Module MA3411: Commentary
Basic Concepts and Results of Group Theory
Michaelmas Term 2009

D. R. Wilkins

Contents

A Commentary: Basic Concepts and Results of Group Theory

A-1 Overview

The first section of the notes for Module MA3411 is a survey of some of the most basic concepts and results of group theory. It is assumed that most of those taking this module have previously taken an introductory course in abstract algebra where most if not all of these matters have been thoroughly discussed. This section of the formal lecture notes for the module is therefore included for the sake of logical completeness, and in order to summarize definitions, concepts and results that ideally should already be familiar to those taking the module.

Some of the basic results discussed only briefly, if at all, in the main lecture notes are treated in greater detail in this commentary.

A-2 The Laws of Indices

Theorem 1.4 states that if x is an element of a group G then (using multiplicative notation) $x^{m+n} = x^m x^n$ and $x^{mn} = (x^m)^n$ for all integers m and n.

Here x^n is defined recursively for positive values of n so that $x^1 = x$ and $x^n = x^{n-1}x$ when $n > 1$. Also x^0 is defined to be the identity element of the group, and x^{-n} is the inverse of x^n for all positive integers n.

A mathematician should be conscious of having worked through a careful proof of the above 'laws of indices' at some stage in her life. A careful proof of the identity x^{m+n} typically breaks down into about seven cases depending on the signs of the integers m, n and $m+n$:—

(i) the case when either $m = 0$ or $n = 0$;

(ii) the case when $m > 0$ and $n > 0$;

(iii) the case when $m < 0$ and $n < 0$;

(iv) the case when $m < 0$, $n > 0$ and $m + n \geq 0$;

(v) the case when $m < 0$, $n > 0$ and $m + n < 0$;

(vi) the case when $m > 0$, $n < 0$ and $m + n \geq 0$;

(vii) the case when $m > 0$, $n < 0$ and $m + n < 0$.

In case (i) the result follows from the fact that x^0 has been defined to be the identity element of the group.

The result in case (ii) follows easily by induction on n, making use of the Associative Law.

The result in case (iii) follows from the previous case on taking inverses. Suppose that $m = -p$ and $n = -q$, where p and q are positive integers. Now $x^{p+q} = x^{q+p} = x^q x^p$. On taking inverses, we see that

$$x^{m+n} = x^{-(p+q)} = (x^{p+q})^{-1} = (x^q x^p)^{-1} = (x^p)^{-1}(x^q)^{-1} = x^m x^n.$$

(Note that, when we take the inverse of a product of elements of a group, where that group is not necessarily Abelian, we obtain the product of the inverses, but with the order of the corresponding factors reversed.)

It remains to deal with the cases (iv)–(vii) in which m and n have opposite signs. Let p and q be non-negative integers. Now $x^{p+q} = x^p x^q = x^q x^p$. Now x^{-p} and x^{-q} are the inverses of x^p and x^q respectively. By taking these equations for x^{p+q}, multiplying on the left or on the right by x^{-p} or x^{-q}, and applying the Associative Law, we may establish the following identities.

$$x^p = x^{p+q}x^{-q} = x^{-q}x^{p+q}, \quad x^q = x^{p+q}x^{-p} = x^{-p}x^{p+q}.$$

Then, on taking inverses, we find that

$$x^{-p} = x^q x^{-(p+q)} = x^{-(p+q)}x^q, \quad x^{-q} = x^p x^{-(p+q)} = x^{-(p+q)}x^p.$$

Suppose that $m < 0$, $n > 0$ and $m+n \geq 0$ (case (iv)). On setting $p = -m$ and $q = m + n$ we see that

$$x^{m+n} = x^q = x^{-p}x^{p+q} = x^m x^n.$$

Next suppose that $m < 0$, $n > 0$ and $m + n < 0$ (case (v)). On setting $p = -m - n$ and $q = n$ we see that

$$x^{m+n} = x^{-p} = x^{-(p+q)}x^q = x^m x^n.$$

Next suppose that $m > 0$, $n < 0$ and $m + n \geq 0$ (case (vi)). On setting $p = m + n$ and $q = -n$ we see that

$$x^{m+n} = x^p = x^{p+q}x^{-q} = x^m x^n.$$

Finally suppose that $m > 0$, $n < 0$ and $m + n < 0$ (case (vii)). On setting $p = m$ and $q = -m - n$ we see that

$$x^{m+n} = x^{-q} = x^p x^{-(p+q)} = x^m x^n.$$

We have therefore verified in all cases that $x^{m+n} = x^m x^n$.

A-3 The General Associative Law

The *General Associative Law* for a group G ensures that the product of a finite list $x_1, x_2, \ldots, x_m$ of elements of G, where $m \geq 3$ depends only on the order in which those elements occur in that list but is otherwise independent of any bracketing that determines the order in which the product is evaluated.

Thus, for example,

$$((x_1x_2)x_3)(x_4x_5) = x_1((x_2(x_3x_4))x_5)$$

for all elements $x_1, x_2, \ldots, x_5$ of G.

Let us consider how such a product of $x_1, x_2, \ldots, x_m$ is evaluated. The value of this product may be expressed in the form st, where s is a product of $x_1, x_2, \ldots, x_p$ for some $p < m$, t is a product of $x_{p+1}, \ldots, x_m$, and where the listed elements occur in the expressions defining the products s and t in the specified order. Indeed if we evaluate the specified product of $x_1, x_2, \ldots, x_m$ by successively forming products of pairs of elements of G, in accordance with the bracketing that determines the order of evaluation of that product, then the final step involves evaluating the product of elements s and t, where s is formed from $x_1, x_2, \ldots, x_p$ and t is formed from $x_{p+1}, \ldots, x_m$ for some $p < m$: this final evaluation cannot be performed until all evaluations of products needed to evaluate s and t have previously been performed, and therefore this evaluation of the product of s and t is the final step in the evaluation of the product of $x_1, x_2, \ldots, x_m$ in accordance with the specified bracketing in the expression defining that product.

