each number on the right hand side is a power of p, so must $|HP|$. Since $HP \geqslant P$, we must have $HP = P$, else HP is a p-group greater than the biggest exponentional of p in G, the p-Sylow group P. From the fact that $HP = P$ we conclude $H \subset P$. □

This theorem can be easily strengthened.

Theorem 4.8. *If H is any p-subgroup of G, and P a p-Sylow subgroup. Then H is contained in some p-Sylow subgroup of G that is conjugate to P.*

Proof. Consider the set X of cosets gP for g in G, and let H act on X by the mapping

$$h(gS) \mapsto hgS$$

The cardinality of X is $|G|/|P| = q$. We know that the number of fixed points of the action is congruent to q modulo p, and since q is relatively prime to q, we know that this number cannot be zero. Thus there exists gP such that $hgP = gP$ for all h in H, and thus $h = gsg^{-1}$ for each and every element h. Thus $H \subset gPg^{-1}$. Since gPg^{-1} is conjugate to P, it too is a p-Sylow subgroup, and hence we obtain the statement above. □

Corollary 4.9. *All p-Sylow subgroups are conjugate.*

Proof. In the previous proof, let H be p-Sylow. Then H is contained in some conjugate p-Sylow subgroup to P. But H is the same size as this conjugate group, and hence H is equal to this conjugate p-Sylow subgroup. □

Corollary 4.10. *If there is only one p-Sylow subgroup, the group is normal.*

Proof. If P is the unique p-Sylow subgroup in a group G, then, for every g in G, $g^{-1}Pg$ is a p-Sylow subgroup. But then this means $g^{-1}Pg = P$. □

Theorem 4.11. *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 the statement that $s \mid q$. □

Theorem 4.12. *If s is the number of p-Sylow subgroups, then* $s \equiv 1 \pmod{p}$

Proof. Let S be a p-Sylow subgroup. S acts on the set of all p-Sylow subgroup X via conjugation. We claim that S is the only fixed point in this action. We know that if S' is a fixed point, then $S \subset N_G(S')$. But then by Lemma (6.7) that $S \subset S'$. Both are the same size, hence $S = S'$. Thus S is the uniqued fixed point of the action. We then have proved our theorem, as $|X| \equiv |X^S| \pmod{p}$, by lemma (12.3), and $|X^S| = |\{S\}| = 1$. □

The theorems above are really powerful to treating groups of finite order. Here is a powerful theorem.

Theorem 4.13. *Let p and q be prime numbers such that $q < p$, and $p \nmid (q-1)$. 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 $s = q$, then $q \equiv 1 \pmod{p}$, 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. if $su = s'u'$, then $s'^{-1}s = u'u^{-1}$, and since the two groups are disjoint, $s'^{-1}s = u'u^{-1} = e$. Thus each su is distinct, and we must have $|S||U|$ elements in SU. Then SU contains qp elements so $SU = G$. We obtain that $G \equiv S \times U$. □

Theorem 4.14. *Let G be a group with cardinality p^2q, where p and q are prime, $p < q$, and $p \nmid (q-1)$. Then G is abelian.*

Proof. If s is the number of p-Sylow subgroups, and r the number of q-Sylow subgroups, then we have the following equations, as in the last proof.

$$r \equiv 1 \pmod{q} \qquad s \equiv 1 \pmod{p} \qquad s \mid q^2$$

□

Theorem 4.15. *Let G be a finite group, and p the smallest prime of G. A subgroup of index p is normal in G.*

Proof. Let H be a subgroup of G of index p. Consider G/H. G acts on G/H by operation on the left. This is a homomorphism from G to S_p. Suppose g is in the kernel of hormomorphism. Then $gg'H = g'H$ for every coset $g'H$. In particular, $gH = H$, hence g is in H. Let the kernel of the homomorphism be K. Then G/K is isomorphic to a subgroup of S_p, and hence its cardinality must divide $p!$. But this means that

$$(G : K) = (G : H)(H : K) = p(H : K) \mid p!$$

hence $(H : K) \mid (p-1)!$. Now p is the smallest factor in $|G|$, and $(H : K) \mid |G|$, hence the only possible conclusion is that $(H : K) = 1$, else $|G|$ has a smaller factor. This means exactly that $H = K$, and hence H is normal in G as it is the kernel of a homomorphism. □

Chapter 5

Solvability

Solvability is the key to Galois' proof of the insolvability of the quintic. Furthermore, solvability is used in many more advanced settings throughout algebra. Thus it makes sense to introduce it in a group theory course before Galois theory to smoothen the transition between the theories.

Definition. Let G be a group. A **series** or **tower** is a finite sequence of groups beginning with G, and such that every sequential group is a subgroup of the previous.

