Notes on the Theory of Groups

by

Marcel K. Goh

1. INTRODUCTION

This set of notes assumes that the reader has had some exposure to linear algebra; the most crucial notions are briefly outlined here. Next, the group axioms are introduced and some basic properties of groups are given.

1.1. Basic Properties of Square Matrices

We shall mainly concern ourselves with matrices of the form

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

where n is some positive integer and each entry $a_{ij} \in \mathbf{R}$. We call the set of all $n \times n$ matrices with real entries $M_n(\mathbf{R})$.

For two matrices A and B, we may form their *sum* $A + B = (a_{ij} + b_{ij})$. (This notation means that at the ith row and jth column, the entry is the sum of a_{ij} and b_{ij}.) Given a real number α, we obtain the scalar product $\alpha A = (\alpha \cdot a_{ij})$ by multiplying every entry in A by α.

We can also *multiply* $n \times n$ matrices A and B with the following formula:

$$A \cdot B = (c_{ij}) \quad \text{where } c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}. \tag{1}$$

Since a matrix represents a linear transformation, multiplying matrices is like composing functions. If S and T are the transformations represented by the matrices B and A, respectively, then the matrix product $A \cdot B$ can be thought of as the following composition of transformations:

$$\mathbf{R}^n \quad \xrightarrow{S} \quad \mathbf{R}^n \quad \xrightarrow{T} \quad \mathbf{R}^n$$

For matrices A and B the commutativity of addition

$$A + B = B + A \tag{2}$$

is valid, and for three matrices A, B, and C, the distributive law

$$A \cdot (B + C) = A \cdot B + A \cdot C \tag{3}$$

and the associativity of multiplication

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C \tag{4}$$

may be proven to hold as well. In particular, (4) is laborious to prove from the definition given in (1), but easy to derive when reasoning about matrices as transformations.

Unlike addition, multiplication is not commutative: $A \cdot B$ does not equal $B \cdot A$ in general. To prove this, we simply note that

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ but } \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

The matrix

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

is called the *identity matrix* and has the property that $AI = IA = A$ for any matrix A. [When no ambiguity can arise, we often omit the $\cdot$ symbol when denoting a product.]

We say that a matrix A is *invertible* if and only if there exists a matrix B such that $AB = BA = I$; otherwise, it is called *singular*. Not all matrices are invertible. For example, the matrix $\mathbf{0}$, all of whose entries are 0, is not invertible in any dimension. The identity matrix is easily seen to be invertible (take $B = I$). Note that if an inverse matrix exists for a given matrix A, then it is unique, for if $AB = AB' = I$ and $B'A = BA = I$, then, multiplying the first identity by B on the left, we arrive at $BAB = BAB'$, i.e. $B = B'$.

Since a 1×1 matrix (a) contains only one real number, it is invertible if and only if $a \neq 0$. A 2×2 matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is invertible if and only if $ad - bc = 0$, since for any such matrix,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -c \\ -c & a \end{pmatrix} = \begin{pmatrix} ad - bc & 0 \\ 0 & ad - bc \end{pmatrix},$$

and we can multiply both sides by $1/(ad - bc)$ to obtain the identity on the right, provided that $ad - bc$ is nonzero.

In general, there exists a function $\det : M_n(\mathbf{R}) \to \mathbf{R}$ such that a matrix $A \in M_n(\mathbf{R})$ is invertible if and only if $\det A \neq 0$. This determinant can be calculated as a sum over $n!$ terms; this formula will not be useful for our purposes.

1.2. The General Linear Group

Let us now restrict our attention to a certain subset of square matrices, namely those whose determinant is nonzero. This set is called the *general linear group of degree n* and is denoted $GL_n(\mathbf{R})$ when all matrix entries are real numbers. Remark that, since $(-1) + (1) = (0)$, this set is not closed under addition; scalar multiplication is also no longer a safe operation, since multiplying any matrix by 0 results in a singular matrix.

In return for these two forfeited closure properties, we get closure under matrix multiplication.

Proposition I. *Suppose that two matrices A and B are invertible. Then their product AB is also invertible.*

Proof. Consider $B^{-1}A^{-1}$ and the product $(B^{-1}A^{-1})(AB)$. By associativity of multiplication, this becomes $B^{-1}(A^{-1}A)B = B^{-1}IB = IB^{-1}B = II = I$. Alternatively, use the fact that $\det(AB) = \det(A)\det(B)$, which is nonzero because both $\det(A)$ and $\det(B)$ are nonzero. ∎

Thus the set $GL_n(\mathbf{R})$, under the operation of matrix multiplication, has a multiplicative inverse A^{-1} for every matrix A. It also contains an identity element I and the multiplication operation is associative. These are the properties of a group.

1.3. Groups

A *group* G is a set on which is defined a rule of combination such that the product of two elements $g, h \in G$, denoted $g \cdot h$ or gh, is also in G. Furthermore, the following three properties must hold:

a) Multiplication must be associative: for all $g, h, k \in G$, $(gh)k = g(hk)$.

b) There exists an identity element $e \in G$ such that $ge = eg = g$ for all $g \in G$. This element is also often denoted 1.

c) For every element $g \in G$, there exists an inverse element g^{-1} such that $gg^{-1} = g^{-1}g = e$.

An immediate consequence of the axioms is that the identity e is unique. For if both e and e' are the identity of a group, then $e = ee' = e'$. Likewise, any element g has a unique inverse, in the sense that if two elements h and h' are both inverses of g, then

$$h = he = h(gh') = (hg)h' = eh' = h',$$

and they were the same to begin with.

The term *order* serves a somewhat dual purpose in group theory. The order of a group G is the number of elements it contains and this value, also denoted $|G|$, need not be finite. On the other hand, we define the order of a group *element* g to be the smallest $k > 0$ such that $g^k = e$. If no such k exists, then g is said to have infinite order.