This observation enables us to prove that the General Associative Law holds using the method of proof by induction: we prove that the product formed from m elements of the group is independent of the bracketing that defines the order of evaluation of the product, assuming that the corresponding result holds for all products involving fewer than m elements of the group. In order to implement this, it helps to define a standard order of evaluation for such a product: we then prove that any product of the given elements is equal to the product formed using this standard order of evaluation, provided that the order in which the elements appear in the expression defining the product is preserved.

Given any finite list $x_1, x_2, \ldots, x_m$ of elements of the group G, let us define $\sigma(x_1, x_2, \ldots, x_m)$ by recursion so that $\sigma(x_1) = x_1$ and

$$\sigma(x_1, x_2, \ldots, x_m) = \sigma(x_1, x_2, \ldots, x_{m-1})x_m$$

for $m > 1$. Thus

$$\sigma(x_1) \;=\; x_1,$$

$$\begin{aligned}
\sigma(x_1, x_2) &= x_1 x_2, \\
\sigma(x_1, x_2, x_3) &= (x_1 x_2) x_3, \\
\sigma(x_1, x_2, x_3, x_4) &= ((x_1 x_2) x_3) x_4, \\
\sigma(x_1, x_2, x_3, x_4, x_5) &= (((x_1 x_2) x_3) x_4) x_5, \\
&\text{etc.}
\end{aligned}$$

The key step now is to verify that the identity

$$\sigma(x_1, \ldots, x_p)\sigma(x_{p+1}, \ldots, x_m) = \sigma(x_1, x_2, \ldots, x_m)$$

holds for all $x_1, x_2, \ldots, x_m \in G$, where $m \geq 2$ and $1 \leq p < m$. We prove this result by induction on m. It certainly holds when $m = 2$ (in which case $p = 1$). The definition of $\sigma(x_1, \ldots, x_m)$ also ensures that the identity holds in all cases when $p = m - 1$. Suppose that $p < m - 1$ and that the corresponding identity holds for all lists involving fewer than m elements of G. Now this inductive hypothesis ensures that

$$\sigma(x_1, \ldots, x_p)\sigma(x_{p+1}, \ldots, x_{m-1}) = \sigma(x_1, x_2, \ldots, x_{m-1}).$$

The definition of $\sigma(x_{p+1}, \ldots, x_m)$ ensures that

$$\sigma(x_{p+1}, \ldots, x_m) = \sigma(x_{p+1}, \ldots, x_{m-1})x_m.$$

The group axioms require that that the group operation be associative, and thus require that the Associative Law is valid for all products involving exactly three elements of the group. Combining these facts, we see that

$$\begin{aligned}
\sigma(x_1, \ldots, x_p)\sigma(x_{p+1}, \ldots, x_m) &= \sigma(x_1, \ldots, x_p)\Big(\sigma(x_{p+1}, \ldots, x_{m-1})x_m\Big) \\
&= \Big(\sigma(x_1, \ldots, x_p)\sigma(x_{p+1}, \ldots, x_{m-1})\Big)x_m \\
&= \sigma(x_1, x_2, \ldots, x_{m-1})x_m \\
&= \sigma(x_1, x_2, \ldots, x_m).
\end{aligned}$$

It follows from the Principle of Mathematical Induction that the identity

$$\sigma(x_1, \ldots, x_p)\sigma(x_{p+1}, \ldots, x_m) = \sigma(x_1, x_2, \ldots, x_m)$$

does indeed hold for all $x_1, x_2, \ldots, x_m \in G$, where $m \geq 2$ and $1 \leq p < m$.

We should now have enough in order to verify that the General Associative Law holds for products of any finite number of elements of a group G. The group axioms ensure that the Associative Law holds for products of three elements of the group. Suppose that the General Associative Law is valid

for all products involving less than m elements of the group, where $m > 3$. Let u be a product of m elements of the group, and let these elements be $x_1, x_2, \ldots, x_m$, where these elements are listed in the order in which they occur in the expression defining the product. Then, as explained above, $u = st$, where s is a product of elements $x_1, x_2, \ldots, x_p$, t is a product of elements $x_{p+1}, \ldots, x_m$, and $1 \leq p < m$. The inductive hypothesis ensures that

$$s = \sigma(x_1, x_2, \ldots, x_p) \quad \text{and} \quad t = \sigma(x_{p+1}, \ldots, x_m).$$

But then

$$u = st = \sigma(x_1, \ldots, x_p)\sigma(x_{p+1}, \ldots, x_m) = \sigma(x_1, x_2, \ldots, x_m).$$

Therefore the General Associative Law is valid for products involving any finite number of elements of the group, by the Principle of Mathematical Induction.

Note that this proof of the General Associative Law requires that the Associative Law is valid for products of three elements, but does not in any way require the existence of identity elements or inverses. The proof, and thus the result, is therefore valid for multiplication in *semigroups*: a *semigroup* by definition is a set on which is defined a binary operation which satisfies the Associative Law (for products involving three elements). What we have shown in fact is that if we have an algebraic structure with a binary operation, and if that binary operation satisfies the Associative Law for products of three elements, then it also satisfies the General Associative Law for products of any finite number of elements: the value of such a product is determined by the order in which the factors appear in the product, but is otherwise independent of the manner in which the expresssion defining the product is bracketed in order to determine the order in which operations are applied to evaluate the product.

A-4 Equivalence Relations and Partitions

Equivalence relations appear throughout pure mathematics: in analysis, in algebra, and in topology and geometry. In particular there are many important examples of equivalence relations arising in Group Theory. Anyone engaged in the serious study of mathematics at university level needs to have a good grasp of the basic concepts and results that concern equivalence relations in general.