To aid in remembering the definition, we write a sequence $(G_0, G_1, \ldots, G_m)$ which forms a tower as

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

This chapter focuses on a very specific type of tower.

Definition. A tower is called a **normal series** if every group in the tower is normal in its predecessor, so for each G_i that is not at the end, we may form the factor group G_i/G_{i+1} with the next element in the sequence. A normal series is **abelian** if each such factor group is abelian, and **cyclic** if every factor group is cyclic.

As with the notation for an ordinary tower, we write a normal series $(H_0, H_1, \ldots, H_m)$ as

$$H = H_0 \rhd H_1 \rhd \cdots \rhd H_m$$

so you needn't remember the definition if you're reading someone elses work; the notation tells you all you need to know!

Theorem 5.1. *Consider a normal tower*

$$H = H_0 \rhd H_1 \rhd \cdots \rhd H_m$$

and a homomorphism φ from G to H. Define a tower on G by letting G_i be $\varphi^{-1}(H_i)$. The tower then formed is a normal series. This tower is abelian/cyclic if and only if the other tower is abelian/cyclic.

Proof. As φ is any mapping from G to H, we have that

$$G_0 = \varphi^{-1}(H_0) = \varphi^{-1}(H) = G$$

Now we know H_{i+1} is normal in H_i for any index i for which H_i is defined. Restrict φ to only elements of G_i. Then φ is of course surjective onto H_i; we are then in the same position as Exercise (12), and we may conclude that G_{i+1} is normal in G_i, and $G_i/G_{i+1} \cong H_i/H_{i+1}$, hence all algebraic properties needed transfer from the factor group of H to the factor group of G. □

The property of having a normal tower is not special. For any group G, simply take the tower $G \supset \{e\}$, and that tower is trivially normal, but its factor groups do not really tell us anything about the group. The longer the tower, the more we separate the properties of the entire group as factor groups. It thus makes sense to take a tower that is maximalized in some way, to strain out as many properties as possible from the group.

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

Definition. 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 subgroup, $S_i/S_{i+1} \cong T_{\varphi(i)}/T_{\varphi(i)+1}$, so the factor groups obtained can really just be considered reorderings of one another.

The following lemma leads to an easy proof on the refinement of normal series. It's proof is perhaps the most technical in this report, but it at least has a nice picture corresponding with the lattice of subgroups to go along with it.

Theorem 5.2 (The Butterfly Lemma (Zassenhaus' Lemma))**.** *Let U and V be subgroups of a group G, and let U', V' be such that $U' \lhd U$, $V' \lhd V$. Then*

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

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

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. Our main strategy is to identify an isomorphism from the first formula to the second in the equation via the first isomorphism theorem. We will define a mapping from $U'(U \cap V)$ to $(U \cap V')/(U' \cap V)(U \cap V')$. Let the following mapping $u'x \mapsto x(U' \cap V)(U \cap V')$ be constructed. This mapping is well defined: If it is true that $ux = u'x'$, then $u'u^{-1} = xx'^{-1} \in U' \cap (U \cap V) = U' \cap V \subset (U' \cap V)(U \cap V')$, hence $x(U' \cap V)(U \cap V') = x'(U' \cap V)(U \cap V')$. Let us hope that the kernel of this mapping is $U'(U \cap V')$. We know that the kernel is precisely those elements representable as $u'x$, where $x \in (U' \cap V)(U \cap V')$, or that $u'x$ is an element of $U'(U' \cap V)(U \cap V') = U'(U \cap V')$, hence the kernel is $U'(U \cap V')$, and we have shown the isomorphism from first formula to second by the first isomorphism theorem, as the map is surjective. As the problem is symmetric, we obtain the isomorphism from third to second, and thus the entire chain of isomorphisms is created by transitivity of isomorphism. □

Do not worry if the statement above is unintuitive. It is only really a mechanic to be used in the next Theorem, and the author knows of no other use of it outside of this context.

Theorem 5.3 (Shreier)**.** *Two normal series in a group G ending with the trivial group have refinements that are equivalent.*

Proof. Consider two normal towers

$$G = G_0 \rhd G_1 \rhd \cdots \rhd G_n = \{e\}$$
$$G = G'_0 \rhd G'_1 \rhd \cdots \rhd G'_m = \{e\}$$

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

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

Similarily, if we define $G'_{i,j} = G'_{i+1}(G_j \cap G'_i)$, with a tower of G'_j generated in a similar fashion. By the butterfly lemma, with $U = G_{i+1}$, $U' = G_i$, $V = G'_{j+1}$, and $V' = G'_j$, we obtain that

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

We must also show the equivalency for $G_{i,m}$, $G_{i+1,0}$, $G'_{i,m}$, and $G'_{i+1,0}$. What are these groups?