Probably the most familiar group is the set of all integers, denoted $\mathbf{Z}$, under the operation of addition. It has 1 as its identity, inverses $-a$ for every whole number a, and the commutative property; likewise, any vector space V is also a group under vector addition. The set of nonzero real numbers forms a group under multiplication. In these examples, the binary operation has the property that for any $g, h \in G$, the products gh and hg are equal. A group where this holds is called an *abelian* or *commutative* group. An example of a non-abelian group is the group Q_8 of quaternions with identity 1, governed by the identities $(-1)^2 = 1$ and $i^2 = j^2 = k^2 = ijk = -1$. It has the following multiplication table:

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

The general linear group $GL_n(\mathbf{R})$ of invertible $n \times n$ matrices is another example of a non-abelian group, whenever $n > 1$.

Group theory is intimately connected to the study of symmetries of an object. Let T be a set and let $\mathrm{Sym}(T)$ denote the set of all bijections from T to itself. This is called the *symmetric group* on T as, under composition of functions, it obeys all the group axioms: It contains the identity transformation Id_T and every bijection f has an inverse bijection f^{-1}. In some sense, the symmetric group is the most general group, because all other groups arise from adding restrictions to these bijections. For instance, $GL_n(\mathbf{R}) \subseteq \mathrm{Sym}(\mathbf{R}^n)$.

1.4. Permutation Groups

A bijection σ from a set T to itself is also called a *permutation*. We will focus on the case where T is finite and we may simply number the elements of $T = \{1, \ldots, n\}$. Then the set of permutations of T is denoted S_n and called the *symmetric group on n letters* or, alternatively, the *permutation group on n letters*. An element σ of this group may be explicitly presented in such a way that we see where each element is taken to by σ:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

It is easy to see that there are $n! = n(n-1)(n-2)\cdots 1$ different permutations of n letters. There are n choices for $\sigma(1)$; subsequently there remain $n-1$ choices for $\sigma(2)$, $n-2$ choices for $\sigma(3)$ and so on until our hand is forced for $\sigma(n)$. So $|S_n| = n!$. Let us now consider the concrete example of the permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 7 & 6 & 4 & 5 & 2 & 3 \end{pmatrix},$$

which is an element of S_7. Notice that the elements 1, 4, and 5 are untouched by the permutation and the remaining four items are permuted in the cycle $2 \mapsto 7$, $7 \mapsto 3$, $3 \mapsto 6$, and $6 \mapsto 2$. This suggests a more concise notation for σ, since this cycle $(2\,7\,3\,6)$ completely determines the permutation. Notation-wise, $(2\,7\,3\,6)$ denotes exactly the same cycle as $(3\,6\,2\,7)$; to avoid confusion, we usually select the one that starts with the smallest number. Not every permutation is a cycle; for example the permutation

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

cannot be represented as a single cycle. However, it contains two cycles, (1 4) and (2 3), so we can simply represent τ as (1 4)(2 3). When two cycles permute disjoint sets of elements, the cycles commute, so (1 4)(2 3) = (2 3)(1 4). Every permutation can be represented as a product of disjoint cycles. The identity permutation can be represented as a singleton cycle, e.g (1), but we will often simply denote it e or (). A cycle that moves k elements is called a k-cycle; a 2-cycle is also called a *transposition*.

Every cycle can be expressed as a product of transpositions in the following manner:

$$(x_1\, x_2\, x_3\, \ldots\, x_{n-1}\, x_n) = (x_1\, x_2)(x_2\, x_3)\cdots(x_{n-1}\, x_n)$$

A consequence of this is that every permutation $\sigma \in S_n$ can be expressed as a product $\sigma = \tau_1\tau_2\cdots\tau_k$ of transpositions. We can classify permutations as *even* or *odd* depending on the parity of k. In particular, the identity permutation () is a product of 0 cycles, so it is even. Thus the set of even permutations of n elements has a group structure. It is called the *alternating group on n letters* and is denoted A_n. Because the set A_n is a a subset of S_n and A_n is a group under the same binary operation that defines S_n, A_n is an example of a subgroup.

1.5. Subgroups

Let G be a group. We call a nonempty subset $H \subseteq G$ a *subgroup* and write $H \leq G$ provided that

a) The set H is closed under the multiplication operation of G.

b) Whenever H contains an element $a \in G$, H contains its inverse a^{-1} as well.

It is immediate from these requirements that a subgroup H contains the identity element. Since H is nonempty, it contains an element h as well as its inverse h^{-1}. Then from closure of multiplication we conclude that $hh^{-1} = e \in H$.

We turn our attention to S_3. This group is not commutative, since the elements $\sigma = (1\,2\,3)$ and $\tau = (1\,2)$ do not commute. Multiplying $\sigma\tau$ we get (1 3) whereas the product $\tau\sigma = (2\,3)$.

Note that for $k \leq n$, $S_k \leq S_n$ because we can simply fix the letters $k+1, k+2, \ldots, n$. An easy corollary, then, is that S_n does not commute for $n \geq 3$. This is because $S_3 \leq S_n$ and we can simply take σ and τ as elements of the larger group S_n that do not commute.

Another example of a subgroup is the set of 2×2 matrices that stabilise the line $y = 0$ (vectors lying on this line remain on this line after transformation). In terms of matrices, this set looks like

$$S = \left\{ \begin{pmatrix} a & c \\ 0 & d \end{pmatrix} : ad \neq 0 \right\}.$$

Showing that this set is closed is a simple matter of computing

$$\begin{pmatrix} a & c \\ 0 & d \end{pmatrix} \begin{pmatrix} a' & c' \\ 0 & d' \end{pmatrix} = \begin{pmatrix} aa' & ac' + cd' \\ 0 & dd' \end{pmatrix},$$

and observing that the determinant of this matrix is nonzero since all of a, a', d, d' were assumed to be nonzero. (Closure under inversion is also easily derived.)