Let X be a set. An *equivalence relation* $\sim$ on X is a binary relation that is reflexive, symmetric and transitive. A binary relation $\sim$ determines, for each pair x, y of elements of the set, whether or not x is related to y. If x is

related to y, then we denote this by writing $x \sim y$. Examples of relations on the set $\mathbb{R}$ of real numbers include $=$, $\neq$, $<$, $\leq$, $>$, $\geq$. (Thus, for examples, real numbers x and y are related with respect to the relation $<$ if and only if x is strictly less than y, in which case we write $x < y$.)

Of the six relations on the set of real numbers listed above, only Equality ($=$) is an equivalence relation. Let $\sim$ be a relation on a set X. The relation $\sim$ is said to be *reflexive* if $x \sim x$ for all $x \in X$. The relation $\sim$ is said to be *symmetric* when the following property holds: if x and y are elements of the set X, and if $x \sim y$, then $y \sim x$. The relation $\sim$ is said to be *transitive* when the following property holds: if x, y and z are elements of the set X, and if $x \sim y$ and $y \sim z$, then $x \sim z$. All three of these properties must hold if this relation $\sim$ is an equivalence relation.

Any equivalence relation $\sim$ on a set X partitions the set into equivalence classes. The *equivalence class* $[x]$ of some element x of X is defined such that

$$[x] = \{y \in X : y \sim x\}.$$

Moreover the equivalence relation $\sim$ partitions the set X into equivalence classes in such a way that, given any element of the set X, there is exactly one equivalence class to which it belongs. Indeed the fact that the equivalence relation is reflexive ensures that $x \in [x]$ for all $x \in X$. Thus every element of the set belongs to at least one equivalence class. If y and z are elements of the set X determining equivalence classes $[y]$ and $[z]$, and if $[y] \cap [z] \neq \emptyset$, then there exists some element x of X which belongs to the equivalence classes $[y]$ and $[z]$. Then $x \sim y$ and $x \sim z$. The fact that the relation is symmetric ensures that $y \sim x$. The fact that the relation is transitive then ensures that $y \sim z$. Moreover, for each element w of X, $w \sim y$ and and only if $w \sim z$. Thus $[y] = [z]$. We have thus shown that if equivalence classes overlap (so that their intersection is non-empty) then they must coincide. Thus no element of the set can belong to more than one (distinct) equivalence class. Now we have shown that every element of the set belongs to at least one equivalence class, and that no element can belong to more than one equivalence class. It follows that a equivalence relation on a set gives rise to a *partition* of the set. One may regard the equivalence classes determined by the relation as the compartments resulting from the partition. (Sets partitioned in this fashion are thus analogous to chests of drawers: every object stored in the chest is stored in exactly one compartment. In other words it is to be found in exactly one drawer. An alternative metaphor is to regard the compartments of the partition as analogous to the rooms of the house. Every occupant of the house is to be found in exactly one room. Of course we are here supposing that occupants are not to be found in the act of passing through a door from one room to another!)

A-5 Equivalence Relations on Groups

Let G be a group with identity element e. We discuss several examples of equivalence relations defined using the group structure on G.

First let H be a subgroup of G. We recall that $e \in H$, and that $xy \in H$ and $x^{-1} \in H$ for all $x, y \in H$. Define a relation $\sim_H$ on G, where elements $x, y \in G$ satisfy $x \sim_H y$ if and only if $y^{-1}x \in H$.

Now $x \sim_H x$ for all $x \in G$, since $x^{-1}x = e$ and $e \in H$. Thus the relation $\sim_H$ is reflexive.

Suppose that x and y are elements of G satisfying $x \sim_H y$. Then $y^{-1}x \in H$. Now $x^{-1}y = (y^{-1}x)^{-1}$. (Recall that the inverse of a product, is the product of the inverses, but with the order of the factors interchanged.) Also the inverse of any element of H itself belongs to H. Therefore $x^{-1}y \in H$, and thus $y \sim_H x$. This proves that the relation $\sim_H$ is symmetric.

Next suppose that x, y and z are elements of G satisfying $x \sim_H y$ and $y \sim_H z$. Then $y^{-1}x \in H$ and $z^{-1}y \in H$. Now $z^{-1}x = (z^{-1}y)(y^{-1}x)$. Also the product of any two elements of H itself belongs to H. Therefore $z^{-1}x \in H$, and thus $x \sim_H z$. This proves that the relation $\sim_H$ is transitive.

The relation $\sim_H$ is reflexive, symmetric and transitive. It is thus an equivalence relation, and therefore gives rise to a partition of the group G into the corresponding equivalence classes. Given any element of the group G, there is exactly one equivalance class to which it belongs.

We now proceed to investigate the nature of these equivalence classes. Let $[x]_H$ denote the equivalence class of $x \in G$ with respect to the relation $\sim_H$. Then

$$\begin{aligned}[x]_H &= \{y \in G : y \sim_H x\} = \{y \in G : x^{-1}y \in H\} \\ &= \{y \in G : y = xh \text{ for some } h \in H\} = xH.\end{aligned}$$

Thus the equivalence classes determined by the equivalence relation $\sim_H$ are the left cosets of H in G.

The basic result concerning equivalence relations on sets and their associated equivalence classes therefore ensures that, given any element of the group G, there is exactly one left coset of the subgroup H to which it belongs.

The reasons for this are revisited in the proof of Lemma 1.7.

A subgroup H of G also determines a relation $\sim_H^*$ on G, where elements x and y of G satisfy $x \sim_H^* y$ if and only if $xy^{-1} \in H$. One can easily verify that this relation is an equivalence relation: its equivalence classes are the right cosets of H in G. We deduce from this that, given any element of the group G, there is exactly one right coset of the subgroup H to which it belongs.

The right cosets of a subgroup H do not coincide with the left cosets of H unless that subgroup H is normal in G.

