Chapter 16

MSM2P2 Symmetry And Groups

(16.1) Symmetry

(16.1.1) Symmetries Of The Square

Definition 1 *A symmetry is a function from an object to itself such that for any two points a and b in the object, the distance between them is preserved, i.e.* $d(a, b) = d(f(a), f(b))$.

By considering this definition — or more easily by 'common sense' — it is evident that the symmetries of a square are as shown in Figure 1. Note that by convention the rotations are anticlockwise. Note also that a

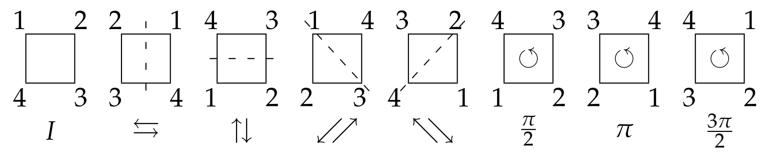

Figure 1: Symmetries of the square

rotation through $\dfrac{-\pi}{2}$ is the same as a rotation of $\dfrac{3\pi}{2}$.

(16.1.2) Representation Using The Complex Plane

Having found some symmetries, it is of interest as to how to represent them mathematically i.e. find the function as described in Definition 1. The easiest way to begin describing symmetries is using the complex plane, from which the equivalent relations in the real plane can be deduced.

Since $z = re^{i\theta}$, a rotation through an angle α must map z to $re^{i(\theta+\alpha)}$. Hence a rotation is represented by multiplying by $e^{i\alpha}$.

It is obvious that reflection in the real axis is represented by complex conjugation, so z maps to $\bar{z}$. From this point, other reflections can be found by combining with a rotation.

For a square, the rotations are through multiples of $\dfrac{\pi}{2}$, so since $e^{i\theta} = \cos\theta + i\sin\theta$ actual values can be calculated. In summary,

1. $z \mapsto z$ represents the identity transformation, I.
2. $z \mapsto iz$ represents a rotation through $\dfrac{\pi}{2}$.
3. $z \mapsto -z$ represents a rotation through π.
4. $z \mapsto -iz$ represents a rotation through $\dfrac{3\pi}{2}$.

5. $z \mapsto \bar{z}$ represents a reflection in the real axis.
6. $z \mapsto -\bar{z}$ represents a reflection in the imaginary axis — a reflection in the real axis followed by a rotation of π.
7. $z \mapsto -i\bar{z}$ represents a reflection in the line $Im(z) = Re(z)$ — a reflection in the real axis followed by a rotation of $\dfrac{\pi}{2}$.
8. $z \mapsto i\bar{z}$ represents a reflection in the line $Im(z) = -Re(z)$ — a reflection in the real axis followed by a rotation of $\dfrac{3\pi}{2}$.

Clearly, the basic rotation $r\colon z \mapsto iz$ and the basic reflection $s\colon z \to \bar{z}$ can be used to build up all the symmetries of the square. Notice that $rs \neq sr$ i.e. symmetries are not commutative.

(16.1.3) Representation In The Real Plane

From the results for the complex plane it is easy to covert the representation of the symmetries into the real plane $\mathbb{R}^2$. Using $z = x + iy = r(\cos\theta + i\sin\theta)$ observe that for a rotation,

$$\begin{aligned} r(\cos\theta + i\sin\theta) &\xmapsto{\times e^{i\alpha}} r\left(\cos(\theta+\alpha) + i\sin(\theta+\alpha)\right) \\ &= r(\cos\theta\cos\alpha - \sin\alpha\sin\theta) + ir(\sin\theta\cos\alpha + \cos\theta\sin\alpha) \\ &= (x\cos\alpha - y\sin\alpha) + i(x\sin\alpha + y\cos\alpha) \\ &= x' + y' \end{aligned}$$

This can be expressed in the obvious way as

$$\begin{pmatrix} x & y \end{pmatrix} \to \begin{pmatrix} x' & y' \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \tag{2}$$

Notice that the determinant of this matrix is 1, and since it represents any rotation, it follows that the determinant of a matrix is 1 if and only if the matrix represents a rotation.

The basic reflection can also be represented in matrix form in quite a trivial way,

$$\begin{pmatrix} x & y \end{pmatrix} \to \begin{pmatrix} x' & y' \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3}$$

Notice that the determinant of this matrix is -1, and by the matrix result

$$\det(AB) = \det(A)\det(B)$$

it follows that a matrix has determinant -1 if and only if it represents a reflection. The determinant of the identity matrix is 1, which is inkeeping with its interpretation as a rotation through an angle of 0.

It is convenient to write the matrix on the right, as then multiple operations are expressed simply by 'stacking up' the matrices. The interpretation of matrices explains why combining symmetries is a non-commutative operation. However, combining symmetries is associative, and this will be of use later.

(16.1.4) Representation Using Permutations

A third way to represent symmetries is using a permutation. A permutation is a matrix in which each column has entry in the top row of the vertex name, and in the bottom row of the new vertex name. Consider the symmetries described in Figure 2. The permutation matrices are calculated as follows.

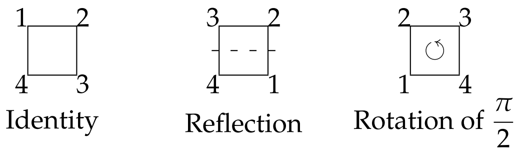

Figure 2: A reflection and a rotation

- For the identity, the permutation is $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$.
- For the reflection, the permutation is $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$.
- For the rotation, the permutation is $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$.

The symmetries of the square can be obtained by any combination of the rotation and reflection as already described. The symmetries may therefore be expressed as

$$\langle z \to iz \, , \, z \to \overline{z} \rangle \quad \text{or} \quad \left\langle \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle \tag{4}$$

For the rotation r there are 4 possibilities, and for the reflection s there are 2 possibilities. Therefore any rotation is of the form

$$s^i r^j \quad \text{for} \quad 0 \leqslant i \leqslant 1 \quad 0 \leqslant j \leqslant 3$$

Since $4 \times 2 = 8$ this gives a total of 8 symmetries, as has been verified empirically.

(16.1.5) Disjoint Cycles & Transpositions

A group may be written more concisely as a product of disjoint cycles. The top row of the permutation matrix is effectively redundant, so instead the elements are listed in the order of what they permute to. For example,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 4 & 1 & 2 & 3 & 5 & 9 & 8 & 6 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 3 & 2 \end{pmatrix} \begin{pmatrix} 5 \end{pmatrix} \begin{pmatrix} 6 & 9 & 7 & 8 \end{pmatrix}$$

and in fact the single 'cycle' containing the 5 may be omitted altogether on the understanding that anything not shown permutes to itself. This transposition is said to have cycle shape $4^2.1$, and this is extended generally in the obvious way. A cycle with just two elements is called a transposition. Note that disjoint cycles commute.

Taking powers of a permutation — i.e. repeating it many times — is easy to calculate, as to calculate the nth power simply take every nth element in each cycle until all are used and repeat for each cycle. This prompts the following.

Definition 5 *The order of a permutation is the power to which it has to be raised in order to produce the identity permutation.*

It is clear that for a single disjoint cycle of length n, its order will be n. However, for a product of disjoint cycles, the order will be the least common multiple of the cycle lengths.

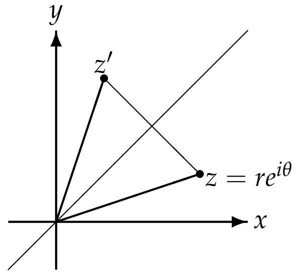

Figure 3: Reflection in any line through the origin.

Theorem 6 *Any permutation can be written as a product of transpositions.*

Proof. It is clear to see that

$$\begin{aligned}\begin{pmatrix}a_1 & a_2 & \dots & a_n\end{pmatrix} &= \begin{pmatrix}a_1 & a_2\end{pmatrix}\begin{pmatrix}a_1 & a_3 & \dots & a_n\end{pmatrix}\\ &= \begin{pmatrix}a_1 & a_2\end{pmatrix}\begin{pmatrix}a_1 & a_3\end{pmatrix}\begin{pmatrix}a_1 & a_4 & \dots & a_n\end{pmatrix}\\ &\vdots\\ &= \begin{pmatrix}a_1 & a_2\end{pmatrix}\begin{pmatrix}a_1 & a_3\end{pmatrix}\dots\begin{pmatrix}a_1 & a_n\end{pmatrix}\end{aligned}$$

as required. □

Since $(1\ 2)(1\ 2) = 1$ it is easy to see that the number of transpositions that form a permutation is not a well-defined quantity. However, whether an odd or even number of transpositions is needed to express a permutation is well-defined. The following definition is therefore made.

Definition 7 *An even permutation is a permutation which, when written as a product of transpositions, has an even number of transpositions in this product. If a permutation is not even, then it is odd.*

A simple example of an even transposition, $(1\ 2)(1\ 2) = 1$, has already been exhibited. However, it is quite possible that an odd permutation could be written as an even permutation, and it is not obvious that odd permutations even exist. Taking this into account makes the definition a little 'one sided'. Of course odd permutations do exist, but this needs to be proved, as is done on page 17.

(16.2) Groups

(16.2.1) Important Groups

Reflections: The Orthogonal Group

Consider a reflection in a line through the origin making an angle of $\dfrac{\phi}{2}$ with the positive x axis, as shown in Figure 3.

It is clear that a reflection in the line making angle $\dfrac{\phi}{2}$ is the same as a rotation through angle $2\left(\dfrac{\phi}{2} - \theta\right) =$

$\phi - 2\theta$. Therefore

$$\begin{aligned} z' &= ze^{i(\phi-2\theta)} \\ &= re^{i\theta}e^{i(\phi-2\theta)} \\ &= re^{-i\theta}e^{i\phi} \\ &= \overline{z}e^{i\phi} \end{aligned}$$

So such a reflection is achieved by means of a reflection in the x axis followed by a rotation through angle ϕ. In matrix form this gives

$$\mathrm{ref}_\phi = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix}$$

notice that $\det \mathrm{ref}_\phi = -1$.