It is not always easy to characterise all subgroups of a given group. For the additive group of integers, however, the following proposition does so nicely.

Proposition S. *The subgroups of* $\mathbf{Z}$ *under addition are precisely given by* $b\mathbf{Z}$*, where* b *is a fixed integer.*

Proof. First, we fix an integer b and show that $b\mathbf{Z}$ is a subgroup. Adding two integers $bm + bn$ gives a third integer $b(m + n)$, which is also in $b\mathbf{Z}$, so the set is closed under the operation of addition; likewise $-(bm) = b(-m)$, so the set is closed under additive inverse.

Now we must show that these $b\mathbf{Z}$ are all the possible subgroups. Let $H \leq \mathbf{Z}$. It is possible that H contains only the identity 0, in which case $H = 0\mathbf{Z}$. If not, let b be the smallest positive integer contained in H. We know from closure that every multiple of b is in H, so $b\mathbf{Z} \subseteq H$. Now let h be any element in H. By the Euclidean division algorithm, we can divide h by b to get $h = mb + r$, where mb is some multiple of b and r is a remainder lying in the range $0 \leq r < b$. Since $r = (-mb) + h$, we have $r \in H$. But then

necessarily we have $r = 0$, since b is the smallest positive integer in H. So h is an integer multiple of b and we have shown that $H \subseteq b\mathbf{Z}$. ∎

Finally, we introduce a specific class of subgroup. If G is a group with an element g, the *cyclic subgroup* generated by g is the set

$$\langle g \rangle = \{g^m : m \in \mathbf{Z}\}.$$

This is a subgroup because $g^m g^n = g^{m+n}$ and $(g^m)^{-1} = g^{-m}$. Note that not all powers are distinct! For example, in the group S_3, the cyclic subgroup generated by τ contains only the identity element and τ itself. If $g^m = e$ and m is the smallest positive integer for which this holds, we say that the *order* of g is m and write $|g| = m$. If no such m exists, then we say g has infinite order.

If G is a group containing an element g such that $\langle g \rangle = G$, then G is called *cyclic* or *singly-generated*. More generally, if $S \subseteq G$ and we let S^{-1} denote the set $\{s^{-1} : s \in S\}$, then the set

$$\langle S \rangle = \{s_1 s_2 \cdots s_n : n \in \mathbf{N}, s_i \in S \cup S^{-1}\}$$

is a subgroup of G, called the subgroup generated by S. If it so happens that $\langle S \rangle = G$, then we way that S is a *generating set* of G.

2. NORMAL SUBGROUPS AND HOMOMORPHISMS

2.1. Cosets and Normal Subgroups

Let H be a subgroup of a group G. For any $g \in G$, we can form a *left coset* of H by multiplying every element of H by g on the left:

$$gH = \{gh : h \in H\}$$

Symmetrically, we could form a *right coset* Hg of H. Now we prove that the cosets partition the group G.

Lemma P. *Let H be a subgroup of a group G. Then every element of G is in exactly one left coset of H.*

Proof. Let $g \in G$ be given. We commence by noting that since H contains the identity element e, the element g is in at least one coset, namely gH. Now we show that any two cosets are either disjoint or equal. Suppose aH and bH are two cosets that are not disjoint. This implies that there exist $h_1, h_2 \in H$ such that $ah_1 = bh_2$. We can manipulate this identity to obtain $a^{-1}b = h_1 h_2^{-1}$. So $a^{-1}b \in H$. Then from closure properties of subgroups, the cosets must be equal, since

$$aH = a(a^{-1}bH) = (aa^{-1})bH = bH.$$

Thus g is in exactly one coset of H. ∎

Any two cosets have the same size, namely the size of H, since the for any coset aH, the function $f : H \to aH$ given by $h \mapsto ah$ establishes a bijection. This gives us some information about H, since it induces a partition of G into equally-sized disjoint cosets. In fact, we have just proved the following famous theorem:

Theorem L (*Lagrange*). *Let H be a subgroup of a group G. Then the order of H must divide the order of G.* ∎

Thus the value $|G|/|H|$ is an integer; it is called the *index* of H in G and denoted $[G : H]$. Lagrange's theorem also tells us that, for any element g of a group G with order n, the order of the cyclic subgroup $\langle g \rangle$ must divide n, so the order of the element g divides n. Consequently, we know $g^n = e$ for any element $g \in G$.

We now know that a subgroup H can partition a group G in two ways, namely into left cosets of the form aH or into right cosets of the form Ha. Remark that in general, the left cosets *do not equal* the right cosets. As an example, let $H = \{(), (1\,2)\} \leq S_3$. Then we have

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

However, there do exist subgroups whose left cosets equal their right cosets. If N is a subgroup of a group G with the property that $aN = Na$ for all $a \in G$, then we call N a *normal* subgroup. Normal subgroups have the property that the product of two cosets is itself a coset:

$$aH \cdot bH = a \cdot Hb \cdot H = a \cdot bH \cdot H = (ab) \cdot HH = (ab)H$$

Now we introduce the concept of conjugacy, which is closely intertwined with normal subgroups. For elements a and b of a group G, we say a is *conjugate* to b if there exists an element $g \in G$ such that $gag^{-1} = b$. It is easy to see that this relation is reflexive (with $g = e$) and symmetric (replace any g with its inverse), so it makes sense to say that a and b are conjugate elements of a group. We can also prove transitivity: If there exists $g \in G$ such that $gag^{-1} = b$ and $h \in G$ such that $hbh^{-1} = c$, then $hgag^{-1}h^{-1} = (hg)a(hg)^{-1} = c$. Thus conjugacy is an equivalence relation that partitions a group into *conjugacy classes*.

Given a group G and any set S of elements from G, we can conjugate the entire set by an element:

$$gSg^{-1} = \{gsg^{-1} : s \in S$$

When the set S is a subgroup, the conjugate set will also be a subgroup, but not necessarily the same subgroup S. It turns out that we can obtain an equivalent definition of normal subgroups in terms of conjugacy.