Let G be a group, and let e denote the identity element of G. We define a relation $\sim$ on G, where elements x and y of G satisfy $x \sim y$ if and only if there exists some element g of G such that $y = gxg^{-1}$. We say that elements x and y of G are *conjugate* if they are related with respect to this equivalence relation.

Now $x \sim x$ for all $x \in G$, since $exe^{-1} = x$. Thus the relation $\sim$ is reflexive.

Let x and y be elements of G satisfying $x \sim y$. Then there exists $g \in G$ such that $y = gxg^{-1}$. Then $x = g^{-1}yg = kyk^{-1}$, where $k = g^{-1}$, and thus $y \sim x$. This proves that the relation $\sim$ is symmetric.

Let x, y and z be elements of G satisfying $x \sim y$ and $y \sim z$. Then there exist $g, h \in G$ such that $y = gxg^{-1}$ and $z = hyh^{-1}$. Then

$$z = h(gxg^{-1})h^{-1} = (hg)x(g^{-1}h^{-1}) = (hg)x(hg)^{-1},$$

and thus $x \sim z$. This proves that the relation $\sim$ is transitive.

We conclude therefore that the relation $\sim$ of conjugacy is an equivalence relation on any group G.

Now let Subgroup(G) be the set of all subgroups of G. We define a relation $\sim$ on Subgroup(G), where subgroups H and K of G are related if and only if there exists some element g of G such that

$$K = gHg^{-1} = \{ghg^{-1} : h \in H\}.$$

Subgroups that are related in this fashion are said to be *conjugate*. This relation on subgroups of G is also an equivalence relation.

A-6 The Importance of Lagrange's Theorem

Lagrange's Theorem might be regarded as the most fundamental theorem concerned with finite groups. It states that the order of any subgroup of a finite group must divide the order of the group.

A common problem in the theory of finite groups is that of classifying, up to isomorphism, all groups of a given order. Typically one starts out by factorizing the given order as a product of prime numbers. One can then draw conclusions concerning the possible subgroups of a group of that order, utilizing the information provided by the factorization of the order.

If the order of the group has few prime factors, then the number of isomorphism classes tends to be somewhat limited. The most limited case is

that in which the order of the group is a prime number: the group must then be a cyclic group. Possibilities are also limited when the order of the group is $2p$ for some odd prime number p. In that case either the group is a cyclic group or else it is isomorphic to a dihedral group. (Given any integer n, with $n \geq 3$, the dihedral group of order $2n$ is the group of symmetries of a regular n-sided polygon in the plane.)

Other theorems provide information on the structure of a group of a given order. One such theorem is the *First Sylow Theorem*: if p is a prime number, and if p^k is the largest power of p that divides the order of the group, then the group is guaranteed to possess at least one subgroup of order p^k. Such a subgroup is referred to as a *Sylow subgroup*. Other Sylow theorems provide more information about the number and nature of such subgroups.

Lagrange's Theorem is an essential ingredient in the proof of such theorems.

A-7 The Proof of Lagrange's Theorem

Let G be a finite group, and let H be a subgroup of G. Lagrange's Theorem states that the order of H must divide the order of G. In order to prove the theorem one considers the left cosets of H in G. (If the subgroup H is not normal in G, then the left cosets of H will not all be right cosets. But one could construct an equally valid proof replacing left cosets by right cosets throughout.) The crucial observations are the following:—

- every element of G belongs to exactly one left coset of H in G;
- all left cosets of H in G have the same number of elements.

The proof that every element of G belongs to exactly one left coset has been given above, in the discussion of equivalence relations in group theory. It is moreover an immediate consequence of Lemma 1.7.

In order to show that all left cosets have the same number of elements, we make use of the basic result that two finite sets have the same number of elements if there exists a bijective function from one to the other. If x is an element of the group G, then the function from the subgroup H to the left coset xH that sends $h \in H$ to xh is both injective and surjective: it is therefore a bijective function from H to xH. Therefore H and xH must have the same number of elements.

A-8 Normal Subgroups and Quotient Groups

Let N be a normal subgroup of a group G. The subgroup N determines an equivalence relation $\sim_N$ on G, where elements x and y of G satisfy $x \sim_N y$

if and only if $y^{-1}x \in G$. The equivalence classes are the cosets of N in G.

Let x_1, x_2, y_1 and y_2 be elements of the group G. Suppose that $x_1 \sim_N x_2$ and $y_1 \sim_N y_2$. Then $x_2^{-1}x_1 \in N$ and $y_2^{-1}y_1 \in N$. Let us consider whether or not it is the case that $x_1y_1 \sim_N x_2y_2$. We need to determine whether or not $(x_2y_2)^{-1}(x_1y_1) \in N$. Now $(x_2y_2)^{-1}(x_1y_1) = y_2^{-1}x_2^{-1}x_1y_1$. We therefore if it is always the case that $y_2^{-1}x_2^{-1}x_1y_1 \in N$ whenever $x_2^{-1}x_1 \in N$ and $y_2^{-1}y_1 \in N$.

Now in the special case where the group G is Abelian, the element $y_2^{-1}x_2^{-1}x_1y_1$ is the product of the elements $x_2^{-1}x_1$ and $y_2^{-1}y_1$ of N, and must therefore itself belong to N. However this argument is not applicable in the more general case when the group operation on G is not necessarily commutative.