These reflections and the rotations can be combined in any way, giving

$$\begin{aligned} \mathrm{rot}_\alpha\,\mathrm{rot}_\beta &: z \xmapsto{\mathrm{rot}_\alpha} ze^{i\alpha} \xmapsto{\mathrm{rot}_\beta} ze^{i\alpha}e^{i\beta} &&= ze^{i(\alpha+\beta} &&= \mathrm{rot}_{\alpha+\beta} \\ \mathrm{rot}_\alpha\,\mathrm{ref}_\beta &: z \xmapsto{\mathrm{rot}_\alpha} ze^{i\alpha} \xmapsto{\mathrm{ref}_\beta} \overline{ze^{i\alpha}}e^{i\beta} &&= \overline{z}e^{i(\beta-\alpha)} &&= \mathrm{ref}_{\beta-\alpha} \\ \mathrm{ref}_\alpha\,\mathrm{rot}_\beta &: z \xmapsto{\mathrm{ref}_\alpha} \overline{z}e^{i\alpha} \xmapsto{\mathrm{rot}_\beta} \overline{z}e^{i\alpha}e^{i\beta} &&= \overline{z}e^{i(\alpha+\beta)} &&= \mathrm{ref}_{\alpha+\beta} \\ \mathrm{ref}_\alpha\,\mathrm{ref}_\beta &: z \xmapsto{\mathrm{ref}_\alpha} \overline{z}e^{i\alpha} \xmapsto{\mathrm{ref}_\beta} \overline{\overline{z}e^{i\alpha}}e^{i\beta} &&= ze^{i(\beta-\alpha)} &&= \mathrm{rot}_{\beta-\alpha} \end{aligned}$$

It is evident, therefore, that any combination of rotations and reflections is a rotation or a reflection. This readily suggests the use of a group. Let

$$O_2 = \left\{ \mathrm{rot}_\alpha,\ \mathrm{ref}_\beta \mid 0 \leqslant \alpha < 2\pi, \quad 0 \leqslant \beta < \pi \right\}$$

then since the rotations and reflections can be expressed as matrices, it follows that O_2 is a group. It is called the orthogonal group in 2 dimensions, and represents the symmetries of a circle with centre at the origin.

Rotations: The Cyclic Group

The rotations of an n-gon form a group by themselves. It may be written as

$$C_n = \langle r \mid r^n = 1 \rangle$$

This is an example of a cyclic group which has one generating element which, when raised to some power, is the identity.

The group C_∞ contains all powers of its generating element, $e, r, r^2, \ldots$. However, to be a group all the inverses, $e, r^{-1}, r^{-2}, \ldots$ must also be in C_∞. Hence all integer powers of r are in C_∞, and so C_n "looks like" $\mathbb{Z}$. The two are said to be 'isomorphic', a term that is defined later.

Definition 8 *The order of an element g of a group is the least integer n such that $g^n = e$.*

Definition 9 *The order or a group $(G, \times)$ is the cardinality of the set G.*

It is clear that the order of the generating element of C_n is n.

Permutations: The Symmetric Group

Permutations can form a group. The symmetric group on n items, S_n is the set of all permutations of n items. Clearly it has order $n!$.

The binary operation is composition of permutations. Notice that every element of the group of symmetries on an n-gon is in S_n. However, there are some permutations in S_n that are not symmetries of a square, say. The square has 8 symmetries but S_4 has $4! = 24$ elements.

Symmetries: The Dihedral Group

The Dihedral group D_n is the group of symmetries of a regular n-gon, of which there are n reflections and n rotations. The order of D_n is therefore $2n$. The two basic symmetries are a reflection which is its own inverse and hence has order 2; and a rotation of order n.

Being a finite group, it is possible to construct a multiplication table for D_n. For large n this is rather an unwieldy process; Table 16.2.1 gives the multiplication table for D_3.

Calculating such a table is rather difficult since the elements of the group do not commute. Recall that D_3 may be defined as

$$D_3 = \left\langle r,\ s \mid r^3 = s^2 = (rs)^2 = 1 \right\rangle$$

using the identity relationship,

$$\begin{aligned}
(rs)^2 &= rsrs = 1 \\
r^{-1}(rsrs)s^{-1} &= r^{-1}s^{-1} \\
sr &= r^{-1}s^{-1} \quad \text{by commutativity} \\
&= r^2 s
\end{aligned}$$

Hence a relationship has been found to express sr the usual way round. Using the same method, similar relationships can be found for other dihedral groups. The multiplication table can now be formed.

	1	r	r^2	s	rs	r^2s
1	1	r	r^2	s	rs	r^2s
r	r	r^2	1	rs	r^2s	s
r^2	r^2	1	r	r^2s	s	rs
s	s	r^2s	rs	1	r^2	r
rs	rs	s	r^2s	r	1	r^2
r^2s	r^2s	rs	s	r^2	r	1

Table 1: Multiplication table of D_3

Matrices: The General Linear Group

The orthogonal group is in fact a special case of a group of all linear transformations, however, its ‘parent’ group is not thought of in terms of symmetry: the general linear group is more of an abstract concept.

Definition 10 *The general linear group over the real numbers is the set*

$$GL_n(\mathbb{R}) = \{A \in M_{nn}(\mathbb{R}) \mid \det A \neq 0\}$$

with matrix multiplication

Defining this set is all very well, but this does not mean that it is a group.

Theorem 11 *The general linear group defined as the set*

$$GL_n(\mathbb{R}) = \{A \in M_{nn}(\mathbb{R}) \mid \det A \neq 0\}$$

together with the binary operation of matrix multiplication, is a group.

Proof. All 4 of the axioms need to be verified (see Definition 15).

closure: $A,\ B \in GL_n(\mathbb{R}) \Rightarrow \det A,\ \det B \neq 0$.
Thus $\det(AB) = (\det A)(\det B) \neq 0$ and so $AB \in GL_n(\mathbb{R})$.

associativity: The associativity of matrix multiplication is a known result*.

identity: I_n, the identity matrix satisfies $AI_n = I_nA = A\ \forall A \in GL_n(\mathbb{R})$. Hence the required element exists.

inverse: Let $A \in GL_n(\mathbb{R})$ then $\det A \neq 0$. Suppose that $\exists A^{-1} \in GL_n(\mathbb{R})$ then since $AA^{-1} = I_n$, $\det(AA^{-1}) = 1$ and so $\det A^{-1} \neq 0$ hence $A^{-1} \in GL_n(\mathbb{R})$.

All four properties hold, so the proof is complete. □

It is clear that $O_2 \subseteq GL_2(\mathbb{R})$, but since O_2 is itself a group under the same binary operation, it is called a subgroup of $GL_2(\mathbb{R})$. This is commonly written as $O_2 \leqslant GL_2(\mathbb{R})$.

(16.2.2) Self-Adjoint Linear Transformations

In light of the fact that $O_2 \leqslant GL_2(\mathbb{R})$, what is so special about the matrices in O_2 that makes it a subgroup? This is partly answered by the following theorem.

Theorem 12 *The orthogonal group O_2 can be defined by*

$$O_2 = \left\{A \in GL_2(\mathbb{R}) \mid AA^T = I_2\right\}$$

i.e. $A \in O_2 \Leftrightarrow AA^T = I_2$.

Proof. Taking each direction in turn,

$\Rightarrow$ If $A \in O_2$ then it is either a rotation or a reflection.

$$\text{rot}_\alpha\,\text{rot}_\alpha^T = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos^2\alpha + \sin^2\alpha & 0 \\ 0 & \cos^2\alpha + \sin^2\alpha \end{pmatrix} = I_2$$

and similarly

$$\text{ref}_\alpha\,\text{ref}_\alpha^T = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix} = \begin{pmatrix} \cos^2\alpha + \sin^2\alpha & 0 \\ 0 & \cos^2\alpha + \sin^2\alpha \end{pmatrix} = I_2$$

Clearly the result holds in both cases.

$\Leftarrow$ Now, $\det A = \det A^T$ so $\det(AA^T) = (\det A)^2 = 1$ since $AA^T = I_2$. Therefore $\det A = \pm 1$.
Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then $A^{-1} = \dfrac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. But $AA^T = I_2$ so $\dfrac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$.
Hence,

*The associativity of matrix multiplication can be shown by considering the summation for the ij entry of $A(BC)$ and showing that it is the same as that for $(AB)C$.

i. If $\det A = 1$ then $a = d$ and $b = -c$ so $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, and so $a^2 + b^2 = 1$.

ii. If $\det A = -1$ then $a = -d$ and $b = c$ so $A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}$, and so $-a^2 - b^2 = -1$.

In both cases the relationship $a^2 + b^2 = 1$ holds and so $\exists \alpha$ with $a = \cos \alpha$ and $b = \sin \alpha$ giving the required form. □

Definition 13 *Let $A \in M_{22}$ represent a linear transformation. The transformation represented by A is said to be self adjoint if $AA^T = I_2$.*

The term "self adjoint" comes from the method for inverting a 3×3 matrix, where the matrix of cofactors is called the adjoint. In this special case, the transpose of the matrix is the adjoint. Since by definition a symmetry must preserve distance, self adjoint linear transformations are of special interest, as is now proved.

Theorem 14 *Self adjoint linear transformations preserve inner products in that*

$$\langle \mathbf{u} \mid \mathbf{v} \rangle = \langle \mathbf{u}A \mid \mathbf{v}A \rangle \Leftrightarrow AA^T = I_2$$

This can of course be extended into more dimensions. Note that when dealing with symmetries in this way it is more convenient to think of $\mathbf{u}$ and $\mathbf{v}$ as row vectors than column vectors, hence the matrix A is postmultiplied in order to perform its transformation.

Proof. Taking each case in turn,

$\Rightarrow$ Let $AA^T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

case 1. Choose $\mathbf{u} = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 1 & 0 \end{pmatrix}$.

$$\mathbf{u}AA^T\mathbf{v}^T = a \quad \text{and} \quad \mathbf{u}\mathbf{v}^T = 1$$

hence $a = 1$.

case 2. Choose $\mathbf{u} = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 0 & 1 \end{pmatrix}$.

$$\mathbf{u}AA^T\mathbf{v}^T = b \quad \text{and} \quad \mathbf{u}\mathbf{v}^T = 0$$

hence $b = 0$.

case 3. Choose $\mathbf{u} = \begin{pmatrix} 0 & 1 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 1 & 0 \end{pmatrix}$.

$$\mathbf{u}AA^T\mathbf{v}^T = c \quad \text{and} \quad \mathbf{u}\mathbf{v}^T = 0$$

hence $c = 0$.

case 4. Choose $\mathbf{u} = \begin{pmatrix} 0 & 1 \end{pmatrix}$ and $\mathbf{v} = \begin{pmatrix} 0 & 1 \end{pmatrix}$.

$$\mathbf{u}AA^T\mathbf{v}^T = d \quad \text{and} \quad \mathbf{u}\mathbf{v}^T = 1$$

hence $d = 1$.

Hence it is evident that $AA^T = I_2$ so the result holds in this direction.