Lemma C. *Normal subgroups are exactly the subgroups that are stable under conjugation.*

Proof. Let H be a normal subgroup of a group G. Let $g \in G$ and $h \in H$ be arbitrarily chosen. Since $gH = Hg$, there is an element $h' \in H$ such that $gh = h'g$. Then we see that $ghg^{-1} = h'$, an element in H. So we conclude that normal subgroups are stable under conjugation.

On the other hand, suppose H is a subgroup such that, for any $g \in G$ and $h \in H$, $ghg^{-1} \in H$. Then $gHg^{-1} \subseteq H$ and we can conclude that $gH \subseteq Hg$. But g^{-1} is also an element of G, so we can apply our hypothesis to get that $g^{-1}Hg \subseteq H$ and $Hg \subseteq gH$. So $gH = Hg$ for all $g \in G$ and H is normal. ∎

Let G be a group, H a subgroup of G, and let S be a non-empty set of elements of G. The set of $h \in H$ such that $hSh^{-1} = H$ is called the *normaliser of S in H*, denoted $N_H(S)$. The set of $h \in H$ such that $hsh^{-1} = s$ for all $s \in S$ is called the *centraliser of S in H*, denoted $Z_H(S)$. It is easy to check that these are both subgroups of H. Note that if S consists only of one element, then the centraliser and normaliser are equal and the inclusion $Z_H(S) \subseteq N_H(S)$ always holds. If $H = G$, we may omit the subscript and simply call the subgroups the normaliser and centraliser of S. It is clear that a subgroup H of G is normal if and only if $N(H) = G$.

Theorem N. *Let S be a set of elements of a group G. If $H \leq G$, the index in H of the normaliser of S in H, $[H : N_H(S)]$, is exactly the number of conjugates of S under H.*

Proof. Write $N = N_H(S)$ for short and suppose that H is the disjoint union of cosets

$$H = N \cup h_1N \cup \cdots \cup h_kN,$$

where $k = [H : N_H(S)]$. Now suppose that h_i and h_j are elements of H such that $h_iSh_i^{-1} = h_jSh_j^{-1}$. This is true if and only if $S = (h_j^{-1}h_i)S(h_j^{-1}h_i)^{-1}$. So $h_j^{-1}h_i$ is in $N_H(S) = N$ and $h_i \in h_jN$. So two conjugates of S are the same if and only if the conjugating elements belong to the same left coset of N. Thus $[H : N_H(S)]$ is exactly the number of conjugates of S under H. ∎

2.2. Simplicity of the Alternating Group

We take a brief detour to prove a famous result in the classification of groups, using only the basic techniques we have learned so far. A group G is *simple* if its only normal subgroups are the trivial group $\{e\}$ and the whole group G. In this section we show that for $n \geq 5$, the alternating group A_n is simple. First we establish that the set of 3-cycles generate A_n.

Lemma G. *Every permutation in A_n can be written as a product of 3-cycles.*

Proof. Since every permutation in A_n can be written as a product of an even number of transpositions, it suffices to show that any product $(a\,b)(c\,d)$ can be rewritten as a product of 3-cycles. There are three cases to consider. Firstly, if the transpositions have two numbers in common, then $(a\,b) = (c\,d)$ and the product is simply the identity permutation. If the transpositions have a single point in common, then the product is a 3-cycle, since $(a\,b)(b\,c) = (a\,b\,c)$. Lastly, in the case that all four points are distinct, we have

$$(a\,b)(c\,d) = (a\,b\,c)(b\,c\,d),$$

a product of 3-cycles. Hence the set of 3-cycles generates A_n. ∎

In fact, we can shrink the generating set to 3-cycles of a specific form.

Lemma H. *The $n-2$ permutations $(1\,2\,3), (1\,2\,4), \ldots, (1\,2\,n)$ generate A_n.*

Proof. By Lemma G, we need only show that any 3-cycle can be rewritten as a product of cycles $(1\,2\,k)$. Since the square of $(1\,2\,k)$ is $(2\,1\,k)$, any cycle that contains both 1 and 2 can be generated by the permutations above. Then, note that any cycle $(a\,b\,c)$ that fixes 1 can be rewritten $(1\,a\,b)(1\,b\,c)$. Now we can rewrite each of these cycles as follows:

$$(1\,j\,k) = (1\,2\,k)(1\,2\,k)(1\,2\,j)(1\,2\,k)$$

So every 3-cycle can be expressed as a product of cycles of the form $(1\,2\,k)$. ∎

This result can be used to show that any normal subgroup of A_n that contains a 3-cycles is necessarily the whole group A_n.

Lemma K. *Let $n \geq 4$ and let N be a normal subgroup of the group A_n. If N contains a 3-cycle, then $N = A_n$.*

Proof. Without loss of generality, suppose that N contains the cycle $(1\,2\,3)$. Then it also contains its inverse $(2\,1\,3)$ and for any $\sigma \in A_n$,

$$\sigma(2\,1\,3)\sigma^{-1} \in N.$$

So let $\sigma = (1\,2)(3\,k)$ for $k \geq 4$ and we have

$$\sigma(2\,1\,3)\sigma^{-1} = (1\,2\,k).$$

By Lemma H, these generate A_n, so N is the entire group. ∎

We may finally prove that A_n is simple for $n \geq 5$. The following theorem is due to E. Galois.

Theorem S. *Let $n \geq 5$ and let $N \geq A_n$ be non-trivial. Then $N = A_n$.*

Proof. Let $\tau \in N$ be a non-identity permutation that fixes as many symbols as possible. We will show, by contradiction, that τ must be a 3-cycle, i.e. it fixes all symbols except three.