$$G_{i,m} = G_{i+1}(G'_m \cap G_i) = G_{i+1}\{e\} = G_{i+1}$$
$$G_{i+1,0} = G_{i+2}(G'_0 \cap G_{i+1}) = G_{i+2}G_{i+1} = G_{i+1}$$
$$G'_{i,m} = G'_{i+1}$$
$$G'_{i+1,0} = G'_{i+1}$$

and hence

$$G_{i,m}/G_{i+1,0} \cong \{e\} \cong G'_{i,m}/G_{i+1}$$

We have verified the tower is normal and equivalent. They also refine the original towers as

$$G_{k,0} = G_k(G'_0 \cap G_{k-1}) = G_k(G \cap G_{k-1}) = G_k G_{k-1} = G_k$$

and similarily for $G'_{k,1}$, so we may embed the original tower in the new one. □

The main corollary requires a new concept, which follows so simply we state it without proof.

Definition. A **composition series** is a normal series which cannot be refined.

Corollary 5.4 (Jordan Hölder)**.** *All composition series of a set G are equivalent.*

All finite groups possess a composition series, as there are only finitely many subgroups of the group. We note this is not true of all groups. Consider the additive group $\mathbf{Z}$. Then every subgroup is of the form $a\mathbf{Z}$ for some a, and every subgroup is normal. Suppose we have a normal series

$$\mathbf{Z} \rhd a_1\mathbf{Z} \rhd a_2\mathbf{Z} \rhd \cdots \rhd a_n\mathbf{Z}$$

Then we can always refine it to

$$\mathbf{Z} \rhd ma_1\mathbf{Z} \rhd a_1\mathbf{Z} \rhd a_2\mathbf{Z} \rhd \cdots \rhd a_n\mathbf{Z}$$

for any integer m greater than one. This shows that there are no composition series because, given any series, we can always refine it.

Composition series can be considered the maximality of a normal series. Simple groups are minimalizations of normality. It is intuitive to connect these concepts. This theorem characterizes this.

Theorem 5.5. *A normal series is a composition series if and only if all factor groups in the series are simple.*

Proof. Consider an arbitrary normal series

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

Suppose G_k/G_{k+1} is not simple, so the factor group posesses a normal subgroup $(G_k/G_{k+1})_S$. By the lattice isomorphism theorem, there is a subgroup S such that $G_k \subset S \subset G_{k+1}$, and S is normal in G_{k+1}. Since G_{k+1} is normal in G_k, G_{k+1} is also normal in S, hence we have a refined normal

series. This proof by contraposition shows that all factor groups are simple in a composition series. Of course, if a normal series is such that every factor group is simple, it must follow that the series cannot be refined, because the existence of a refinement shows exactly that there is a normal subgroup between the two, hence the tower is a composition series. □

We now proceed to specialize to a particular type of normal series. First, a lemma.

Theorem 5.6. *From any abelian tower of an abelian group we can construct a cyclic tower.*

Proof. Let us prove this for all abelian groups, by by induction on the order of the group. For a base case, we note any abelian tower on the trivial group $\{e\}$ is cyclic. Now, consider an abelian 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

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

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

$$G = G_0/\langle g \rangle \rhd G_1/\langle g \rangle \rhd \cdots \rhd G_m/\langle g \rangle$$

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. Thus the statement holds for all finite abelian groups. □

Corollary 5.7. *An abelian tower on any group admits a cyclic refinement.*

Proof. Suppose for a group G we have an abelian tower

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

The lattice isomorphism theorem establishes a bijection between subgroups of G/X and subgroups of G that contain X. Consider a pair G_i and G_{i+1} in the tower. We have an abelian tower $G_i/G_{i+1} \rhd \{e\}$, and G_i/G_{i+1}, so we have a cyclic refinement of this tower. By Theorem (7.1), we can bring this

refinement back to G, and this will also be cyclic, beginning with G_i, and ending with G_{i+1}. Thus we can refine our original abelian tower with the cyclic tower constructed from each pair G_i and G_{i+1} to form a new abelian tower. □