Let us then analyse the structure of the expression $y_2^{-1}x_2^{-1}x_1y_1$ in the non-commutative case. We suppose that $x_2^{-1}x_1 \in N$ and $y_2^{-1}y_1 \in N$. Note that x_2^{-1} and x_1 are adjacent to one another in the expression $y_2^{-1}x_2^{-1}x_1y_1$. Now in order to achieve something useful with the factor y_2^{-1} it would help to introduce a factor y_1 immediately following it in the expression. This can be done by introducing a factor y_1 followed immediately by y_1^{-1} at the appropriate position in the product, since the product $y_1y_1^{-1}$ is the identity element of the group G. On doing this, we find that

$$y_2^{-1}x_2^{-1}x_1y_1 = (y_2^{-1}y_1)(y_1^{-1}x_2^{-1}x_1y_1).$$

Now $y_2^{-1}y_1 \in N$. What about the second factor on the right hand side, which may be written in the form $y_1^{-1}(x_2^{-1}x_1)y_1$? We know that $x_2^{-1}x_1 \in N$. Moreover the definition of a normal subgroup requires that $g^{-1}ng \in N$ for all $n \in N$ and $g \in G$. Thus, because the subgroup N is normal, we can conclude that $y_1^{-1}x_2^{-1}x_1y_1 \in N$. It then follows that $y_2^{-1}x_2^{-1}x_1y_1$ is a product of two elements of the subgroup N, and must therefore belong to this subgroup.

We have thus shown that, if x_1, x_2, y_1 and y_2 are elements of the group G, and if $x_1 \sim_N x_2$ and $y_1 \sim_N y_2$, then $x_1y_1 \sim_N x_2y_2$. Let us denote by $[x]_N$ the equivalence class of an element x of G under the equivalence relation $\sim_N$. Every element of the group G belongs to exactly one of these equivalence classes. Moreover $[x_1]_N = [x_2]_N$ if and only if the elements x_1 and x_2 of G satisfy $x_1 \sim_N x_2$. The result we have proved may therefore be restated as follows: if x_1, x_2, y_1 and y_2 are elements of the group G, and if $[x_1]_N = [x_2]_N$ and $[y_1]_N = [y_2]_N$, then $[x_1y_1]_N = [x_2y_2]_N$. It is an immediate consequence of this result that there is a well-defined binary operation $*$ on the set of equivalence classes, where $[x]_N * [y]_N = [xy]_N$. (The fact that this binary operation $*$ is well-defined is an immediate consequence of the fact that the equivalence class $[xy]_N$ does not depend on the choice of elements x and y representing the equivalence classes $[x]_N$ and $[y]_N$, but is completely determined by these equivalence classes $[x]_N$ and $[y]_N$ themselves.)

Let us denote by G/N the set of these equivalence classes, which are the cosets of N in G. The binary operation $*$ on this set G/N of cosets is associative:

$$\begin{aligned}([x]_N * [y]_N) * [z]_N &= [xy]_N * [z]_N = [(xy)z]_N = [x(yz)]_N = [x]_N * [yz]_N \\ &= [x]_N * ([y]_N * [z]_N)\end{aligned}$$

for all $x, y, z \in G$. The equivalence class of the identity element of G is the subgroup N itself, and this equivalence class is clearly an identity element for the binary operation $*$ on G/N. Also $[x]_N * [x^{-1}]_N$ is the identity element of G/N for every $x \in G$, and thus every coset has an inverse with respect to the binary operation $*$ on cosets.

We may therefore regard the set G/N of cosets of the normal subgroup N as a group in its own right. It is the *quotient* of the group G by the normal subgroup N. The coset of N containing an element x of G may be written as xN. Thus $[x]_N = xN$ for all $x \in G$. Also $(xN)(yN) = (xy)N$. We see therefore that the construction of the quotient group given here using basic properties of equivalence relations and equivalence classes is consistent with that given in the lecture notes for the module.

A-9 Equivalence Relations and Induced Functions

Let X and Y be sets, and let $f: X \to Y$ be a function from X to Y. Let $\sim$ be an equivalence relation on the set X. Then this relation $\sim$ partitions the set X into equivalence classes: given any element of the set X there is exactly one equivalence class to which it belongs; elements x and x' of the set X belong to the same equivalence class if and only if $x \sim x'$. We denote by $[x]$ the equivalence class of an element x of X.

Let $\hat{X}$ denote the set of equivalence classes of elements of X under this equivalence relation $\sim$, and let $q: X \to \hat{X}$. be the function that sends each element x of X to its equivalence class $[x]$. Then elements x and x' of X satisfy $q(x) = q(x')$ if and only if $x \sim x'$.

Under what conditions does the function $f: X \to Y$ induce a function $\hat{f}: \hat{X} \to Y$ with the property that $f = \hat{h} \circ q$? We note that if the function $\hat{f}$ is to be well-defined then we must have $f(x) = f(x')$ for all $x, x' \in X$ satisfying $q(x) = q(x')$. This is a necessary condition for the existence of a function $\hat{f}: \hat{X} \to Y$ for which $f = \hat{f} \circ q$. It turns out that this condition is also sufficient. Suppose that the function $f: X \to Y$ has the property that $f(x) = f(x')$ for all $x, x' \in X$ satisfying $q(x) = q(x')$. Given $\hat{x} \in \hat{X}$, we define $\hat{f}(\hat{x}) = f(x)$, where x is any element of X for which $q(x) = [x] = \hat{x}$. This works, for if x' is another element of X that could have been chosen in

place of X, then $q(x') = q(x)$, and therefore $f(x) = f(x')$. We summarize these observations as follows.

> *Let $\sim$ be an equivalence relation on a set X, let $\hat{X}$ be the corresponding set of equivalence classes, and let $q: X \to \hat{X}$ be the function that maps each element x of X to its equivalence class $[x]$. Let $f: X \to Y$ be a function from X to some set Y. Then f induces a function $\hat{f}: \hat{X} \to Y$ on the set $\hat{X}$ of equivalence classes, where $f = \hat{f} \circ q$, if and only if $f(x) = f(x')$ for all $x, x' \in X$ satisfying $x \sim x'$. Thus the function $f: X \to Y$ induces a corresponding function $\hat{f}: \hat{X} \to Y$ defined on equivalence classes if and only if the function f is constant over each equivalence class.*

The applications of this fundamental principle are ubuquitous throughout pure mathematics.