Suppose that τ moves more than three symbols. Then either τ has a cycle with more than three symbols, or τ consists of disjoint transpositions. Without losing generality, we use the symbols 1, 2, 3, 4, and 5. In the first case, we have

$$\tau = (1\,2\,3\,\ldots)\ldots$$

and we can conjugate by $(3\,4\,5)$ to get

$$\tau' = (3\,4\,5)\tau(3\,4\,5)^{-1} = (1\,2\,4\,\ldots)\ldots.$$

Then $\tau^{-1}\tau'$ fixes 1 where τ did not, a contradiction.

In the case that τ only contains transpositions, we have

$$\tau = (1\,2)(3\,4)\ldots,$$

and again we conjugate by $(3\,4\,5)$ to get

$$\tau' = (3\,4\,5)\tau(3\,4\,5)^{-1} = (1\,2)(4\,5)\ldots)\ldots.$$

Then $\tau^{-1}\tau'$ fixes 1 and 2 and τ does not, leading to a contradiction.

The contradiction in both cases implies that τ moves three symbols or less. In the alternating group, this implies that τ is a 3-cycle. So N contains a 3-cycle and by Lemma K, $N = A_n$. ∎

With this theorem in hand, it becomes easy to check that A_4 is the only alternating group that possesses a normal subgroup.

2.3. Isomorphisms and Homomorphisms

Consider the group $G_1 = \{\pm 1, \pm i\}$ under complex multiplication alongside the group $G_2 = \langle \rho \rangle \leq S_4$, where ρ is the permutation that takes $1 \mapsto 2$, $2 \mapsto 3$, $3 \mapsto 4$, and $4 \mapsto 1$. The groups have the following multiplication tables:

G_1:

	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

G_2:

	e	ρ	ρ^2	ρ^3
e	e	ρ	ρ^2	ρ^3
ρ	ρ	ρ^2	ρ^3	e
ρ^2	ρ^2	ρ^3	e	ρ
ρ^3	ρ^3	e	ρ	ρ^2

It doesn't take long to realise that these two multiplication tables are the same, up to relabeling $1 \equiv e$, $i \equiv \rho$, $-1 \equiv \rho^2$, and $-i \equiv \rho^3$. This is an example of the concept of isomorphisms between groups.

Formally, an *isomorphism* is a bijection $f : G \to G'$ from a group to another such that

$$f(x \cdot y) = f(x) \cdot f(y).$$

The multiplication on the left-hand side is taking place in G while the right-hand multiplication takes place in G'. In the above example, the function $f : G_1 \to G_2$ that takes $i^k \mapsto \rho^k$ for $k = 0, 1, 2, 3$ gives an explicit isomorphism. If there exists an isomorphism between two groups G and H, we say that they are *isomorphic* and write $G \simeq H$.

Any two cyclic groups of order n are isomorphic. We will not formally prove this, but it is clear that if G_1 and G_2 are cyclic groups of the same order generated by g_1 and g_2 respectively, then the function $f : G_1 \to G_2$ given by $f({g_1}^k) = {g_2}^k$ for any integer k will be a well-defined isomorphism.

As a perhaps surprising example, the group of real numbers under addition and the group of positive real numbers under multiplication are isomorphic to one another. The function $f(x) = e^x$ is a bijection from $\mathbf{R}$ to $\mathbf{R}^+$ and $e^{x+y} = e^x e^y$ for real numbers x and y.

If two groups G_1 and G_2 are isomorphic, then

a) The groups have the same order, i.e. $|G_1| = |G_2|$.

b) Either G_1 and G_2 are both abelian or they are both non-abelian.

c) Both groups have the same number of elements of every order.

These properties are useful in showing that two groups are not isomorphic. For example, S_3 has no element of order 6, so it is not isomorphic to the cyclic group of order 6, which has two such elements: 1 and 5.

We may obtain a generalisation of an isomorphism by relaxing the requirement that the function be bijective. Any map $f : G_1 \to G_2$ between groups that satisfies $f(x \cdot y) = f(x) \cdot f(y)$ for all $x, y \in G_1$ is called a *homomorphism*, and an isomorphism is simply a bijective homomorphism. It follows from the definition that any homomorphism maps the identity of the first group to the identity of the second, and that inverses are mapped to inverses. The simplest example of a homomorphism is the trivial homomorphism that maps every element of G_1 to the identity of G_2. Another homomorphism that is not an isomorphism is the map from $\mathbf{Z}$ to S_2 that takes all even integers to the identity and all odd integers to the permutation that switches 1 and 2.

2.4. The Isomorphism Theorems

Let $f : G \to H$ be a homomorphism. The *kernel* of f, denoted $\ker(f)$, is a the set of all elements in the G that map to the identity e_H of the target group H. The *image* of f, denoted $\mathrm{Im}(f)$, is the set of all elements in H that equal $f(g)$ for some $g \in G$. In particular, the homomorphism f is an isomorphism if and only if $\ker(f) = \{e\}$ and $\mathrm{Im}(f) = G$. It is an easy exercise to verify that $\ker(f)$ and $\mathrm{Im}(f)$ are always subgroups of G and H respectively.

Now we want to describe the preimages of $\mathrm{Im}(f)$. For each element $h \in \mathrm{Im}(f) \leq H$, the subset $\{g : f(g) = h\}$ is its *fibre*. The following lemma establishes that the kernel of a homomorphism is normal and relates the fibres of the image to cosets of the kernel. Since the image of a homomorphism is a subgroup, we can assume that the homomorphism is surjective without losing generality.