$\Leftarrow$ Consider the dot product $\mathbf{u} \cdot \mathbf{v}$, then,

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= \mathbf{u}\mathbf{v}^T \\ &\xmapsto{A} (\mathbf{u}A)(\mathbf{v}A)^T \\ &= \mathbf{u}AA^T\mathbf{v} \quad \text{but by hypothesis, } AA^T = I \\ &= \mathbf{u}\mathbf{v}^T\end{aligned}$$

Clearly the result holds.

Both implications hold, so the theorem is proven. □

Linear Transformations As Symmetries

It has been shown above that self adjoint linear transformations preserve inner products. Now,

$$|\mathbf{u}| = \sqrt{\langle \mathbf{u} \mid \mathbf{u} \rangle}$$

and

$$\langle \mathbf{u} \mid \mathbf{v} \rangle = |\mathbf{u}||\mathbf{v}| \cos\theta$$

so length and cosine of angle are conserved under these linear transformation. Notice that it is the cosine of the angle that is preserved and not the angle, as under a reflection it is possible that $\theta \mapsto -\theta$ but since cosine is an even function, the cosine is still preserved.

(16.3) Abstract Group Theory

Having observed a number of groups which have meaningful practical interpretations, from this point concern will shift to the theoretical properties of a group.

(16.3.1) Axioms Of A Group

Mathematics is often concerned with the study of abstract structures such as vector spaces. Groups are another such structure, and as usual a structure is defined in terms of its axioms — unconditional truths.

Definition 15 *A group G is a set, equipped with a binary operation, such that the following 4 axioms hold.*

1. *The set is closed under the binary operation.*
2. *The binary operation is associative.*
3. *There exists an identity element e such that for any $g \in G$, $eg = ge = g$.*
4. *Every element of G has an inverse in G, so that if $g \in G$ then $\exists g' \in G$ such that $gg' = g'g = e$.*

Having defined the axioms, it is of interest as to what can be deduced from them. This, essentially, is the subject studied from here onwards.

Theorem 16 *For a group G with elements g, h, and k,*

1. *The identity element of G is unique.*
2. *The inverse of any element of G is unique.*

3. *$(h^{-1})^{-1} = h$.*
4. *$(gh)^{-1} = h^{-1}g^{-1}$.*
5. *The cancellation laws hold in that if $gh = gk$ or $hg = kg$ then $h = k$.*

Proof. Taking each part in turn,

1. Suppose that e and e' are two identities of G.
   Since e is an identity, $ee' = e'$.
   Since e' is an identity, $ee' = e$.
   Hence $e = ee' = e'$ and the theorem holds.
2. Suppose that g' and g'' are inverses of g.
   Then $g'gg'' = (g'g)g'' = eg'' = g''$ by the associative property, the inverse property, and the identity axioms respectively.
   And $g'gg'' = g'(gg'') = g'e = g'$ by the associative property, the inverse property, and the identity axioms respectively.
   Hence $g' = g'gg'' = g''$ and the theorem holds.
3. Now, $gg^{-1} = g^{-1}g = e$ so g is an inverse for g^{-1}. By the preceding result, inverses are unique, so $(g^{-1})^{-1} = g$ and so the theorem holds.
4. $$\begin{aligned}
(gf)(f^{-1}g^{-1}) &= \left((gf)f^{-1}\right)g^{-1} && \text{by the associative axiom.}\\
&= \left(g(ff^{-1})\right)g^{-1} && \text{by the associative axiom.}\\
&= (ge)g^{-1} && \text{by the inverses axiom.}\\
&= gg^{-1} && \text{by the identity axiom.}\\
&= e && \text{by the iverses axiom.}
\end{aligned}$$

   Hence $f^{-1}g^{-1}$ is a right inverse for gf. It is readily shown that it is also a left inverse by a similar process. Hence the theorem holds.
5. $$\begin{aligned}
\text{suppose}\quad gh &= gk && \text{by the inverses axiom, } g^{-1} \text{ exists, so premultiply by it}\\
g^{-1}(gh) &= g^{-1}(gk) &&\\
(g^{-1}g)h &= (g^{-1}g)k && \text{by the associative axiom.}\\
eh &= ek && \text{by the inverses axiom.}\\
h &= k && \text{by the identity axiom.}
\end{aligned}$$

   Hence the result holds for left cancellation. A similar process shows that the result also holds for right cancellation, and hence the theorem holds. □

(16.3.2) Combining Groups

Composition

As functions can be composed, so can groups.

Theorem 17 *Let $(G,\circ)$ and $(H,\cdot)$ be groups. Then the cartesian product*[†] *$G \times H$ is a group under the operation*

$$(g_1,h_1) \odot (g_2,h_2) = (g_1 \circ g_2, h_1 \cdot h_2)$$

Proof. To show that something is a group, the axioms need to be checked.

closure: Certainly $G \times H$ is closed under $\odot$ since G is closed under $\circ$ and H is closed under $\cdot$.

associativity:

$$\begin{aligned}((g_1,h_1) \odot (g_2,h_2)) \odot (g_3,h_3) &= (g_1 \circ g_2, h_1 \circ h_2) \odot (g_3,h_3) \\ &= ((g_1 \circ g_2) \circ g_3, (h_1 \cdot h_2) \cdot h_3) \\ &= (g_1 \circ (g_2 \circ g_3), h_1 \cdot (h_2 \cdot h_3)) \\ &= (g_1,h_1) \odot (g_2 \circ g_3, h_2 \cdot h_3) \\ &= (g_1,h_1) \odot ((g_2,h_2) \odot (g_3,h_3))\end{aligned}$$

Hence $\odot$ is associative.

identity: Consider (e_G, e_H).
Clearly $(g,h) \odot (e_G,e_H) = (g \circ e_G, h \cdot e_H) = (g,h)$
Similarly $(e_G,e_H) \odot (g,h) = (g,h)$
Hence the identity element is (e_G,e_H).

inverse: Consider (g^{-1},h^{-1}).
Clearly $(g,h) \odot (g^{-1},h^{-1}) = (g \circ g^{-1}, h \cdot h^{-1}) = (e_G,e_H)$
Similarly $(g^{-1},h^{-1}) \odot (g,h) = (e_G,e_H)$
Hence the inverse of (g,h) is (g^{-1},h^{-1}). □

Hence $G \times H$ is a group, and is called the direct product of G and H and clearly $|G \times H| = |G||H|$.

Consider the groups $\hat{G} = (g, e_H)$ and $\hat{H} = (e_G, h)$. Notice that $(g_1,e_H) \odot (g_2,e_H) = (g_1 \circ g_2, e_H)$. Furthermore, $|\hat{G}| = |G|$ so $\hat{G}$ is effectively the same group as G.

Any element of $G \times H$ can be written as $\hat{g} \odot \hat{h}$ for $\hat{g} \in \hat{G}$ and $\hat{h} \in \hat{H}$. Note also that

$$\hat{g} \odot \hat{h} = (g,e_H) \odot (e_G,h) = (g,h) = (e_G,h) \odot (g,e_H) = \hat{h} \odot \hat{g}$$

so every element of $\hat{G}$ commutes with every element of $\hat{H}$.

(16.3.3) Subgroups

Definition 18 *A non-empty subset H of a group G which is itself a group under the same binary operation as G is a subgroup of G, $H \leqslant G$.*

Note that since H is a group, $e_H e_H = e_H$ but since $H \subseteq G$, $e_H = e_H e_G$. Hence by cancellation, the identity if H is the same as the identity of G.

Furthermore, if h' is the inverse of h in H, then $hh' = e_H = e_G$ and hence it is also the inverse of h in G. Inverses are therefore the same in H as they are in G.

Theorem 19 *Let H be a non-empty subset of a group G. Then H is a subgroup of G if and only if*

[†] The Cartesian product of the sets G and H is the set $G \times H = \{(g,h) \mid g \in G \quad h \in H\}$.

(i) $hk \in H \quad \forall h, k \in H$.

(ii) $h^{-1} \in H \quad \forall h \in H$.

Proof. The '$\Rightarrow$' proof is obvious.

For the '$\Leftarrow$' proof, consider each axiom in turn.

closure: By hypothesis H is closed, so no proof is required.

associativity: Since the binary operation of the group G is associative and $H \subseteq G$ and H takes the same binary operation as G, it follows that the binary operation of H is associative.

identity: Consider some $h \in H$ which can be done since $H \neq \emptyset$. Then by hypothesis $h^{-1} \in H$. But also by hypothesis, the product $hh^{-1} \in H$ i.e. $e_H = e_G \in H$. Hence H contains an identity element.

inverse: By hypothesis every element of H has its inverse in H, so no proof is needed. □

However, it is possible to find a more efficient sufficient condition for a subset to be a subgroup.

Theorem 20 *Let H be a non-empty subset of a group G. Then H is a subgroup of G if and only if $hk^{-1} \in H$ $\forall h, k \in H$.*

Proof. Again, the '$\Rightarrow$' proof is obvious.

For the '$\Leftarrow$' proof, consider each axiom in turn.

closure: By hypothesis H is closed, so no proof is required.

associativity: Since the binary operation of the group G is associative and $H \subseteq G$ and H takes the same binary operation as G, it follows that the binary operation of H is associative.

identity: Consider $h \in H$, and in the hypothesis put '$k = h$', then by hypothesis $hh^{-1} = e_H = e_G \in H$.

inverse: Putting '$h = e_H$' and '$k = h$', then by hypothesis $h^{-1} \in H$. Hence h has in inverse in H. □

Theorem 21 *If $H \leqslant G$ and $K \leqslant G$ then $H \cap K \leqslant G$.*

Proof. Consider $x, y \in H \cap K$.
Since $x, y \in H$ and H is a group, $xy^{-1} \in H$.
Since $x, y \in K$ and K is a group, $xy^{-1} \in K$.
Hence $xy^{-1} \in H \cap K$ and so by Theorem 20 $H \cap K \leqslant G$ as required. □

Since a subgroup must be closed under the binary operation, it is clear that if x is an element of some subgroup, then all powers of it must also be in the subgroup; not only $x^2, x^3, \ldots$ but also $x^{-1}, x^{-2}, \ldots$. However, if x has finite order as described by Definition 8 then this list of other elements that must be in the subgroup is also finite.
If G is a group and $x \in G$ then the subgroup generated by x is $\langle x \rangle = \{x^r \mid 1 \leqslant r < m\}$ where m is the order of x. But is this really a subgroup?

closure: Follows by definition, since any power of x can be expressed as x^r such that $1 \leqslant r < m$.

associativity: Since x is already a member of some other group, and $\langle x \rangle$ takes the same binary operation, $\langle x \rangle$ must certainly be associative.

identity: $x^0 = e \in \langle x \rangle$, inverse: Consider x^r for $1 \leqslant r < m$ then clearly also $x^{m-r} \in \langle x \rangle$. But $x^r x^{m-r} = x^{m-r} x^r = x^m = e$ hence any element of $\langle x \rangle$ has its inverse in $\langle x \rangle$.