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

Here we provide an explicit example before moving to the abstract. Consider the group $\mathrm{GL}_n(\mathbf{F})$. Let $N_n(\mathbf{F})$ be the set of elements that are zero both on and below the diagonal. For any r between 1 and n, the set $U_r = I_n + (N_n(\mathbf{F}))^r$ is a subgroup of $GL_n(\mathbf{F})$ (the determinant of all the matrices is 1). For U_k, define a mapping from U_k to the additive group $\mathbf{F}^{n-k}$ by taking the k'th upper diagonal. That is, if a matrix $M_n = [m_{i,j}]$. Then $M_n \mapsto (a_{1,k}, a_{2,k+1}, \ldots, a_{n-k,n})$. This is an homomorphism because U_k is a matrix of the form

$$\begin{pmatrix}
1 & 0 & \ldots & 0 & a_{1,k} & \ldots & \ldots & a_{1,n} \\
0 & 1 & \ldots & 0 & 0 & a_{2,k+1} & \ldots & a_{2,n} \\
0 & 0 & \ddots & 0 & 0 & 0 & \ddots & \vdots \\
0 & 0 & \ldots & 1 & 0 & 0 & \ldots & a_{n-r,n} \\
0 & 0 & \ldots & \ldots & \ddots & \ldots & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

and hence for any two matrices $M = [a_{i,j}]$ and $N = [b_{i,j}]$ in U_k, $MN = [c_{i,j}]$ fits the equations $c_{n,k+n-1} = a_{n,k+n-1} + b_{n,k+n-1}$ (the identity matches up with the r'th column). The kernel of the homomorphism is U_{k+1}, hence U_{k+1} is normal in U_k, and $U_k/U_{k+1} \cong F^{k-r}$ and the factor group is abelian. Thus the sequence (U_k) is an abelian tower, and U is solvable.

Here is a simpler example. Let G be an abelian group. Then the series $G \rhd \{e\}$ is an abelian tower, because $G/\{e\} \cong G$, and is hence abelian. Thus G is solvable.

Theorem 5.8. *A subgroup of a solvable group is solvable.*

Proof. Consider a solvable group G, and a subgroup H. 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 sequence (H_k), where $H_k = G_k \cap H$. We know that, since G_k is normal in G_{k+1}, so too are H_k and H_{k+1} The second isomorphism theorem tells us that

$$(H \cap G_{i+1})/(H \cap G_i) = (H \cap G_{i+1})/(H \cap G_i \cap G) \cong (H \cap G_{i+1})G_i/G_i \subset G_{i+1}/G_i$$

and thus H_i/H_{i+1} is abelian. □

Theorem 5.9. *Let G be an arbitrary group, and H an arbitrary normal subgroup. G is solvable if and only if both H and G/H are.*

Proof. Let G be a solvable group, with an abelian tower

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

Given this abelian tower, consider the canonical mapping π from G to G/H, and define a new sequence (H_k) such that $H_k = \pi(G_k)$. We know H_k is normal in H_{k+1} by Exercise 13. Furthermore, we know that $H_k = G_k/H$, hence, by the third isomorphism theorem,

$$H_k/H_{k+1} = (G_k/H)/(G_{k+1}/H) \cong G_k/G_{k+1}$$

Conversely, suppose that H and G/H is solvable. Then by Theorem (7.1) we can construct an abelian tower on G, which ends with $H = \pi^{-1}(e)$. Combine this with the abelian series on H, and we obtain that G is solvable. □

Definition. Let G be a group. A **commutator** is an element of G that can be written $ghg^{-1}h^{-1}$, for two elements g and h in G, which we also write as $[g,h]$. Define the **commutator** or **derived subgroup** $D(G)$ of the group G to be the group generated by the set of commutators in G.

Lemma 5.10. *For any G, $D(G)$ is normal in G.*

Proof. Let g be an element of G, and $hkh^{-1}k^{-1}$ an element of $D(G)$,

$$ghkh^{-1}k^{-1}g^{-1} = (ghg^{-1})(gkg^{-1})(gh^{-1}g^{-1})^{-1}(gkg^{-1})^{-1}$$

Hence it is an element of the commutator. We leave it to the reader to prove that, if gkg^{-1} holds for every k in a set K which is a subset of a group G, from which g reside, then $\langle K \rangle$ is normal in G. □

Lemma 5.11. *For any group G, $G/D(G)$ is commutative.*

Proof. For any gh, $g^{-1}h^{-1}gh$ is in $D(G)$, hence

$$gD(G)hD(G) = ghD(G) = ghg^{-1}h^{-1}hgD(G) = hD(G)gD(G)$$

and we have calculated that the group is commutative. □

Lemma 5.12. *For any homomorphism from G to H such that H is commutative, $D(G)$ is a subset of the kernel of H.*

Proof. Let φ be the homomorphism above, and let g and h be arbitrary elements of G. By doing the following calculation,

$$\varphi(ghg^{-1}h^{-1}) = \varphi(g)\varphi(h)\varphi(g^{-1})\varphi(h^{-1}) = \varphi(g)\varphi(g^{-1})\varphi(h)\varphi(h^{-1}) = e$$

Since these elements generate $D(G)$, every element in $D(G)$ is composed of elements like this, which all cancel out in φ, hence $D(G)$ is in the kernel of φ. □