Lemma F (*Fibre lemma*). *Let $f : G \to H$ be a surjective group homomorphism with kernel K. Then for any $h \in H$ and $g \in G$ that satisfies $f(g) = h$, we have $f^{-1}(h) = gK = Kg$, implying that K is a normal subgroup of G.*

Proof. Let $h \in H$ and let X denote $f^{-1}(h)$, the fibre of h. Now take $g \in G$ such that $f(g) = h$. Since $f(k) = e_H$ for all $k \in K$, we have $f(sk) = f(s)f(k) = h = f(k)f(s) = f(ks)$. So $sK \subseteq X$ and $Ks \subseteq H$. To prove the reverse inclusions, we let $x \in X$ (so $f(x) = h$). Then since f is a homomorphism, we have $f(g^{-1}) = h^{-1}$ and

$$f(xg^{-1}) = f(x)f(g^{-1}) = hh^{-1} = e_H = h^{-1}h = f(g^{-1})f(x) = f(g^{-1}x).$$

This implies that $xg^{-1} \in K$, so multiplying on the right by g, we find that $x \in Kg$. Similarly, $g^{-1}x \in K$ so $x \in gK$. ∎

Now for any surjective map f from sets G to H, we can define an equivalence relation on elements $a, b \in G$:

$$a \sim b \quad \text{if and only if} \quad f(a) = f(b)$$

The set of equivalence classes is denoted $G/\!\sim$ and there exists a bijection $f' : G/\!\sim \to H$. Now if f is a group homomorphism, then we can pull back the group structure of H to $G/\!\sim$ using the bijection. Then the set

$$G/K = \{gK : g \in G\}.$$

has a group structure and is called the *quotient group*. Multiplication in the quotient group is given by $gK \cdot hK = (g \cdot h)K$, which is well-defined because K is normal. So the group structure of G/K depends only on the kernel K and not the homomorphism that induced it. The following theorem sums up what we have discovered so far.

Theorem I (*First Isomorphism Theorem*). *Let $f : G \to H$ be a surjective group homomorphism. Then $K = \ker(f)$ is a normal subgroup of G and there exists a group isomorphism $f' : G/K \to H$.* ∎

This result is fundamental and is often used in conjunction with other laws of homomorphisms to establish isomorphisms. The other two isomorphism theorems provide further insight into the properties of group homomorphisms.

Theorem J (*Second Isomorphism Theorem*). *Let G be a group with subgroups A and B. Suppose that $A \leq N_G(B)$. Then AB is a group and B is a normal subgroup of AB. Then there exists a surjective homomorphism $f : A \to AB/B$ with kernel $A \cap B$, given by $f(a) = aB$.*

Proof. First we show that AB is a group. Since $A \leq N_G(B)$, for any product $a_1b_1a_2b_2$, we can find $b' \in B$ such that $a_2b' = b_1a_2$. Then $a_1b_1a_2b_2 = a_1a_2b'b_2 \in AB$. Similarly, AB is stable under inverses, since for any product ab, we can find $b' \in B$ such that $(ab)^{-1} = b^{-1}a^{-1} = a^{-1}b'$. Because both A and B are subgroups of $N_G(B)$, so is AB and in particular, B is a normal subgroup of AB. Then $f : A \to AB/B$ is the restriction of the quotient map $AB \to AB/B$ to A, so it is surjective.

Now let $a \in \ker(f)$. By definition, $f(a) = aB = B$, so $a \in B$. Thus $a \in A \cap B$. Now let $a \in A \cap B$. Then $aB = B$ so $a \in \ker(f)$. Thus $\ker(f) = A \cap B$. In fact, we can now apply the first isomorphism theorem to deduce that $A/\ker(f) \cong AB/B$. ∎

Theorem K (*Third Isomorphism Theorem*). *Suppose that $A \leq B \leq G$ is a chain of subgroups and furthermore, both A and B are normal subgroups of G. Then the kernel of the surjective homomorphism $f : G/A \to G/B$ is B/A.*

Proof. The homomorphism f is well defined, since if $sA = tA$, then $t^{-1}s \in A$, implying that $t^{-1}s \in B$ and $sB = tB$. It is also surjective, since for any $sB \in G/B$, we have $sA \in G/A$ and $f(sA) = sB$.

Now let $sA \in \ker(f)$. We have $f(sA) = sB = B$, so $s \in B$. This means that $sA \in B/A$. Conversely, if $s \in B$ then $f(sA) = sB = B$. So $\ker(f) = B/A$ and by the first isomorphism theorem, we can express this as the isomorphism $(G/A)/(B/A) \cong G/B$. ∎

[Include correspondence theorem.]

3. AUTOMORPHISMS

An isomorphism from a group to itself is called an *automorphism*. Given a group G, we can construct a set $\mathrm{Aut}(G)$ of all automorphisms of G and in fact, it is easily verifiable that $\mathrm{Aut}(G)$ is a group under function composition.

3.1. Inner and Outer Automorphisms

For any element a of a group G, the map f_a defined by

$$f_a(s) = asa^{-1}$$

is an automorphism. Such automorphisms, which conjugate the group by a fixed element, are called *inner* automorphisms. An automorphism that is not inner is called an *outer* automorphism. There is a homomorphism $\phi : G \to \mathrm{Aut}(G)$ given by $a \mapsto f_a$; that is, $f_a f_b = f_{ab}$. This follows from associativity:

$$f_a f_b(s) = a(bsb^{-1})a^{-1} = (ab)s(ab)^{-1} = f_{ab}(s)$$

The image of the homomorphism ϕ is denoted $\mathrm{Inn}(G)$.

The *centre* of a group G is defined as the centraliser of G in itself:

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

Lemma Z. *The centre of G is the kernel of the map from G to $\mathrm{Aut}(G)$ given by $a \mapsto f_a$.*

Proof. Let $a \in Z(G)$ and note that for an arbitrary element $s \in G$,

$$f_a(s) = asa^{-1} = aa^{-1}s = s$$

so f_a is the identity automorphism. Likewise, if some $f_a \in \mathrm{Aut}(G)$ is the identity automorphism, then $f_a(s) = asa^{-1} = s$ for all $s \in G$. So $as = sa$ for all $s \in G$ and $a \in Z(G)$. ∎

From Lemma 2.4.C, we know that normal subgroups are exactly the subgroups that are stable under conjugacy, i.e. stable under inner automorphism. We can extend this into a more general definition. If $H \leq G$ and $f(H) = H$ for all $f \in \mathrm{Aut}(G)$, then we say that H is a *characteristic* subgroup. (Of course, every characteristic subgroup is also normal.)