Hence $\langle x \rangle$ is called the subgroup generated by x.

If x does not have finite order, then all elements of $\langle x \rangle$ are distinct and there is a one-to-one correspondence with the elements of $\mathbb{Z}$. Furthermore, multiplying any two elements of $\langle x \rangle$ is done by adding powers, so $\langle x \rangle$ behaves like the group $(\mathbb{Z}, +)$. The two are said to be isomorphic, a term which is properly defined in Definition 24. The symbol '$\cong$' is used to express isomorphism. Notice that when x does have finite order m, $\langle x \rangle \cong C_m$.

Definition 22 *For a subset S of a group G, define $\langle S \rangle$ — the subgroup generated by S — to be the 'smallest' subgroup of G that contains S i.e.*

$$\bigcap_{\substack{S \subseteq H \\ H \leqslant G}} H$$

which is the intersection of all subgroups of which S is a subset.

Cyclic Subgroups

It has been seen that generating a subgroup using an element of a group produces a cyclic group, regardless of the type of group the element came from. It seems reasonable, therefore, that any subgroup of a cyclic group should be cyclic.

Theorem 23 *A subgroup of a cyclic group is cyclic*

Proof. Let G be a cyclic group, $x \in G$, and $H \leqslant G$; then H is a cyclic subgroup of G.
Consider $P = \{j \in \mathbb{Z} \mid x^j \in H\}$. Note that $j \in P \Leftrightarrow -j \in P$. Let k be the smallest positive member of P.
If $k = 1$, then $G = H$.
Suppose $x^l \in H$ with $l > k$. Then $l = ak + b$, say, for $b < k$. Hence

$$x^l = x^{ak+b} = \left(x^k\right)^a x^b \tag{$*$}$$

Now, $x^{-k} \in H$ therefore $\left(x^{-k}\right)^a \in H$ and also $x^l \in H$. Since H is a subgroup, $x^b = x^l \left(x^{-k}\right)^a \in H$. But by the minimality of k this is a contradiction.
Hence $b = 0$ so equation $(*)$ gives $x^l = \left(x^k\right)^a$ and so it follows that $H = \left\langle x^k \right\rangle$ which is cyclic. $\square$

Homomorphisms & Isomorphisms

It was seen above that if $x \in G$ has infinite order, then G is "sort of the same as" $\mathbb{Z}$. The concept of groups behaving in similar ways like this is formalised as follows.

Definition 24 *Let $(G, *)$ and $(H, \circ)$ be groups, and let ϕ be a map $\phi\colon G \to H$.*

1. *ϕ is a homomorphism $\Leftrightarrow (g_1 * g_2)^\phi = g_1^\phi \circ g_2^\phi$ for all $g_1, g_2 \in G$.*
2. *ϕ is an isomorphism if it is also a bijection. This is written as $G \cong H$.*

Theorem 25 *If ϕ is a homomorphism then*

1. *$e_G^\phi = e_H$.*
2. *$\left(g^{-1}\right)^\phi = \left(g^\phi\right)^{-1}$.*

3. $G^\phi = \{g_\phi \mid g \in G\} \leqslant H$.
4. *G abelian $\Rightarrow G^\phi$ abelian.*
5. $\ker\phi = \{g \in G \mid g^\phi = e_H\} \leqslant G$.

Proof. 1. Working from the definitions,

$$\begin{aligned} e_G * e_G &= e_G \\ (e_G * e_G)^\phi &= e_G^\phi \\ e_g^\phi \circ e_g^\phi &= e_G^\phi \\ &= e_G^\phi \circ e_H \\ e_G^\phi &= e_H \quad \text{by cancellation} \end{aligned}$$

2. Now, $\left(g * g^{-1}\right)^\phi = e_G^\phi = e_H$
   Also, $\left(g * g^{-1}\right)^\phi = g^\phi (g^{-1})^\phi$
   Hence $e_H = g^\phi (g^{-1})^\phi$ and so $(g^{-1})^\phi$ is a right inverse for g^ϕ.
   Similarly, considering $\left(g^{-1} * g\right)^\phi$ shows that $(g^{-1})^\phi$ is a left inverse for g^ϕ.
   Hence $(g^{-1})^\phi$ is an inverse for g^ϕ, i.e. $\left(g^{-1}\right)^\phi = \left(g^\phi\right)^{-1}$.

3. Let $g_1^\phi, g_2^\phi \in G^\phi$.

closure: $g_1^\phi \circ g_2^\phi = (g_1 * g_2)^\phi$ but $g_1 * g_2 \in G$ and hence $(g_1 * g_2)^\phi \in G^\phi$.

associative: $G^\phi \subseteq H$ and H is a group, hence G^ϕ is associative.

identity: By part 1 of the theorem $e_G^\phi = e_H$ and clearly $e_G^\phi \in G^\phi$. Hence G^ϕ has an identity.

inverse: By part 2 of the theorem, $\left(g^{-1}\right)^\phi = \left(g^\phi\right)^{-1}$ and certainly $\left(g^{-1}\right)^\phi \in G^\phi$, hence G^ϕ has inverses.

Hence by verifying the group axioms, G^ϕ is a group. Since $G^\phi \subseteq H$ it follows that $G^\phi \leqslant H$.

4. If G is abelian then,
$$g_1^\phi \circ g_2^\phi = (g_1 * g_2)^\phi = (g_2 * g_1)^\phi = g_2^\phi \circ g_1^\phi$$
   Hence if G is abelian, then so is G^ϕ.

5. Let $k, l \in \ker\phi$ so $k^\phi = l^\phi = e_H$.
   (i) $(k * l)^\phi = k^\phi \circ l^\phi = e_H \circ e_H = e_H$ so $\ker\phi$ is closed under multiplication.
   (ii) $\left(k^{-1}\right)^\phi = \left(k^\phi\right)^{-1}$ by part 2 of the theorem. But $\left(k^\phi\right)^{-1} = e_G^\phi = e_H$ so $k^{-1} \in \ker\phi$.
   Hence by the test for a subgroup $\ker\phi \leqslant G$. □

Since isomorphisms are bijections, they have inverses which are themselves bijections i.e. isomorphisms. Recall from Chapter **??** that

- A function $f\colon X \to Y$ is injective if $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$. In the conventional sense, any horizontal line will intersect the function at most once.
- A function $f\colon X \to Y$ is surjective if $\forall y \in Y\ \exists x \in X$ such that $y = f(x)$. Any horizontal line will therefore intersect the function at least once.
- f is bijective if it is both injective and surjective, so any horizontal line will intersect the function exactly once.

Theorem 26 *If $\phi\colon G \to H$ is an isomorphism, then $\phi^{-1}\colon H \to G$ is also an isomorphism.*

Proof. Consider $h, l \in H$ such that $g^\phi = h$ and $k^\phi = l$ where $g, k \in G$,

$$(hl)^{\phi^{-1}} = \left(g^\phi k^\phi\right)^{\phi^{-1}} = \left((gk)^\phi\right)^{\phi^{-1}} = gk = h^{\phi^{-1}} l^{\phi^{-1}}$$

so ϕ^{-1} obeys the homomorphism property. Since it is the inverse of a bijection it is also a bijection, hence ϕ^{-1} is an isomorphism. $\square$

Theorem 27 *A homomorphism ϕ is injective $\Leftrightarrow$ $\ker \phi = \{e_G\}$.*

Proof. $\Rightarrow$ Suppose ϕ is injective. Let $g \in \ker \phi$, but always $e_G \in \ker \phi$. Hence by injectivity $g = e_G$.

$\Leftarrow$ Suppose $\ker \phi = \{e_G\}$ and suppose $g^\phi = h^\phi$. Then

$$\begin{aligned} \left(gh^{-1}\right)^\phi &= g^\phi \left(h^{-1}\right)^\phi \\ &= g^\phi \left(h^\phi\right)^{-1} \\ &= h^\phi \left(h^\phi\right)^{-1} \\ &= e_H \end{aligned}$$

Hence $gh^{-1} \in \ker \phi$ and so $gh^{-1} = e_G$ therefore $g = h$. Hence ϕ is injective. $\square$

Cayley's Theorem

At the beginning of the chapter it was seen that the symmetries of a figure could be represented by a dihedral group, or a subgroup of a symmetric group. This suggests that dihedral groups, and indeed many other kind of groups, are isomorphic to subgroups of a symmetric group. In fact this is true for all groups, as is now proved.

Theorem 28 (Cayley's Theorem) *Let G be a group with set of elements X. Then G embeds (is a subgroup of) S_X, the symmetric group on X.*
If $|X| = n$ then G embeds in S_n.

Proof. For some particular $g \in X$ define the permutation of X

$$\pi_g : x \mapsto xg \qquad \forall x \in X$$

A permutation must be bijective, and this is now verified.

(i) For injectivity,

$$x^{\pi_g} = y^{\pi_g} \;\Leftrightarrow\; xg = yg \;\Leftrightarrow\; x = y$$

The last step following by cancellation. Hence π_g is injective.

(ii) For surjectivity it must be shown that $\forall x \in X \exists y \in X$ such that $y^{\pi_g} = x$.
Put $y = xg^{-1}$ which is in X since $x, g \in X$ and X is the set of elements of a group. Now,

$$y^{\pi_g} = yg = xg^{-1}g = x$$

Hence π_g is surjective.

Hence π_g is bijective and so is a permutation, as claimed.