A-10 Induced Homomorphisms on Quotient Groups

Let G be a group, and let N be a normal subgroup of G. Then N determines an equivalence relation $\sim_N$ on G, where elements x and y of G satisfy $x \sim y$ if and only if $y^{-1}x \in N$. The equivalence classes of the relation are the cosets of N in G: elements x and y of G satisfy $x \sim_N y$ if and only if $xN = yN$. We denote by G/N the quotient group that consists of the set of cosets of N in G, with the operation of multiplication of cosets defined so that $(xN)(yN) = xyN$ for all $x, y \in N$.

Now let $\theta: G \to K$ be a homomorphism from the group G to some group K. The definition of *homomorphism* then requires that $\theta(uv) = \theta(u)\theta(v)$ for all $x, y \in N$. Moreover $\theta(e_G) = e_K$, where e_G and e_K denote the identity elements of the groups G and K respectively. Also $\theta(x^{-1}) = \theta(x)^{-1}$ for all $x \in X$. The *kernel* of the homomorphism θ is the subgroup $\ker\theta$ of G defined by

$$\ker\theta = \{x \in G : \theta(x) = e_K\}.$$

The kernel of any homomorphism is a normal subgroup (see Lemma 1.18).

Now the equivalence classes of the relation $\sim_N$ are the cosets of N in G. It follows that a function defined on G induces a corresponding function on G/N if and only if it is constant over each coset of N in G. We now show that a homomorphism $\theta: G \to K$ has this property if and only if $N \subset \ker\theta$.

Suppose that the homomorphism $\theta: G \to K$ is constant on each coset of N. Then in particular it is constant over the normal subgroup N itself: any normal subgroup is a coset of itself. But the normal subgroup N contains the identity element e_G of G, and $\theta(e_G) = e_K$. It follows therefore that if

the homomorphism $\theta: G \to K$ is constant over each coset of N, then it must map the whole of N to the identity element e_K of the group K, and thus $N \subset \ker \theta$.

Conversely suppose that $\theta: G \to K$ is a homomorphism with the property that $N \subset \ker \theta$. Then θ maps the whole of the normal subgroup N to the identity element e_K of K, and is thus constant over N. The rigidity imposed by the requirement that θ be a homomorphism now ensures that θ is constant over every other coset of N in G. Indeed let x and y be elements of G. Suppose that $y \in xN$. Then $y = xn$ for some $n \in N$. Then $\theta(y) = \theta(x)\theta(n)$, because θ is a homomorphism. But $n \in \ker \theta$, and therefore $\theta(n) = e_K$. It follows that $\theta(y) = \theta(x)$ for all $y \in xN$. Thus if $N \subset \ker \theta$ then the homomorphism $\theta: G \to K$ is constant over each coset of N in G, and therefore induces a function $\hat{\theta}: G/N \to K$, defined on the group G/N of cosets of N in G, where $\hat{\theta}(xN) = \theta(x)$ for all $x \in G$. Moreover this induced function $\hat{\theta}$ is itself a homomorphism. Indeed the group operation on G/N is defined so that $(xN)(yN) = (xy)N$ for all $x, y \in G$. Therefore

$$\hat{\theta}((xN)(yN)) = \hat{\theta}(xyN) = \theta(xy) = \theta(x)\theta(y) = \hat{\theta}(xN)\hat{\theta}(yN).$$

Now a homomorphism is injective if and only if its kernel is the trivial subgroup consisting of the identity element of its domain. In particular the induced homomorphism $\hat{\theta}: G/N \to K$ is injective if and only if the only coset of N in G that gets mapped under θ to the identity element of K is the normal subgroup N itself. Thus $\hat{\theta}: G/N \to K$ is injective if and only if $N = \ker \theta$.

We have therefore established the basic facts concerning induced homomorphisms stated in Proposition 1.19.

A-11 An example of an Induced Homomorphism

Let $\mathbb{R}^2$ denote the set of ordered pairs of real numbers. Then $\mathbb{R}^2$ is an Abelian group under the operation of addition of ordered pairs, where

$$(x, y) + (u, v) = (x + u, y + v)$$

for all $x, y, u, v \in \mathbb{R}$. Let

$$N = \{(x, y) \in \mathbb{R}^2 : x + y = 0\} = \{(x, -x) : x \in \mathbb{R}\}.$$

Then N is a subgroup of $\mathbb{R}^2$. Moreover it is a normal subgroup, because *every* subgroup of an Abelian group is a normal subgroup.

We now identify the cosets of N in $\mathbb{R}^2$. Let $(x_0, y_0) \in \mathbb{R}^2$. Then the coset $(x_0, y_0) + N$ of N in $\mathbb{R}^2$ that contains (x_0, y_0) consists of all ordered pairs

of real numbers that can be expressed in the form $(x_0, y_0) + (u, v)$ for some $(u, v) \in N$. Therefore

$$\begin{aligned}(x_0, y_0) + N &= \{(x_0 + u, y_0 - u) : u \in \mathbb{R}\} \\ &= \{(x, y) \in \mathbb{R} : x + y = x_0 + y_0\}.\end{aligned}$$

Thus the cosets of N in $\mathbb{R}^2$ are the subsets of $\mathbb{R}^2$ that are of the form

$$\{(x, y) \in \mathbb{R}^2 : x + y = c\}$$

for some real number c. If we regard ordered pairs of real numbers as the Cartesian coordinates of points in the plane, then the cosets of N are represented by the diagonal lines with equations of the form $y = c - x$, where c is some real constant. The elements of the quotient group $\mathbb{R}^2/N$ are represented by these diagonal lines. The group operation on this quotient group is the binary operation that sends the lines $x + y = a$ and $x + y = b$ to the line $x + y = a + b$ for all real numbers a and b.