Corollary 5.13. *If G is a group with normal group N, and G/N is abelian, then $D(G) \subset N$.*

Commutator groups give us the key to unravelling the notion of solvability. We know $D(G)$ is normal in G, and we also know $D(D(G))$ is normal in $D(G)$, and so on and so forth, and each factor group created is abelian. Define $D^n(G)$ recursively by $D^n(G) = D(D^{n-1}(G))$. Via this, for each n we get a normal series

$$G \rhd D(G) \rhd D^2(G) \rhd \cdots \rhd D^{n-1}(G) \rhd D^n(G)$$

If it eventually holds that $D^n(G) = \{e\}$ for some n, then we obtain an abelian series, and G is solvable. What is amazing is this statement holds in reverse.

Theorem 5.14. *If a group G is solvable, $D^n(G) = \{e\}$ for some n.*

Proof. Suppose G is solvable, and hence has an abelian normal series

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

For each r, $D(G_r) \subset G_{r+1}$, as G_r/G_{r+1} is abelian. We claim $D^r(G) \subset G_r$. How do we prove this? Well $D(G) \subset G_1$, hence $D^2(G) \subset D(G_1) \subset G_2$. Thus the claim follows by induction. It follows that $D^n(G) \subset \{e\}$, and the two groups are hence equal as the only subgroup of the trivial group is itself. □

The main use of this theorem is not to show other groups are solvable, but to show that some groups are not solvable. Solvability began solely to answer questions in Galois field theory, which considers permutations of polynomial equations. This is just a representation of the symmetric group. You should see the connection between the following theorem and the insolvability of the quintic equation, at least in the numbers used.

Theorem 5.15. *S_n is not solvable when $n \geqslant 5$.*

Proof. Let i,j,k,r,s be 5 distinct characters that are being permuted in S_n. Let $\sigma = (i\ j\ k)$, and let τ be $(k\ r\ s)$. Then

$$[\sigma, \tau] = \sigma\tau\sigma^{-1}\tau^{-1} = (r\ k\ i)$$

As each r, k, and i were arbitrary, we know all three cycles are in $D(S_n)$. As only three cycles were used in the commutators above, all three cycles are also in $D^2(S_n)$, and so on inductively, hence we will never have $D^m(S_n) = \{\mathbf{1}\}$. Thus S_n is not solvable. □

Theorem 5.16. *If G is a p-group, G is solvable.*

Proof. Let G be of cardinality p^m. We proved in Lemma (6.6) that for any k between 1 and $m-1$ there is a subgroup of order p^k. In particular, there is a subgroup of order p^{m-1}. Denote this group G_1. G_1 is normal in G, and G/G_1 is of order p, so the group must be cyclic as p is prime. By induction, we must do this for G_1, G_2, etc. to construct a normal series where each factor group is cyclic. □

Chapter 6

Direct Products and Abelian Groups

This chapter presents methods for constructing new groups from smaller ones. By doing this, we will be able to break down a group into smaller, component groups via the reverse of this technique. We call a construction of this form the direct product.

Definition. Consider an indexed family of groups $\{G_i\}_{i\in I}$. The direct product of these groups, denoted $\times_{i\in I} G_i$ is the group formed by taking the cartesian product of elements in each G_i. The operation defined on the group is

$$\left[\bigtimes_{i\in I} g_i\right] \circ \left[\bigtimes_{i\in I} h_i\right] = \left[\bigtimes_{i\in I} g_i h_i\right]$$

where the operation is just taken coordinatewise.

Each individual group G_i in the direct product of a group can be studied in order to understand the entire direct product. If we can identify a group as isomorphic to the direct product of a set of groups, then we can understand the group by understanding each individual group from which it is structured. The following method shows how we can deconstruct a group into its direct products.

Theorem 6.1. *If a group contains two subgroups who are disjoint but for the identity, commute with each other, and whose product contains the whole group, then the whole group is isomorphic to the direct product of the two subgroups.*

Proof. Let G be a group, with two subgroups H and K such that $H \cap K = \{e\}$, H and K commute, and $HK = G$. Define a mapping φ from $H \times K$ to G by the calculation $(h,k) \mapsto hk$. Since H and K commute,

$$\varphi((hh',kk')) = hh'kk' = hkh'k' = \varphi(h,k)\varphi(h',k')$$

It follows that φ is a homomorphism. If $hk = e$, then $k = h^{-1}$, so k is in both H and K, which means $k = e$ as H and K are disjoint but for the identity. We have shown exactly that the kernel of the function φ is trivial. Furthermore, φ is surjective, as $HK = G$. It has been shown that φ is an isomorphism, and therefore G is isomorphic to $H \times K$. □

An example of a direct product that we are very familiar is is the additive vector group $\mathbf{F}^n$, which is isomorphic to the direct product $(\times_{k=1}^{n} \mathbf{F})$. In general, for any ring R, the module R^n is isomorphic to $(\times_{k=1}^{n} R)$.