Let $H = \{\pi_g \mid g \in X\}$. Claim that H is a group, which is shown from the axioms closure: For $g, h \in G$,

$$\begin{aligned} x^{\pi_g \pi_h} &= (xg)^{\pi_h} \\ &= (xg)h = x(gh) \quad \text{since } g, h \in X \text{ which is a group} \\ &= x^{\pi_{gh}} \end{aligned}$$

Now, $gh \in X$ so $\pi_{gh} \in H$. Hence H is closed.

associative: In a similar way to above,

$$\begin{aligned} x^{\left(\pi_g \pi_h\right)\pi_k} &= (x(gh))^{\pi_k} \\ &= x(gh)k \\ &= xg(hk) \\ &= (xg)^{\pi_h \pi_k} \\ &= x^{\pi_g(\pi_h \pi_k)} \end{aligned}$$

Note that strictly speaking a number of intermediate steps have been omitted here.

identity: Claim that π_{e_G} is the identity of H.

$$\begin{aligned} x^{\pi_g \pi_e} &= (xg)^{\pi_e} \\ &= xge = xg \\ &= x^{\pi_g} \end{aligned}$$

This shows that π_e is a right identity, and clearly it is also a left identity. Hence the result.

inverse: Since $g \in X$, $g^{-1} \in X$ so the permutation $\pi_{g^{-1}}$ exists in H. Now,

$$x^{\pi_g \pi_{g^{-1}}} = (xg)^{\pi_{g^{-1}}} = (xg)g^{-1} = x = x^{\pi_e}$$

Hence the result

All axioms have been verified, so H is indeed a group. In fact H is a group of $|X|$ permutations, each of which permute $|X|$ items. Hence $H \leqslant S_X$.

Now define the function $\phi \colon G \to H$ by $g^\phi = \pi_g$. Since H is a group,

$$(gh)^\phi = \pi_{gh} = \pi_g \pi_h = g^\phi h^\phi$$

so ϕ is a homomorphism. To complete the proof it must now be shown that ϕ is an isomorphism.

(i) Consider $\ker \phi = \left\{ g \in X \mid \pi_g = \pi_e \right\}$.

$$g \neq e \Rightarrow \pi_g \colon x \to xg \neq x \therefore \pi_g \neq \pi_e$$

Hence $\ker \phi = \{e\}$ so by Theorem 27, ϕ is injective.

(ii) Trivially ϕ is surjective since $\pi_g \in H \Rightarrow \exists g \in G$ such that g defines π_g.

Since ϕ is bijective it is an isomorphism, so $G \cong H$. Also, $H \leqslant S_X$, hence the result. □

Cayley's Theorem itself is of negligible practical use. The methods employed in the proof, however, are. In summary,

- For each element of the group, define a permutation of the group for it. Show that this is a permutation by showing that it is bijective.
- Show that the set of these permutations is a group, and is therefore a subgroup of a symmetric group.
- Show that the group is isomorphic to the group of permutations which its elements define.

How Cayley's Theorem is used is best illustrated by means of an example.

Example 29 *Find a subgroup of S_4 that is isomorphic to C_4 and contains the permutation $\begin{pmatrix}1 & 4 & 3 & 2\end{pmatrix}$.*

Proof. Solution Observe that one possible way to label the elements is

$$\begin{pmatrix} e & x & x^2 & x^3 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

So $\begin{pmatrix}1 & 4 & 3 & 2\end{pmatrix}$ represents multiplication on the right by x^3 i.e. π_{x^3}.
The permutations must therefore be

$$\begin{aligned} \pi_e &= (1)(2)(3)(4) \\ \pi_x &= \begin{pmatrix}1 & 2 & 3 & 4\end{pmatrix} \\ \pi_{x^2} &= \begin{pmatrix}1 & 3\end{pmatrix}\begin{pmatrix}2 & 4\end{pmatrix} \\ \pi_{x^3} &= \begin{pmatrix}1 & 4 & 3 & 2\end{pmatrix} \end{aligned}$$

□

(16.3.4) Odd And Even Permutations

Existence Of Odd Permutations

It seems odd to have developed so much theory without yet having seen whether odd permutations exist: Recall that Definition 7 says nothing about the existence of odd permutations, if indeed they do exist. Enough theory has been covered at this point to show that in fact they do.

Consider the product $\prod\limits_{i<j}(x_i - x_j)$, a special case of which is the determinant of the VanDermonde matrix,

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = \pm \prod_{i<j}(x_i - x_j)$$

For the sake of clarity, observe that

$$\begin{aligned} \prod_{i<j}(x_i - x_j) = (x_1 - x_2)(x_1 - x_3)(x_1 - x_4)&\dots(x_1 - x_n) \\ (x_2 - x_3)(x_2 - x_4)&\dots(x_2 - x_n) \\ (x_3 - x_4)&\dots(x_3 - x_n) \\ &\ddots\vdots \\ &(x_{n-1} - x_n) \end{aligned}$$

Consider the effect of a single adjacent transposition, $\tau_1 = \begin{pmatrix} i & i+1 \end{pmatrix}$. This will swap factors in the same columns, and the 'extra' factor on line i will be negated. Hence the value of the whole product is negated.

Consider now the effect of two adjacent transpositions. The value of the product will be negated twice and so stay the same. It is easy to see that oddly many adjacent transpositions negate the product, while evenly many do not effect it.

Consider now some arbitrary transposition, $\begin{pmatrix} i & j \end{pmatrix}$. Since $\begin{pmatrix} i & j \end{pmatrix} = \begin{pmatrix} j & i \end{pmatrix}$ it can be assumed that $i < j$. Now, this can be expressed as a product of adjacent transpositions as follows. Using the matrix analogy it is evident that:

- First of all move column i to column $j+1$, giving

$$\begin{pmatrix} i & i+1 \end{pmatrix} \begin{pmatrix} i+1 & i+2 \end{pmatrix} \dots \begin{pmatrix} j-1 & j \end{pmatrix}$$

  There is one transposition for each of the columns strictly between i and j, and another to put i above j, so there are $j - i - 1 + 1 = j - i$ transpositions here.

- Now move column j down to where column i was. The last transposition will be the one with column i, which used to be $i+1$. Hence the transpositions are

$$\begin{pmatrix} j-1 & j-2 \end{pmatrix} \begin{pmatrix} j-2 & j-3 \end{pmatrix} \dots \begin{pmatrix} i+1 & i \end{pmatrix}$$

  There is one transposition for each of the columns (that used to be) strictly between i and j, i.e. $j - i - 1$.

- In total this makes $2(j-i) - 1$ adjacent transpositions.

Hence any transposition is a product of oddly many adjacent transpositions, and so the product negates. Now, the product cannot both negate and stay the same, hence an odd transposition—as is claimed to have been found here—necessarily cannot be expressed as evenly many adjacent transpositions. By the definition this truly is an odd permtation. Furthermore,

- A transposition is a product of oddly many adjacent transpositions. Hence
  - a product of oddly many transpositions is a product of oddly many adjacent transpositions, and so is odd. (An odd number of odd numbers is odd.)
  - a product of evenly many transpositions is a product of evenly many adjacent transpositions and so is even. (An even number of odd numbers is even.)

Any permutation can be expressed as a product of disjoint cycles, and observe that any cycle can be written as a product of transpositions,

$$\begin{pmatrix} 1 & 2 & 3 & \dots & n \end{pmatrix} = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 3 \end{pmatrix} \begin{pmatrix} 3 & 4 \end{pmatrix} \dots \begin{pmatrix} n-1 & n \end{pmatrix}$$

There are $n-1$ transpositions here—one beginning with each number from 1 to $n-1$. Hence a cycle of even length is odd, and a cycle of odd length is even.

Since for products of transpositions

$$\begin{aligned} \text{even} \times \text{even} &= \text{even} \\ \text{even} \times \text{odd} &= \text{odd} \\ \text{odd} \times \text{even} &= \text{odd} \\ \text{odd} \times \text{odd} &= \text{even} \end{aligned}$$

So the parity of a permutation can be deduced from the parity of its disjoint cycles.

The Alternating Group

Consider the symmetric group S_n and let $\pi \in S_n$. Claim the set $A_n = \{\pi \in S_n \mid \pi \text{ is even}\}$ is a subgroup of S_n. Let

$$\pi = \tau_1\tau_2\ldots\tau_{2k} \qquad \sigma = \upsilon_1\upsilon_2\ldots\upsilon_{2l}$$

where $k, l \in \mathbb{N}$, and τ_i and υ_j are transpositions. Then by using Theorem 19

(i) Since a transposition is its own inverse,

$$\pi^{-1} = \tau_{2k}\tau_{2k-1}\ldots\tau_1$$

which is a product of evenly many transpositions and so $\pi^{-1} \in A_n$.

(ii) Clearly

$$\sigma\pi = \tau_1\tau_2\ldots\tau_{2k}\upsilon_1\upsilon_2\ldots\upsilon_{2l}$$

is an even permutation since it is a product of evenly many transpositions. Hence $\sigma\pi \in A_n$.

Hence $A_n \leqslant S_n$.

Suppose that ρ is an odd permutation. Consider the set $A_n\rho = \{\sigma\rho \mid \sigma \in A_n\}$. Now, if $\sigma\rho$ is even,

$$\begin{aligned} \sigma\rho &\in A_n \quad \text{but } \sigma^{-1} \in A_n \text{ so} \\ \sigma^{-1}\sigma\rho &\in A_n \\ \rho &\in A_n \end{aligned}$$

which is a contradiction since ρ is odd and A_n contains only even permutations. The set $A_n\rho$ therefore contains only odd permutations.

Now suppose that π and ρ are odd permutations, and σ is an even permutation. Clearly a product of two odd permutations is an even permutation, so $\pi\rho^{-1} \in A_n$, say $\pi\rho^{-1} = \sigma$. But then

$$\pi\rho^{-1} = \sigma \;\Rightarrow\; \pi = \sigma\rho \in A_n\rho$$

A product of even and odd permutations is therefore odd, and it is clear that every even permutation $\sigma \in A_n$ has a corresponding odd permutation, $\sigma\rho \in A_n\rho$ and vice versa i.e. a bijection. Hence $|A_n| = |A_n\rho|$. But since A_n and $A_n\rho$ are certainly disjoint, and since a permutation can be only odd or even, $S_n = A_n\dot{\cup}A_n\rho$ and so $\frac{1}{2}|S_n| = \frac{1}{2}n! = |A_n| = |A_n\rho|$.

As has already been shown, $A_n \leqslant S_n$. The same is not true of $A_n\rho$ which is not closed under multiplication since the product of two odd permutation is an even permutation.

(16.3.5) Numerology Of Groups

Modular Arithmetic

Finite groups have particular numbers of elements, and it is quite reasonable to suppose that some numbers are 'not allowed' since, say, including another element would violate one of the axioms unless some more elements are included as well.

In particular with cyclic groups, say in C_5, $x^{12} = x^7 = x^2$ so 12, 7, and 2 are all somehow 'the same'. This concept is the subject of modular arithmetic.

Definition 30 *Define a relation on $\mathbb{Z}$ by*

$$r \sim s \;\Leftrightarrow\; m\,|\,(r-s) \;\Leftrightarrow\; r-s = mk \;\Leftrightarrow\; r = s+mk$$

for some $k \in \mathbb{Z}$.