A homomorphism $\theta: \mathbb{R}^2 \to K$ from $\mathbb{R}^2$ to a group K induces a homomorphism on the quotient group $\mathbb{R}^2/N$ if and only if $N \subset \ker \theta$. The homomorphism θ is then constant on each diagonal line of the form $x + y = c$. One such homomorphism is the homomorphism $\theta: \mathbb{R} \to \mathbb{C}^*$ mapping $\mathbb{R}^2$ to the group C^* of non-zero complex numbers under multiplication, where

$$\theta(x, y) = \exp(2\pi i(x + y)) = \cos(2\pi(x + y)) + i \sin(2\pi(x + y)).$$

The kernel of θ is the subgroup

$$\{(x, y) \in \mathbb{R}^2 : x + y \in \mathbb{Z}\}.$$

Note that $N \subset \ker \theta$. The induced homomorphism on $\mathbb{R}^2/N$ maps the coset represented by the diagonal line $x + y = c$ to the complex number $e^{2\pi ic}$ for all real numbers c.

A-12 Group Actions

Let $GL_3(\mathbb{R})$ denote the group of all 3×3 matrices with real coefficients whose determinant is non-zero. The group operation on $GL_3(\mathbb{R})$ is matrix multiplication. The group $GL_3(\mathbb{R})$ acts on the set $\mathbb{R}^3$: we represent elements of $\mathbb{R}^3$ as column vectors, and these column vectors are multiplied on the left by matrices belonging to $GL_3(\mathbb{R})$. This *action* of $GL_3(\mathbb{R})$ on $\mathbb{R}^3$ is thus determined by a function $\lambda: GL_3(\mathbb{R}) \times \mathbb{R}^3 \to \mathbb{R}^3$, where $\lambda(A, \mathbf{v}) = A\mathbf{v}$ for all $A \in GL_3(\mathbb{R})$ and $\mathbf{v} \in \mathbb{R}^3$. Note that $(AB)\mathbf{v} = A(B\mathbf{v})$ for all $A, B \in GL_3(\mathbb{R})$ and $\mathbf{v} \in \mathbb{R}^3$. Also $I\mathbf{v} = \mathbf{v}$ where I is the identity matrix in $GL_2(\mathbb{R})$.

This action of the group $GL_3(\mathbb{R})$ on the set $\mathbb{R}^3$ is a typical example of a left action of a group on a set. Let G be a group with identity element e_G, and let X be a set. A *left action* of the group G on the set X is determined by a function $\lambda\colon G \times X \to X$ with the following properties:

(i) $\lambda(gh, x) = \lambda(g, \lambda(h, x))$ for all $g, h \in G$ and $x \in X$;

(ii) $\lambda(e_G, x) = x$ for all $x \in X$.

It is convenient to denote $\lambda(g, x)$ by $g.x$. We then require that $(gh).x = g.(h.x)$ and $e_G.x = x$ for all $g, h \in G$ and $x \in X$. Given such a left action, we say that the group G *acts* on the set X *on the left*.

There is a corresponding definition for a *right action* of a group G on a set X. Such a right action is specified by a function $\rho\colon X \to G \to X$, where $\rho(x, gh) = \rho(\rho(x, g), h)$ and $\rho(x, e_G) = x$ for all $g, h \in G$ and $x \in X$. In the case of a right action, we can denote $\rho(x, g)$ by $x.g$. We then require that $x.(gh) = (x.g).h$ and $x.e_G = x$ for all $g, h \in X$ and $x \in X$.

The symmetric group Σ_n is defined to be the group of all permutations of the set $\{1, 2, \ldots, n\}$. (A *permutation* of this set is a bijective function from the set to itself.) The group Σ_n acts on the set $\{1, 2, \ldots, n\}$ on the left, where $\sigma.m = \sigma(m)$ for all $\sigma \in \Sigma_n$ and $m \in \{1, 2, \ldots, n\}$.

The group $\mathbb{Z}$ of integers acts on the set $\mathbb{R}$ of real numbers on the left. This action is determined by the function that sends (m, x) to $m + x$ for all $m \in \mathbb{Z}$ and $x \in \mathbb{R}$. We can therefore regard the group $\mathbb{Z}$ of integers as acting on the real line: a positive integer m translates points of the real line to the right by an amount m; a negative integer n translates points to the left by an amount $-m$.

Let G be a group with identity element e_G which acts on a set X on the left. Let the image of $(g, x) \in G \times X$ under the function defining the action be denoted by $g.x$, so that $(gh).x = g.(h.x)$ and $e_G.x = x$ for all $g, h \in G$ and $x \in X$. Then each element g of the group G determines a corresponding function $\lambda_g\colon X \to X$ from the set X to itself, where $\lambda_g(x) = g.x$. Now

$$\lambda_{g^{-1}}(\lambda_g(x)) = \lambda_{g^{-1}}(g.x) = g^{-1}.(g.x) = (g^{-1}g).x = e_G.x = x$$

for all $g \in G$ and $x \in X$. It follows that, for each $g \in G$, the function $\lambda_g\colon X \to X$ is a bijection whose inverse is $\lambda_{g^{-1}}\colon X \to X$.

Now a *permutation* of a set is by definition a bijective function from the set to itself. We see that, whenever a group G acts on a set X on the left, each element of the group determines a corresponding permutation of the set X. Now the set $\mathrm{Symm}(X)$ of permutations of a set X is a group, where the group operation is composition of permutations. One can readily check

that, when a group G acts on a set X on the left, the function that sends each element g of the group G to the corresponding permutation λ_g of X is a group homomorphism from the group G to the group $\mathrm{Symm}(X)$. Conversely every group homomorphism from G to $\mathrm{Symm}(X)$ determines a left action of the group G on the set X.

We see therefore that left actions of a group G on a set X correspond to group homomorphisms from G to the group $\mathrm{Symm}(X)$ of permutations of the set X.