Lemma C. *The centre $Z(G)$ of a group G is characteristic.*

Proof. It must be shown that if $g \in Z(G)$, then for all $f \in \mathrm{Aut}(G)$, $f(g) \in Z(G)$. We prove this by contraposition. Let $g \in G$ and suppose there exists $f \in \mathrm{Aut}(G)$ such that $f(g) \notin Z(G)$. Then there exists $s \in G$ such that $f(g) \cdot s \neq s \cdot f(g)$. Now we apply the inverse automorphism f^{-1} to both sides to get that $g \cdot f^{-1}(s) \neq f^{-1}(s) \cdot g$. Hence there exists an element of G, namely $f^{-1}(s)$, that g does not commute with, so $g \notin Z(G)$. ∎

By the first isomorphism theorem, $\mathrm{Inn}(G) \simeq G/Z(G)$; in particular, if G is abelian, we have $Z(G) = G$ and $\mathrm{Inn}(G)$ is trivial. Because it arose as the image of a group homomorphism ϕ, we know that $\mathrm{Inn}(G) \leq \mathrm{Aut}(G)$. In fact, we can prove that $\mathrm{Inn}(G)$ is normal in $\mathrm{Aut}(G)$.

Lemma N. *Let G be a group. Then $\mathrm{Inn}(G)$ is a normal subgroup of $\mathrm{Aut}(G)$.*

Proof. Let $f_a \in \mathrm{Inn}(G)$ be the map that conjugates by an element $a \in G$. Let $g \in \mathrm{Aut}(G)$ be arbitrary. Then

$$g \circ f_a \circ g^{-1} = g(a \cdot g^{-1}(s) \cdot a^{-1}) = g(a) \cdot s \cdot g(a)^{-1},$$

and we see that $g \circ f_a \circ g^{-1} = f_{g(a)} \in \mathrm{Inn}(G)$. So $\mathrm{Inn}(G)$ is stable under conjugation in $\mathrm{Aut}(G)$. ∎

Now by the first isomorphism theorem, we have a group $\mathrm{Out}(G) = \mathrm{Aut}(G)/\mathrm{Inn}(G)$, called the *outer automorphism group*. Note that this is not the set of outer automorphisms! (The set of outer automorphisms do not form a group because the identity automorphism is inner.)

?.? Structure of Abelian Groups

[Under construction.]

?.?. Group Actions on Sets

Let G be a group and S a set of elements. We say that G *acts on* S if there exists a map from $G \times S \to S$ (whose pairs (g, s) are written $g(s)$ or gs) such that, for all $s \in S$, $es = s$ and $(gh)s = g\big(hs\big)$ for every $g, h \in G$. For any element $s \in S$, the *orbit* is set of elements in S to which s may be taken by G:

$$Gs = \{gs \in S : g \in G\}$$

If there exists an s such that $Gs = S$, then we say that G acts *transitively* on S. In fact, because the orbits of a group action partition the set S into equivalence classes, transitivity of a group action implies there is only one orbit, and for any $s, s' \in S$, there exists some $g \in G$ such that $gs = s'$.

The *stabiliser* of s is the set of all elements in G that leave s fixed:

$$\mathrm{Stab}(s) = \{g \in G : gs = s\}$$

Consider a model example of a transitive action. Let G be a group with a subgroup $H \leq G$; let S be the set $\{aH : a \in G\}$ of left cosets of H. The group G acts on S by taking a coset aH to gaH. This action is transitive: For any a, a', there exists some g such that $g(a) = a'$, so we can take any aH to $a'H$ via g as well. The closure property of subgroups gives us $G_H = H$, and the stabiliser of an arbitrary coset aH is the set $\{g \in G : gaH = aH\}$. We find that this is exactly the conjugate subgroup aHa^{-1}. More generally, if G acts transitively on a set S and $gs = s'$, then $\mathrm{Stab}(s') = g\,\mathrm{Stab}(s)g^{-1} \subseteq G$.

We can equivalently describe a group action of G on a set S as a homomorphism from G to $\mathrm{Sym}(S)$. If this map is injective, we say the action is *faithful*. Most group actions that we will study are faithful, and a very simple one is given by the following theorem.

Theorem C (*Cayley*). *Let G be a finite group. Then G is isomorphic to a subgroup of $\mathrm{Sym}(G)$.*

Proof. For every element g of G, let $\sigma_g : G \leftarrow \mathrm{Sym}(G)$ be given by $x \mapsto gx$. The map σ_g is injective, since in G, $g_1 x = g_2 x$ implies that $g_1 = g_2$. For any given $x \in G$, we have $\sigma_g(g^{-1}x) = x$, so the map is also surjective. Then we also have $\sigma_{gh}(x) = (gh)x = g(hx) = \sigma_g(\sigma_h(x)) = (\sigma_g \circ \sigma_h)(x)$, establishing the isomorphism. ∎

The injection from G to $\mathrm{Sym}(G)$ is faithful, since if $\sigma_{g_1} = \sigma_{g_2}$, then $\sigma_{g_1}(e) = g_1 e = g_2 e = \sigma_{g_2}(e)$, and $g_1 = g_2$. We also observe that σ_e is the identity permutation. This action of a group G on itself is called the *left Cayley action* or *left regular action*. We could also have described an isomorphism given by $g \mapsto (x \mapsto g)$, but this does not give a homomorphism since now, $\sigma_{gh} = \sigma_h \circ \sigma_g$. This can be mended by letting σ_g be given by $\sigma_g(x) = xg^{-1}$. This is called the *right regular action*.