Claim that this is an equivalence relation, which is shown as follows.

reflexive: $r - r = 0 = 0m$ so $m\,|\,(r-r)$ i.e. $r \sim r$.

symmetric: $r \sim s \Leftarrow r - s = mk$ for $k \in \mathbb{Z}$. Clearly $s - r = m(-k)$ so $s \sim r$.

transitive: Suppose $r \sim s$ and $s \sim t$. Then

$$\begin{aligned} \exists k \in \mathbb{Z} \text{ such that } r - s &= mk \\ \exists l \in \mathbb{Z} \text{ such that } s - t &= lk \\ \text{adding gives } r - t &= (m+l)k \end{aligned}$$

Since $l, m \in \mathbb{Z}$, $(m+l) \in \mathbb{Z}$ and so $r \sim t$.

Hence this is an equivalence relation.

Being an equivalence relation, it partitions $\mathbb{Z}$ into disjoint equivalence classes,

$$[r] = \bar{r} = \{s \mid s \sim r\}$$

It is evident that

- $[0]$ is the multiples of m.
- $[1]$ is 1 plus a multiple of m.
- $[r]$ is r plus a multiple of m. This is written as $m\mathbb{Z} + r$.

Since $[m] = [0]$, $[m+1] = [1]$ etc. there are at most m equivalence classes. Furthermore, if $0 \leqslant a, b \leqslant m-1$ then $m\,|\,(b-a) \Leftrightarrow a = b$ so all the equivalence classes are distinct. Hence

$$\mathbb{Z} = m\mathbb{Z} \,\dot\cup\, m\mathbb{Z} + 1 \,\dot\cup \ldots \dot\cup\, m\mathbb{Z} + m - 1$$

The set of all these equivalence classes is denoted by $\mathbb{Z}_m$, the integers modulo m.

Addition and multiplication in $\mathbb{Z}_m$ are defined in the obvious way, $[a] + [b] = [a+b]$ and $[a][b] = [ab]$. However it is vital to check that these are well-defined in that they do not change if different values from $[a]$ and $[b]$ are chosen. Consider $a' = \alpha m + a$ and $b' = \beta m + b$ so

$$\begin{aligned} [a'] + [b'] &= [a' + b'] \\ &= [\alpha m + a + \beta m + b] \\ &= [(\alpha + \beta)m + (a+b)] \\ &= [a+b] \\ &= [a] + [b] \end{aligned} \qquad \begin{aligned} [a'][b'] &= [a'b'] \\ &= [(\alpha m + a)(\beta m + b)] \\ &= [m(\alpha\beta m + \alpha + \beta) + ab] \\ &= [ab] \\ &= [a][b] \end{aligned}$$

In $\mathbb{Z}_m$ under addition, the identity element is $[0]$ and and inverses are of the form $[m-a]$ since

$$[a]+[m-a]=[a+m-a]=[m]=[0]$$

It is clear that $\mathbb{Z}_m$ is an abelian group under addition.

Clearly $\mathbb{Z}$ is not a group under multiplication since there are no inverses—the rationals. However, this is not so with the integers modulo m since for example if $m=7$, $[4][2]=[8]=[1]$ and $[1]$ is the identity. However, there is a problem with $[0]$ since $[r][0]=[0]\ \forall r$, and inparticular for $r=1$. Clearly this will not do, so consider $\mathbb{Z}_m \setminus [0]$.

Having excluded $[0]$ there is a problem if $[r][s]=[m]$ for $1<r,s<m$. Since $[m]=[0]\not\subseteq \mathbb{Z}_m$ this means that the closure axiom is violated. In order to prevent this it is required that m is a prime, p.

Assertion 31 *If* $\gcd(x,y)=1$ *then* $\exists a,b\in\mathbb{Z}$ *such that* $ax+by=1$.

Consider $r\in\mathbb{Z}_p\setminus[0]$ and assume that $1\leqslant r\leqslant p-1$. Since p is prime, r and p are certainly coprime and so $\exists a,b\in\mathbb{Z}$ such that $ar+bp=1$. Hence

$$[1]=[ar+bp]=[a][r]+[b][p]=[a][r]$$

So $[r]$ has an inverse, $[a]\in\mathbb{Z}_p\setminus[0]$. Hence all the axioms hold and $\mathbb{Z}_p\setminus[0]$ is a group under multiplication.

Example 32 *Write down the multiplication table for the group* $\mathbb{Z}_5\setminus[0]$ *under multiplication.*

Proof. Solution The table suggests, and rightly so, that $\mathbb{Z}_p\setminus[0]\cong C_{p-1}$. Any element other than the identity,

	[1]	[2]	[3]	[4]
[1]	[1]	[2]	[3]	[4]
[2]	[2]	[4]	[1]	[3]
[3]	[3]	[1]	[4]	[2]
[4]	[4]	[3]	[2]	[1]

$[1]$, could be chosen as the generating element for example using $[3]$,

$$[3]^1=[3]\qquad [3]^2=[9]=[4]\qquad [3]^3=[27]=[2]\qquad [3]^4=[81]=[1]$$

□

Definition 33 *Where* $\mathbb{Z}_m$ *is the set of integers modulo m, let*

$$\mathcal{U}(\mathbb{Z}_m)=\{[a]\in\mathbb{Z}_m \mid \exists[b] \text{ such that } [a][b]=[1]\}$$

$\mathcal{U}(\mathbb{Z}_m)$ *is called the units of* $\mathbb{Z}_m$.

For example, the units of $\mathbb{Z}_{12}$ can be found by considering the multiples of 12, plus 1, i.e.

$$1\quad 13\quad 25\quad 37\quad 49\quad 61\quad 73\quad 85\quad 97\quad 109\quad 121\quad 133$$

The units of $\mathbb{Z}_{12}$ must certainly have the property that their square is 1 i.e. one of the numbers above. From this it is deduced that $\mathcal{U}(\mathbb{Z}_{12})=\{[1],[5],[7],[11]\}$.

$\mathcal{U}(\mathbb{Z}_8)=\{[1],[3],[5],[7],[11]\}$ is a set of three elements of order 2 and the identity. This is Klein's four-group, V_4. Some sets of units are groups while some are not.

(16.3.6) Group Actions

The chapter started by discussing the symmetries of a square which may be represented by D_4. A group may therefore act on a set—a group action.

Definition 34 *The group action of a group G on a set X is a function* $f\colon G \times X \to X$ *where*

1. $x^e = x \; \forall x \in X$.
2. $x^{gh} = (x^g)^h \;\; \forall x \in X \; \forall g, h \in G$.

When an element g acts on an element x is is common to write x^g.

Definition 35 *The orbit of a point x of a set X acted on by a group G is the set*

$$\operatorname{orb}_G(x) = \{x^g \mid g \in G\}$$

Definition 36 *The stabiliser of a point x of a set X acted on by a group G is the set*

$$\operatorname{stab}_G(x) = \{g \in G \mid x^g = x\}$$

Clearly $e \in \operatorname{stab}_G(x) \; \forall x \in X$ and for all groups.

Theorem 37 *For a group G acting on a set X with element x,* $\operatorname{stab}_G(x) \leqslant G$.

Proof. $e \in \operatorname{stab}_G(x)$ so $\operatorname{stab}_G(x) \neq \emptyset$ so suppose $h, k \in \operatorname{stab}_G(x)$.
Therefore $x^h = x^k = x$. Now use the test for a subgroup as given in Theorem 19

i. To show that $k^{-1} \in \operatorname{stab}_G(x)$,

$$\begin{aligned} x^{k^{-1}} &= \left(x^k\right)^{k^{-1}} \\ &= x^{kk^{-1}} && \text{by the second property of group actions} \\ &= x^e \\ &= x && \text{by the first property of group actions} \end{aligned}$$

Hence $k^{-1} \in \operatorname{stab}_G(x)$

ii. To show that $hk \in \operatorname{stab}_G(x)$,

$$\begin{aligned} \left(x^h\right)^k &= a^k && \text{since } h \in \operatorname{stab}_G(x) \\ &= x && \text{since } k \in \operatorname{stab}_G(x) \end{aligned}$$

Hence $\left(x^h\right)^k = x^{hk} \in \operatorname{stab}_G(x)$ and the second condition holds.

Both conditions hold, so by the test for a subgroup the theorem is proved. □

Cosets & Lagrange's Theorem

Definition 38 *The right coset of a subgroup H of a group G is the set*

$$Hg = \{hg \mid h \in H\} \quad \textit{for some } g \in G$$

A left coset has the obvious definition.

Theorem 39 (Lagrange's Theorem) *If $H \leqslant G$ then $|H| \mid |G|$.*

Proof. Define the relation $x \sim y \iff xy^{-1} \in H$ for some $H \leqslant G$. This is now shown to be an equivalence relation.

reflexive: $x \sim x \iff xx^{-1} \in H$ but $xx^{-1} = e$ which is certainly in H. Hence $x \sim x$.

symmetry: Suppose $x \sim y$ then

$$\begin{aligned} xy^{-1} \in H &\iff \left(xy^{-1}\right)^{-1} \in H \quad \text{since } H \text{ is a group it has inverses} \\ &\iff yx^{-1} \in H \\ &\iff y \sim x \end{aligned}$$

transitive: Suppose $x \sim y$ and $y \sim z$ then

$$\begin{aligned} \left(xy^{-1} \in H\right) \wedge \left(yz^{-1} \in H\right) &\iff \left(xy^{-1}\right)\left(yz^{-1}\right) \in H \quad \text{since } H \text{ is a group} \\ &\iff xz^{-1} \in H \\ &\iff x \sim z \end{aligned}$$

Hence this is verified as being an equivalence relation. Now consider the equivalence classes.

$$\begin{aligned} [y] &= \{x \in G \mid x \sim y\} \\ &= \left\{x \in G \mid xy^{-1} \in H\right\} \\ &= \left\{x \in G \mid xy^{-1} = h \in H\right\} \\ &= \{x \in G \mid x = hy \text{ for } h \in H\} \\ &= \{hy \mid h \in H\} \\ &= Hy \end{aligned}$$

So the equivalence classes are the cosets. Note that H is any old subgroup of G. Since the equivalence classes are all disjoint, for any particular H is is possible to find elements $x_1, x_2, \ldots, x_r$ such that

$$G = \dot{\bigcup_{i=1}^{r}} Hx_i$$

Now claim that $|Hx| = |H|$ and consider the mapping $\phi\colon H \to Hx$ given by $h^{\phi} = hx$.
By cancellation, $h_1 x = h_2 x \iff h_1 = h_2$ so ϕ is injective.
For surjectivity, observe that by definition $k \in Hx$ means that $\exists h \in H$ such that $k = hx$. Hence ϕ is surjective, and so is a bijection. The claim is therefore true.