Another example is the group $E_{p^n} = (\times_{k=1}^{n} \mathbf{Z}/p\mathbf{Z})$, where p is a prime. The group is of order p^n; in general, a group $(\times_{k=1}^{n} G_i)$ has cardinality $\prod_{k=1}^{n} |G_i|$. Each non-trivial element is of order p. Consider the group E_{p^2} for some prime p. We know that if G and H are two subgroups of order p, then $G \cap H = \{e\}$ unless the two groups are equal. Thus if m is the number of p subgroups, then E_{p^2} possesses at least $(p-1)m + 1$ distinct elements. Thus $(p-1)m + 1 \leqslant p^2$, from which we conclude that m is bounded above by $p + 1$. We leave it to the reader to identify the $p + 1$ distinct subgroups which bound m below, from which we can conclude that m is exactly equal to $p + 1$.

For each group G_k in a direct product $\times_{i \in I} G_i$, we have a homomorphism π_k from $\times_{i \in I} G_i$ to G_k defined by the surjective mapping $(\times_{i \in I} g_i) \mapsto g_k$. The kernel of this mapping is the set of all elements $(\times_{i \in I} g_i)$ such that $g_k = e$, and hence we can quotient this kernel out to get a direct isomorphism to G_k. Think of G_k as the coordinate axis in the direct product group.

Direct products are the key to classifying a certain class of abelian groups. The ideas of this classification you have probably learned before you even read this article; there is a distinct connection to the ideas of linear algebra. Here is the special class of abelian groups we will classify.

Definition. A group is finitely generated if it is generated from a finite set.

It will help to introduce some notation to deal with splitting up components of abelian groups. We note the formal definition in the infinite case is not used for now, but we include it for thoroughness.

Definition. Given a collection of abelian groups $(G_i)_{i \in I}$, we define the **direct sum** $\bigoplus_{i \in I} G_i$ to be the subgroup of the direct product of those groups consisting of all elements where there are only finitely many elements that are non-identity elements. In the case of a finite product of elements, the direct sum is equivalent to the direct product.

You can probably see how abelian groups connect to vector spaces. In some sense, vector spaces are the canonical abelian if you consider their addition as the fundamental operation that defines them. The definitions below should be familiar to you from a study of vector spaces.

Definition. If an abelian group is generated by a set S, then that set is a **basis** if every element in the group is uniquely represented by a sum of elements in S. If a group has a basis, we say the group is **free**.

For every set S, there is an abelian group whose basis is S. Let us construct this group. Consider the set of mappings from S to $\mathbf{Z}$. In particular, consider the mappings that assign 1 to some element s in S, and 0 to every other element. Then this set forms a basis to all of the function group, and we can consider S to be the basis of this set. The group we have constructed is called the free abelian group generated by S, which is commonly denoted $F_{ab}(S)$. Every free group is isomorphic to the free abelian group generated by its basis.

Theorem 6.2. *Every abelian group is isomorphic to a factor group of a free abelian group.*

Proof. Consider an abelian group G. Take a generating set S of G (in the worst case, we may take G as the generating set). Form the abelian group $F_{ab}(S)$. Define a homomorphism φ from $F_{ab}(S)$ to G by $\varphi(\sum_{k=1}^{n} n_k g_k) = \sum_{k=1}^{n} n_k g_k$. This homomorphism is surjective, hence G is isomorphic to the factor group by the kernel of the homomorphism with $F_{ab}(S)$. □

In particular, if an abelian group is finitely generated, this group is isomorphic to a factor group of $\mathbf{Z}^n$ for some n. This means if we want to classify all finitely generated abelian groups, we first must must classify subgroups on $\mathbf{Z}^m$ for every m. We will now build up the mechanics of how we can classify this.

Lemma 6.3. *If a homomorphism φ maps from an abelian group G onto a free abelian group H, then G is isomorphic to the direct sum of the kernel of φ and H.*

Proof. Let $h_{i\,i\in I}$ be a basis for H. For each h_i, consider some g_i in G such that $f(g_i) = h_i$. Take the group C generated by the set of elements g_i. We claim C is isomorphic to H. We know that φ restricted to C is still surjective, and if $\varphi(\sum_{i\in I} n_i g_i) = 0$, then $\sum_{i\in I} n_i h_i = 0$, hence all n_i are zero, which means $\sum_{i\in I} n_i g_i = 0$. Hence φ is injective when restricted to C, and we obtain an isomorphism. Let K be the kernel of φ. We have shown $C \cap K = 0$. Now we must show $C + K = G$. Let x be an arbitary element of G, and let $f(x) = \sum_{i\in I} n_i g_i$. Then $x - \sum_{i\in I} n_i h_i$. Thus $x - \sum_{i\in I}$ is in K, and $x \in K + C$. It follows that G is isomorphic to the direct sum of C and K. □

The next theorem allows us to characterize all subgroups of free groups, which connects to our objective of classifying subgroups of $\mathbf{Z}^n$.