A-13 Actions, Orbits and Stabilizers

Let G be a group with identity element e_G which acts on a set X on the left. This action is determined by a function from $G \times X$ to X sending $(g, x) \in G \times X$ to some element $g.x$ of X, where $(gh).x = g.(h.x)$ and $e_G.x = x$ for all $g, h \in G$ and $x \in X$. Each element g of the group G determines a permutation $\lambda_g \colon X \to X$ of the set X, where $\lambda_g(x) = g.x$ for all $x \in X$.

Let x be an element of X. The *orbit* of x is defined to be the set $\{g.x : g \in G\}$. (Thus an element y of X belongs to the orbit of x if and only if $y = g.x$ for some $g \in X$. The *stabilizer* of x is defined to be S_x, where $S_x = \{g \in G : g.x = x\}$. Now $e_G \in S_x$, because the definition of a group action requires that $e_G.x = x$. Let g and h be elements of S_x. Then $(gh).x = g.(h.x) = g.x = x$, and therefore $gh \in S_x$. Also $g^{-1}.xg^{-1}.(g.x) = (g^{-1}g).x = e_G.x = x$, and therefore $g^{-1} \in S_x$. This proves that the stabilizer S_x of x is a subgroup of G.

Now a left action of a group G on a set X determines a relation $\sim$ on X, where $x \sim y$ if and only if there exists $g \in G$ for which $y = g.x$. This relation is reflexive, since $e_G.x = x$ for all $x \in X$, where e_G denotes the identity element of the group G.

Let x and y be elements of X satisfying $x \sim y$. Then $y = g.x$ for some $g \in G$. But then $x = g^{-1}.y$, and therefore $y \sim x$. This shows that the relation $\sim$ on X is symmetric.

Now let x, y and z be elements of X satisfying $x \sim y$ and $y \sim z$. Then there exist elements g and h of G such that $y = g.x$ and $z = h.y$. But then $z = h.(g.x) = (hg).x$, and therefore $x \sim z$. This shows that the relation $\sim$ on X is transitive.

We have shown that the relation $\sim$ on the set X determined by a left action of a group G on X is reflexive, symmetric and transitive. It is thus an equivalence relation on the set X. It therefore partitions the set into equivalence classes: these equivalence classes are the *orbits* for the action of G on X. An element of X belongs to at most one orbit for this action;

elements x and y of X belong to the same orbit if and only if there exists some element g of the group G for which $y = g.x$.

Let x be an element of X, and let g and h be elements of G. Then

$$\begin{aligned} g.x = h.x &\iff h^{-1}.(g.x) = x \iff (h^{-1}g).x = x \\ &\iff h^{-1}g \in S_x \iff gS_x = hS_x. \end{aligned}$$

Thus $g.x = h.x$ if and only if g and h belong to the same left coset of S_x in X. Thus there is a bijection between the set of left cosets of S_x in G and the orbit of the element x.

Suppose now that the group G is finite. Then the number of left cosets of S_x in G is the index $[G : S_x]$ of S_x in G, where $[G : S_x] = |G|/|S_x|$. We see then that, for each element x of the set X, the number of elements in the orbit of x is equal to the index $[G : S_x]$ of the stabilizer of x in G, and is thus equal to $|G|/|S_x|$.

A-14 Natural Actions of a Group on Itself

There are several natural actions of a group G on itself.

A group G acts on itself by left multiplication: the function that sends (g, x) to gx for all elements g and x of G has all the properties required of a left action.

Similarly there is a right action of a group on itself, where the group acts by right multiplication. Now any right action of a group on a set determines a corresponding left action, where acting on the left by an element g of the group G corresponds to acting on the right by g^{-1}. We now carry this through in order to define a left action of a group on itself. Thus, given any elements g and x of the group G, let us denote xg^{-1} by $g\#x$. Then

$$g\#(h\#x) = g\#(xh^{-1}) = (xh^{-1})g^{-1} = x(h^{-1}g^{-1}) = x(gh)^{-1} = (gh)\#x.$$

It follows that the function that sends (g, x) to xg^{-1} for all $g, x \in G$ defines a left action of the group G on itself.

Another natural action of a group on itself is the *adjoint action*, discussed below.

A-15 Conjugacy and the Adjoint Action

Let G be a group with identity element e_G. Given elements x and y of G, let $\mathrm{ad}_x(y) = xyx^{-1}$. Now $\mathrm{ad}_{e_G}(z) = z$ and

$$\begin{aligned} \mathrm{ad}_x(\mathrm{ad}_y(z)) &= \mathrm{ad}_x(yzy^{-1}) = x(yzy^{-1})x^{-1} = (xy)z(y^{-1}x^{-1}) \\ &= (xy)z(xy)^{-1} = \mathrm{ad}_{xy}(z) \end{aligned}$$

for all $x, y, z \in G$. Therefore the function that sends $(x, y) \in G \times G$ to $\mathrm{ad}_x(y)$ is a left action of the group G on itself.

Now any action of a group on a set determines an equivalence relation on that set, where two elements of the set are related if and only if they belong to the same orbit. The equivalence relation on a group G determined in this fashion by the adjoint action is the relation of conjugacy: elements x and y of the group G are *conjugate* if and only if there exists some element g of G such that $y = gxg^{-1}$. In particular the orbits for the adjoint action are the conjugacy classes. Also the stabilizer of any element x of the group under the adjoint action is the centralizer of x.

We have noted above that, when a finite group acts on a set, the number of elements in the orbit of some element of the set is equal to the index of the stabilizer of that element. We can apply this to the adjoint action: given any finite group G, and given any element h of G, the number of elements in the conjugacy class of h is equal to the index $[G : C(h)]$ of the centralizer $C(h)$ of h in G. (Here $[G : C(h)] = |G|/|C(h)|$.) This result is included in the lecture notes as Lemma 1.21.