Now,

$$G = \dot{\bigcup_{i=1}^{r}} Hx_i \quad \text{so} \quad |G| = \left|\dot{\bigcup_{i=1}^{r}} Hx_i\right| = \sum_{i=1}^{r} |Hx_i| = \sum_{i=1}^{r} |H| = r|H|$$

Hence $|G|$ is a multiple of $|H|$ whenever $H \leqslant G$ and so the theorem is proved. □

Lagrange's Theorem has some quite surprising consequences. If $O(x) = m$, the order of x, then $\langle x \rangle \cong C_m$ and $|\langle x \rangle| = m$. But $\langle x \rangle \leqslant G$, and hence $O(x) \mid |G|$ and this must hold for all $x \in G$.

For an element x of order m it is clear that $x^2, x^3, \ldots, x^{m-1}$ are also elements of order m, and so such elements come in 'batches' of $m-1$. Using this, notice that a group of order 12 cannot contain only the identity and elements of order 6 since 5 does not divide 11.

The Orbit-Stabiliser Theorem

Lagrange's theorem shows that the order of a subgroup divides the order of the group it comes from. It was also shown that the stabiliser of an element under a group action is a subgroup so there is clearly some kind of relationship to explore here.

Let G be a group acting on a set Λ. Recall

$$\operatorname{orb}_G(a) = \{a^g \mid g \in G\} \qquad \operatorname{stab}_G(a) = \{g \in G \mid a^g = a\}$$

Theorem 40 (The Orbit-Stabiliser Theorem) *For a group G acting on a set Λ,*

$$|G| = |\operatorname{stab}_G(a)|\,|\operatorname{orb}_G(a)|$$

Proof. Recall that $\operatorname{stab}_G a \leqslant G$. From the proof of Lagrange's Theorem it follows that $|G|$ can be partitioned as

$$G = \dot{\bigcup_{i=1}^{r}} (\operatorname{stab}_G(a))\, g_i$$

where $g_1 = e$. It also follows that

$$|G| = \left| \dot{\bigcup_{i=1}^{r}} (\operatorname{stab}_G(a))\, g_i \right| = \sum_{i=1}^{r} |(\operatorname{stab}_G(a))\, g_i| = r\,|\operatorname{stab}_G(a)| \tag{41}$$

Consider now the coset $(\operatorname{stab}_G(a))\, g$ for some $g \in G$ and for fixed $a \in \Lambda$. Claim that $(\operatorname{stab}_G(a))\, g = \{y \in G \mid a^y = a^g\}$. To prove this it is shown that $z \in (\operatorname{stab}_G(a))\, g \Leftrightarrow a^z = a^g$.

- $\Rightarrow$ Suppose $z \in (\operatorname{stab}_G(a))\, g$, then $z = xg$ for $x \in \operatorname{stab}_G(a)$. Hence $a^z = a^{xg} = a^g$ and the implication is proved.
- $\Leftarrow$ Suppose $a^z = a^g$ then $a^{zg^{-1}} = a$ so $zg^{-1} \in \operatorname{stab}_G a$.
  Say $zg^{-1} = x$ so $z = xg \in (\operatorname{stab}_G(a))\, g$, as required.

Hence the claim holds. From this it is evident that each a^{g_i} is one of the elements of $\operatorname{orb}_G a$, so

$$\operatorname{orb}_G a = \{a^{g_1} = a^e, a^{g_2}, \ldots, a^{g_r}\} \quad \text{giving} \quad |\operatorname{orb}_G a| = r$$

Hence from equation (41),

$$|G| = |\operatorname{stab}_G(a)|\,|\operatorname{orb}_G a|$$

and hence the theorem is proved. □

Clearly each coset $(\operatorname{stab}_G(a))\, g$ is the set of all elements of G which take a to a particular point in its orbit.

Definition 42 $[G : H] = \dfrac{G}{H} = \{Hg \mid g \in G\}$

The relationship mentioned above is a special case of a more general result. It turns out that Hg contains all the elements of G which take a to a^g. This is shown as follows.

Define $\phi\colon [G : H] \to \operatorname{orb}_G a$ by $\phi\colon Hg \mapsto a^g$.
First of all it is necessary to show that this is well-defined.

$$Hg = Hk \Leftrightarrow Hgk^{-1} = H \quad \text{so} \quad gk^{-1} = e \quad \text{and so} \quad a^{gk^{-1}} = a \quad \text{giving} \quad a^g = a^k$$

Hence $g \in Hk$ and $k \in Hg$ so ϕ is well-defined.
Clearly ϕ is surjective, since any element of $\operatorname{orb}_G a$, a^g say, is the image of the corresponding coset, Hg.
Finally, by reversing the above argument it is evident that ϕ is injective. Hence ϕ is bijective.

The well-definition and bijectivity of ϕ show that the assertion that Hg is the subset of G of all elements that take a to a^g is proved.

Symmetries Of The Cube

The cube has 48 symmetries, 24 rotational and another 24 reflective. They can be described in terms of the faces, the vertices, or the diagonals. Figure 16.3.6 shows on the left the labeled vertices of the cube, and on the right a diagram showing the faces.

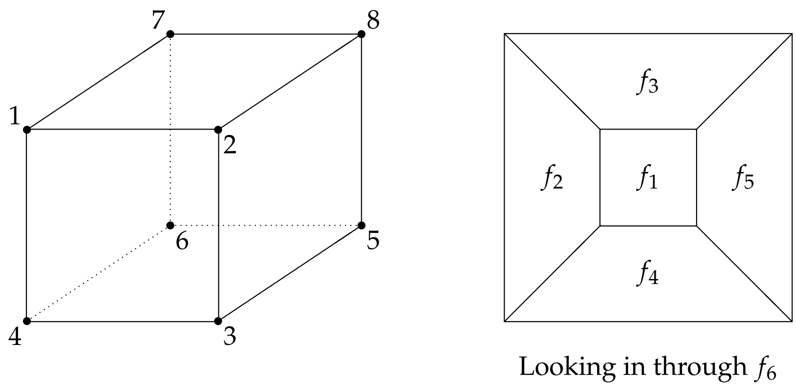

Figure 4: A cube, showing labeled vertices and faces.

For example, it is clear that $\operatorname{stab}_G 1$ is the three rotations which permute vertices 2, 4, and 7, and the reflections in the three diagonal planes through 1. Including the identity, $|\operatorname{stab}_G 1| = 6$. Clearly vertex 1 has all 8 positions in its orbit, so $|\operatorname{orb}_G 1| = 8$. From Lagrange's theorem this gives $|G| = 6 \times 8 = 48$, as was claimed.

Definition 43 *If when a group G acts on a set Λ there is just one orbit (so that an element $a \in \Lambda$ can be sent to any other element of Λ i.e. $\operatorname{orb}_G a = \Lambda$), then it is said that G acts transitively on Λ.*

Definition 44 *The fix in Λ of an element $g \in G$ is the set*

$$\operatorname{fix} g = \{a \in \Lambda \mid a^g = a\}$$

The fix is therefore like the stabiliser, but with G and Λ changed round.

Theorem 45 $\sum\limits_{a\in\Lambda} |\operatorname{stab}_G a| = \sum\limits_{g\in G} |\operatorname{fix} g|$

Proof. Consider the set $P = \{(a, g) \mid a \in \Lambda,\ g \in G,\ a^g = a\}$.

- For each a the number of elements with the property $a^g = a$ is precisely $|\operatorname{stab}_G(a)|$. Hence $|P| = \sum\limits_{a\in\Lambda} |\operatorname{stab}_G a|$.

- For each g the number of elements with the property $a^g = a$ is precisely $|\text{fix}\, g|$. Hence $|P| = \sum_{g\in G} |\text{fix}\, g|$.

Since the same things have been counted in these two different ways, the theorem is proved. □

As an example, consider the rotational symmetries of the cube, as are given in Table 16.3.6.

Type of symmetry	No. of	Cycle shape when acting on Vertices	Faces	Diagonals
$\frac{\pi}{2}$ through centres of faces	6	4^2	$1^2 4$	4
π through centres of faces	3	2^4	$1^2 2^2$	2^2
π through centres of edges	6	2^4	2^3	$1^2 2$
$\frac{2\pi}{3}$ through vertices	8	$1^2 3^2$	3^2	1.3
Identity	1	1^8	1^6	1^4

Table 2: Rotational symmetries of the cube.

Theorem 45 can be applied as follows. To find $\sum_{g\in G} |\text{fix}\, g|$ simply count the number of cycles of length 1 which appear in (any one column of) Table 16.3.6. For all three of the columns this is 24.

In general, $\Lambda = \dot{\bigcup}_{i=1}^{r} \text{orb}_G\, a_i$ i.e. there are several orbits. Hence

$$\sum_{a\in\Lambda} |\text{stab}_G\, a| = \sum_{i=1}^{r} \sum_{a\in a_i^G} |\text{stab}_G\, a_i| = \sum_{i=1}^{r} \left|a_i^G\right| |\text{stab}_G\, a_i| = r|G|$$

with the last step following from the orbit-stabiliser theorem. In the case of the cube there is only one orbit i.e. G acts transitively. This gives $\sum_{a\in\Lambda} |\text{stab}_G\, a| = |G| = 24$ which has already been shown from Theorem 45.

(16.3.7) Conjugacy & Centrality

Theory

Notice that the symmetries of the cube fell naturally into a number of categories e.g. the rotations through π about an axis through opposite faces.

Definition 46 *For a group G, $x, y \in G$ are conjugate $\Leftrightarrow \exists g \in G$ such that $x = g^{-1}yg$.*

It can be shown that conjugacy is an equivalence relation as follows.

reflexive: Choosing $g = e$, $e^{-1}xe = x$ so $x \sim x$ as required.

symmetric: $x \sim y \;\Rightarrow\; x = g^{-1}yg \;\Rightarrow\; y = gxg^{-1} = \left(g^{-1}\right)^{-1} xg^{-1}$. Now putting "$g = g^{-1}$" it is evident that $y \sim x$, as required.

transitive: Suppose $x \sim y$ and $y \sim z$ then $\exists g, h \in G$ such that

$$\begin{aligned} x &= g^{-1}yg \quad \text{and} \quad y = h^{-1}zh \\ &= g^{-1}h^{-1}zhg \\ &= (hg)^{-1} z(hg) \end{aligned}$$

So where $k = hg$, $\exists k \in G$ such that $x = k^{-1}zk$ and hence $x \sim z$.