Suppose that a group G acts on itself by conjugation:

$$s \mapsto gsg^{-1}.$$

Then the orbit of an element s is its conjugacy class and its stabiliser is its centraliser. Let $\overline{H}$ denote the set of all subgroups of G. Then G also acts on $\overline{H}$ by conjugation: $g(H) = gHg^{-1}$. In this action, the stabiliser of a subgroup H is $N(H)$, its normaliser.

?.?. The Sylow Theorems

Lagrange's Theorem, presented in Section 2.4, states that if G is a group with $|G| = n$, then the order of any element of G divides n. The converse of this theorem does not hold: There does not necessarily exist a subgroup with size m for all m dividing n. For example, A_4 has no subgroup of order 6. However, the results of this section provide a partial converse to Lagrange's Theorem. To this end, we first introduce the concept of double cosets.

Let H and K be subgroups, not necessarily distinct, of a group G. For a fixed element g of G, the set

$$HgK = \{hgk : h \in H \text{ and } k \in K\}$$

is called a *double coset*.

First we show that, just like ordinary cosets, double cosets partition a group. First, every element g of a group belongs to some double coset of H and K, namely HgK. Then we have the following lemma:

Lemma D. *Let H and K be subgroups of a group G. For elements $x, y \in G$, the cosets HxK and HyK are either equal or disjoint.*

Proof. If HxK and HyK are not disjoint, then we can find $g \in HxK \cap HyK$. Suppose that $g = h_1xk_1 = h_2yk_2$; so $x = h_1^{-1}h_2yk_2k_1^{-1}$ and $y = h_2^{-1}h_1xk_1k_2^{-1}$. Then for any $hxk \in HxK$,

$$hxk = hh_1^{-1}h_2yk_2k_1^{-1}k,$$

meaning that $HxK \subseteq HyK$. Likewise,

$$hyk = hh_2^{-1}h_1xk_1k_2^{-1}k$$

and $HyK \subseteq HxK$. ∎

A double coset HgK contains left cosets of K that are of the form $(hg)K$ as well as right cosets of H of the form $H(gk)$. In fact, we can calculate the exact quantity of such left and right cosets.

Lemma Q. *The number of right cosets of H in HgK is $[gKg^{-1} : H \cap gKg^{-1}]$ and the number of left cosets of K in HgK is $[H : H \cap gKg^{-1}]$.*

Proof. We form a bijection $f : HgK \to HgKg^{-1}$ that takes an element hgk to $hgkg^{-1}$. This creates a correspondence between left cosets $h(gKg^{-1})$ of gKg^{-1} and left cosets $(hg)K$ of K as well as a correspondence between right cosets $H(gkg^{-1})$ of H and right cosets $H(gk)$ of H in HgK.

Let $D = (H \cap gKg^{-1}) \leq gKg^{-1}$. Then gKg^{-1} can be written as a union of disjoint right cosets

$$gKg^{-1} = D \cup Dx_2 \cup \cdots \cup Dx_n,$$

where $n = [gKg^{-1} : D]$. Each x_i belongs to gKg^{-1} and the claim is that $H, Hx_2, \ldots, Hx_n$ are exactly the right cosets of H that are in HgK. These cosets are distinct since if $Hx_i = Hx_j$, then $x_ix_j^{-1} \in H$, but since both x_i and x_j are in gKg^{-1}, this means that $x_ix_j^{-1} \in D$ and $Dx_i = Dx_j$, contradicting our choice of coset representatives. Now every right coset of H in $HgKg^{-1}$ is of the form Hx, where $x = dx_i$ for some $d \in D$. But since $D \subseteq H$, $Hx = Hdx_i = Hx_i$. So the number of right cosets of H in $HgKg^{-1}$ is $[gKg^{-1} : D] = [gKg^{-1} : H \cap gKg^{-1}]$ and by the bijection above, this is the number of right cosets of H in HgK.

In a similar manner, it can be shown that the number of left cosets of gKg^{-1} in $HgKg^{-1}$ is $[H : D] = [H : H \cap gKg^{-1}]$ and this is exactly the number of left cosets of K in HgK. ∎

So much for double cosets. We saw earlier that if n is the order of a group G and d divides n, there does not necessarily exist a subgroup of order d. However, if d is prime or a power of a prime, then there is such a subgroup. We begin with a useful theorem, due to Cauchy.

Theorem C (*Cauchy*). *If G is a group that has order dividing a prime p, then G contains an element of order p.*

Proof. Let G be a group with order mp, where p is prime. The proof is by strong induction on m. If $m = 1$, then G is cyclic, generated by an element of order p.

Now suppose that the theorem holds for all smaller multiples of p. Then if G contains a proper subgroup H whose index $[G : H]$ is not divisible by p, then the order of H is divisible by P and, by induction, H contains an element of order p. In the case that all proper subgroups of G have indices divisible by p, we split G into a union of disjoint conjugacy classes:

$$G = \{e\} \cup C_2 \cup \cdots \cup C_k$$

Suppose that each conjugacy class C_i has order $n_i \geq 1$. Let x_i be a representative from each C_i. Then $n = n_1 + \cdots + n_k$, where, by Theorem 1.4.N, $n_i = [G : Z(\{s_i\})]$. If $n_i \neq 1$, then n_i is the index of a proper subgroup of G, and by hypothesis is divisible by p. Surely $n_1 = 1$, so the number of distinct i for which $n_i = 1$ divides p. But $n_i = 1$ if and only if $s_i \in Z(G)$, so the centre of G has order dividing p. By the Fundamental Theorem of Finite Abelian Groups, this implies that it contains an element of order p. ∎

Cauchy's Theorem tells us that if p divides the order of G, at least one subgroup of G has order p. The first of Sylow's theorems proves a stronger result.

Theorem S (*First Sylow Theorem*).

REFERENCES

The contents of this document are heavily based on the following three sets of lectures: Math 122 given by Benedict Gross at Harvard University, Fall 2003; MATH 235 given by Dani Wise at McGill University, Fall 2018; and MATH 456 given by Mikaël Pichot at McGill University, Fall 2019.