Theorem 6.4. *Every subgroup of a free abelian group with a finite basis is free, with a basis of size less than or equal to the size of the entire group.*

Proof. We prove by induction on the size of the group. If $n = 1$, the group is cyclic, and thus every subgroup is cyclic, generated by a single element which forms the basis provided the group is infinite. Now suppose that for $m \leqslant n$ this theorem holds. Let G be a free abelian group with basis

$\{g_1, g_2, \ldots, g_n\}$, and consider a subgroup H. We have a homomorphism π_1 from G to $\langle g_1 \rangle$ defined by the mapping

$$\pi_1(\sum_{k=1}^{n} l_k g_k) = l_1 g_1$$

Consider the restriction of this homomorphism from H, and the resultant kernel H'. Then the range of this restricted homomorphism, and hence is of the form $\langle a g_1 \rangle$ for some integer a. The kernel H' is a subgroup contained in the group $\langle g_2, \ldots, g_n \rangle$, and hence has a basis $h_1, h_2, \ldots, h_q$, where $q \leqslant n-1$. If $a \neq 0$. By Lemma (8.3), there is a subgroup C of H isomorphic to $\langle a g_1 \rangle$, and $H = H' \cdot C$. Now C is either zero or infinite cyclic, which proves that H is free. □

Corollary 6.5. *Every pair of bases of a finitely generated free abelian group is of the same cardinality.*

Proof. Let G a finitely generated free abelian group with two bases of size T and Q respectively. Using the basis corresponding to T, we conclude the group G/pG is a sum of T cyclic groups of order p, and is thus of cardinality p^T. Conversely, using the basis of Q, we conclude the basis is of order p^Q. But then $p^T = p^Q$, hence $T = Q$. □

The number of elements in the basis of a free abelian group is called the **rank** of the group. The problem with the proof above is it is not so easy to construct such a basis. For the next theorem, we will use the fact that any subgroup of a free group is finitely generated, but only to provide an algorithm to conclude our objective of classifying all subgroups of $\mathbf{Z}$.

Theorem 6.6. *Let G be a finitely generated abelian group, generated by a set of n elements. Then*

$$G \cong \mathbf{Z}/a_1\mathbf{Z} \oplus \mathbf{Z}/a_2\mathbf{Z} \oplus \ldots \mathbf{Z}/a_r\mathbf{Z} \oplus \mathbf{Z} \oplus \cdots \oplus \mathbf{Z}$$

where $r \leqslant n$ and such that $a_i \mid a_{i+1}$ for each a_i, and the number of $\mathbf{Z}$ groups in the direct product is $n - r$. This formulation is unique for any such subgroup.

Proof. Consider the group G defined above. We know that $G \cong \mathbf{Z}^n/K$ for some subgroup K of $\mathbf{Z}^n$. Suppose we have an automorphism φ on $\mathbf{Z}^n$. Then this induces a mapping from K to another subgroup K', and $\mathbf{Z}^n/K \cong$

$\mathbf{Z}^n/K'$. Our strategy is thus to simplify $\mathbf{Z}/K$ via these automorphisms to determine that each such group $\mathbf{Z}/K$ is isomorphic to one of the sets above. What's good about this algorithm is that it gives us a method to find this isomorphism.

Let K be a subgroup of $\mathbf{Z}^n$. Then we know that K is finitely generated by a set of elements $\{k_1, k_2, \ldots, k_l\}$. Each k_i is an array of n integers $(k_{i,1}, k_{i,2}, \ldots, k_{i,n})$, as it is an element of $\mathbf{Z}^n$. This motivates that we construct the matrix

$$\begin{pmatrix} k_{1,1} & k_{1,2} & \ldots & k_{1,n} \\ k_{2,1} & k_{2,2} & \ldots & k_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ k_{l,1} & k_{l,2} & \ldots & k_{l,n} \end{pmatrix}$$

Can we create row and column operations which correspond to isomorphisms of $\mathbf{Z}^n$. We wouldn't be constructing this matrix if not! these are the operations we require:

- We may interchange two rows i and j. This corresponds to swapping the order of two generators in the set, which does not change the subgroup K we are operating on.
- Multiplying a row i by negative one corresponds to swapping a generator k_i with its inverse, $-k_i$. We note that this also does not change the subgroup K we are operating on.
- Adding row i to row j, where $i \neq j$, corresponds to replacing a generator k_i with $k_i + k_j$. These generators are equivalent, so K is the same.
- Interchanging Columns i and j corresponds to an automorphism of $\mathbf{Z}^n$ where we interchange two coordinates.
- Multiplying a column i by negative one corresponds to an automorphism of $\mathbf{Z}^n$ where a specific coordinate is inverted in every element.
- Adding a column i to a column j corresponds to an automorphism of $\mathbf{Z}^n$. This is perhaps the only non-trivial automorphism to see. We map a vector $(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n)$ to $(x_1, \ldots, x_i, \ldots, x_i + x_j, \ldots, x_n)$. Then $(x_1 + y_1, \ldots, x_i + y_i, \ldots, x_j + y_j, \ldots, x_n + y_n)$ is mapped to $(x_1 + y_1, \ldots, x_i + y_i, \ldots, x_i + x_j + y_i + y_j, \ldots, x_n + y_n)$, which is precisely the

addition of the individual mappings, hence the mapping is a homomorphism. Verification that this mapping is an automprhism is left to the reader.

These actions are sufficient to reduce any matrix to the 'Smith Normal Form', a matrix of the form

$$\begin{pmatrix} \alpha_1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \alpha_n & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}$$

where the only non-zero entries are on the diagonal, and each α_i divides α_{i+1}. How is this useful to us? It means precisely that every subgroup K can be by automorphisms transformed into $\alpha_1\mathbf{Z}\oplus\alpha_2\mathbf{Z}\oplus\cdots\oplus\alpha_n\mathbf{Z}\oplus\{0\}\oplus\cdots\oplus\{0\}$, and thus our original finitely generated abelian group is isomorphic to $\mathbf{Z}/\alpha_1\mathbf{Z}\oplus\mathbf{Z}/\alpha_2\mathbf{Z}\oplus\cdots\oplus\mathbf{Z}/\alpha_n\mathbf{Z}\oplus\mathbf{Z}\oplus\cdots\oplus\mathbf{Z}$. All that is left is show our method of reduction of an arbitrary integer matrix to Smith normal form. For now, we suppose it true, and we will establish the technique after this proof is complete. □

The technique to reducing an integer matrix to smith normal form turns out to be quite simple. Clearly, we need only provide a technique to reduce a matrix to the form

$$\begin{pmatrix} \alpha & 0 \\ 0 & M \end{pmatrix}$$

Where M is a submatrix of one less column, and such that α divides every entry in M. By induction, the rest of the method is taken care of.

The first step of our algorithm is to check if the matrix you are reducing is the zero matrix; if this is true, we are done before we have even started. Otherwise, move the element in the matrix of smallest absolute value to the top left hand corner of the matrix, which we call the pivot. Secondly, repeatedly add or subtract the pivot row from each subsequent row such that the absolutel value of each pivot row and column entry is reduced. Do this for the pivot column from all other columns also.

Eventually, either all entries in the pivot row and column will be zero, or one will have absolute value smaller than the pivot entry. In this case, move this entry to the top left corner, and continue the process. We can only reduce the absolute value of an entry finitely many times before we are done, so eventually, the pivot row and column will be reduced to zero beside from the pivot entry.

Finally, check if the pivot entry divides every other entry in the matrix. If so, we can recurse to the submatrix. Otherwise, take the row that is not divisible by the pivot. Add this row to the first row, and return to adding and subtracting the rows and columns. This will reduce the size of the pivot, meaning we must eventually terminate.

It is best to learn an algorithm by computing out an example by hand. Here is an example. Consider a homomorphism from $\mathbf{Z}^3$ to a group G with kernel $\langle (6,3,3), (4,5,7), (3,2,2) \rangle$. What group is G isomorphic to. First, we form the matrix

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

We bring the smallest entry, the one with the value of two, up to the pivot entry,

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

then we reduce the row sizes

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

and the column sizes

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

We move the 1 on the first row to the pivot, and then reduce to get the matrix

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

Continuing by induction, you should end up with a matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{pmatrix}$$

Which means K is isomorphic to $\mathbf{Z} \oplus \mathbf{Z} \oplus 6\mathbf{Z}$, and $\mathbf{Z}^3/K$ is isomorphic to $\mathbf{Z}/\mathbf{Z} \oplus \mathbf{Z}/\mathbf{Z} \oplus \mathbf{Z}/6\mathbf{Z} \cong \mathbf{Z}/6\mathbf{Z}$.

Exercise 6.1. *What is the order of* $(\times_{i \in I} g_i)$ *in relation to the order of each* g_i *in the direct product.*