Any group is therefore partitioned into disjoint conjugacy classes. Each class is of the form

$$\begin{aligned}
\mathrm{Cl}_G(x) &= \{y \in G \mid x \sim y\} \\
&= \left\{y \in G \mid x = g^{-1}yg\right\} \\
&= \left\{gxg^{-1} \mid g \in G\right\} \\
&= \left\{g^{-1}xg \mid g \in G\right\} \\
&= \{x^g \mid g \in G\} = x^G
\end{aligned}$$

The final expression follows by the symmetry of the equivalence relation. Notice that this looks rather like an orbit, and this is no coincidence since conjugation in this way is a group action where G acts on itself. This is readily shown since where $x^g = g^{-1}xg$, it is trivial that $x^e = x$ so the first group action property holds. For the second,

$$\begin{aligned}
x^{gh} &= \left(h^{-1}g^{-1}\right)x(gh) \quad \text{since } (gh)^{-1} = h^{-1}h^{-1} \\
&= h^{-1}\left(g^{-1}xg\right)h \\
&= (x^g)^h
\end{aligned}$$

So this equivalence relation is also a group action. The orbits are the conjugacy classes.

Theorem 47 *Elements in the same conjugacy class have the same order.*

Proof. Suppose $x \sim y$ so that $x = g^{-1}yg$, and that $O(y) = m$. Then

$$\begin{aligned}
x^m &= \left(g^{-1}yg\right)^m \\
&= \left(g^{-1}yg\right)\left(g^{-1}yg\right)\dots\left(g^{-1}yg\right) \\
&= g^{-1}y^m g \\
&= g^{-1}eg \\
&= e
\end{aligned}$$

Hence $O(x) \leqslant O(y)$. But since also $y \sim x$ it must be the case that $O(y) \leqslant O(x)$ and hence the theorem is proved. □

It follows that the identity element is in a conjugacy class of its own.

Definition 48 *The centraliser in a group G of an element x is the set*

$$\mathrm{C}_G(x) = \{g \in G \mid xg = gx\}$$

The centaliser is the stabiliser of the group action since $xg = gx$ means that $x = g^{-1}xg$. The centraliser may be expressed as $\{g \in G \mid x^g = x\}$.

Theorem 49 $\mathrm{C}_G(x) \leqslant G$.

Proof. Certainly $\mathrm{C}_G(x) \neq \emptyset$ since $e \in C_G(x)$.
Consider $g, h \in \mathrm{C}_G(x)$ then

$$(gh)x = g(hx) = g(xh) = (gx)h = (xg)h = x(gh)$$

so $gh \in C_G(x)$ so it is closed under taking products.
Also,

$$xg = gx \;\Rightarrow\; g^{-1}(xg)g^{-1} = g^{-1}(gx)g^{-1} \;\Rightarrow\; g^{-1}x = xg^{-1}$$

so $g^{-1} \in C_G(x)$.
Hence by the test for a subgroup the theorem is proved. □

Suppose that g and h are in different conjugacy classes (orbits), then

$$\begin{aligned} & g^{-1}xg \neq h^{-1}xh \quad \text{for some } x \in G \\ \Leftrightarrow\; & \left(hg^{-1}\right) x \left(gh^{-1}\right) \neq x \\ \Leftrightarrow\; & \left(gh^{-1}\right)^{-1} x \left(gh^{-1}\right) \neq x \\ \Leftrightarrow\; & gh^{-1} \notin C_G(x) \\ \Leftrightarrow\; & C_G(x)\, gh^{-1} \neq C_G(x) \\ \Leftrightarrow\; & C_G(x)\, g \neq C_G(x)h \end{aligned}$$

What this shown (due to the double implication) is that there are the same number of conjugacy classes (orbits) as there are cosets of the centraliser (stabiliser).
The proof of the orbit-stabiliser theorem shows that the order of any subgroup divides $|G|$ by a factor equal to the number of cosets of that subgroup. Hence

$$|G| = |C_G(x)|\,|\mathrm{Cl}_G(x)|$$

Of course, this result could have been deduced from the orbit-stabiliser theorem.

The centraliser of an element x is the set of all elements of the group with which it commutes, clearly it is of interest as to whether it is ever the case that $C_G(x) = G$ i.e. if x commutes with every element of G.

Definition 50 *The centre of a group G is the set of elements of G which commute with every other element of G, so*

$$Z(G) = \{x \in g \mid xg = gx \;\forall g \in G\}$$

It is readily shown that the centre of a group is also a subgroup. The centraliser $C_G(x)$ contains the gs which commute with some particular x, so considering each x in turn, it is clear that

$$Z(G) = \bigcap_{x \in G} C_G(x)$$

Now, suppose $\left|x^G\right| = |\mathrm{Cl}_G(x)| = 1$ then

$$\begin{aligned} & \left|\left\{g^{-1}xg \mid g \in G\right\}\right| = 1 \\ \text{so} \quad & g^{-1}xg = x \quad \forall g \in G \\ \text{i.e.} \quad & xg = gx \quad \forall g \in G \end{aligned}$$

Hence $\left|x^G\right| = 1 \;\Leftrightarrow\; x \in Z(G)$.

Definition 51 *The conjugacy classes are defined in terms of an equivalence relation, so for $x_1, x_2, \ldots, x_r$ in G it follows that*

$$G = x_1^G \,\dot\cup\, x_2^G \,\dot\cup \ldots \dot\cup\, x_r^G$$

hence, define the class equation of G to be

$$|G| = \left|x_1^G\right| + \left|x_2^G\right| + \cdots + \left|x_r^G\right|$$

Theorem 52 *If $|G| = p^m$ where p is prime, then $|Z(G)| \geqslant p$ i.e. the centre of G is non-trivial.*

Proof. Consider the class equation of G, $G = x_1^G \,\dot\cup\, x_2^G \,\dot\cup \ldots \dot\cup\, x_s^G \,\dot\cup \ldots \dot\cup\, x_r^G$ written in such a way that $\left|x_i^G\right| = 1$ for all $1 \leqslant i \leqslant s$.

$$\text{Hence } \begin{cases} x_i \in Z(G) & 1 \leqslant i \leqslant s \\ x_i \notin Z(G) & s < i \leqslant r \end{cases} \qquad \text{and so } G = Z(G) \,\dot\cup\, x_{s+1}^G \,\dot\cup \ldots \dot\cup\, x_r^G$$

which gives

$$p^m = |G| = |Z(G)| + \left|x_{s+1}^G\right| + \cdots + \left|x_r^G\right| \tag{53}$$

Now, for all i the Orbit-Stabiliser theorem gives that $\left|x_i^G\right| \mid p^m$ and since p is prime this gives $\left|x_i^G\right| = p^{a_i}$.

Certainly p divides $|G|$ and so must divide the right hand side of equation (53). Each $\left|x_i^G\right|$ has been shown to be divisible by p, so it is concluded that $p \mid |Z(G)|$. □

Application To Symmetric Groups

For σ and τ in S_n it is clearly undesirable to have to calculate $\tau^{-1}\sigma\tau$ and indeed there is a useful result regarding this.
Let $\sigma = \begin{pmatrix} a_1 & a_2 & \ldots & a_r \end{pmatrix} \begin{pmatrix} b_1 & b_2 & \ldots & b_s \end{pmatrix} \cdots \begin{pmatrix} c_1 & c_2 & \ldots & c_t \end{pmatrix}$
then $\sigma^\tau = \tau^{-1}\sigma\tau = \begin{pmatrix} a_1^\tau & a_2^\tau & \ldots & a_r^\tau \end{pmatrix} \begin{pmatrix} b_1^\tau & b_2^\tau & \ldots & b_s^\tau \end{pmatrix} \cdots \begin{pmatrix} c_1^\tau & c_2^\tau & \ldots & c_t^\tau \end{pmatrix}$
This is because $(a_i^\tau)^{\tau^{-1}\sigma\tau} = a_i^{\sigma\tau} = a_{i+1}^\tau$, where usually $a_i^\sigma = a_{i+1}$. Note this means that σ and σ^τ have the same cycle shape, so a conjugacy class contains all the permutations of the same cycle shape. Note also that if $\sigma = \begin{pmatrix} a_1 & a_2 & \ldots & a_r \end{pmatrix} \begin{pmatrix} b_1 & b_2 & \ldots & b_s \end{pmatrix} \cdots \begin{pmatrix} c_1 & c_2 & \ldots & c_t \end{pmatrix}$
and $\rho = \begin{pmatrix} a_1' & a_2' & \ldots & a_r' \end{pmatrix} \begin{pmatrix} b_1' & b_2' & \ldots & b_s' \end{pmatrix} \cdots \begin{pmatrix} c_1' & c_2' & \ldots & c_t' \end{pmatrix}$
then $\rho = \sigma^\tau$ where

$$\tau = \begin{pmatrix} a_1 & a_2 & \ldots & a_r & b_1 & b_2 & \ldots & b_s & c_1 & c_2 & \ldots & c_t \\ a_1' & a_2' & \ldots & a_r' & b_1' & b_2' & \ldots & b_s' & c_1' & c_2' & \ldots & c_t' \end{pmatrix}$$

It has already been shown that $|G| = |\mathrm{Cl}_G(x)||C_G(x)|$, but for symmetric groups it is also true that where σ has cycle shape $a_1^{r_1} a_2^{r_2} \ldots a_s^{r_s}$ then

$$|C_G(x)| = a_1^{r_1} a_2^{r_2} \ldots a_s^{r_s} r_1! r_2! \ldots r_2! \tag{54}$$

Knowing the number of elements in a centraliser can make it easy to specify precisely what the elements are.

(16.3.8) Normal Subgroups

Definition 55 *Let N be a subgroup of a group G. Then if*

$$g^{-1}ng \in N \quad \forall n \in N \quad \forall g \in G$$

then N is said to be a normal subgroup of G, written $N \lhd G$.

In particular the kernel of any homomorphism is a normal subgroup of the domain of the homomorphism. For $\phi\colon G \to H$ and $k \in \ker\phi$,

$$\begin{aligned}\left(g^{-1}kg\right)^{\phi} &= \left(g^{-1}\right)^{\phi} k^{\phi} g^{\phi} \\ &= \left(g^{-1}\right)^{\phi} g^{\phi} \\ &= e\end{aligned}$$

Hence $k^G \subseteq \ker\phi$ and $\ker\phi = \bigcup_{k\in\ker\phi} k^G$, a union of complete conjugacy classes.

Theorem 56 *If $N \lhd G$ then $gN = Ng\ \forall g \in G$, i.e. left and right cosets of N are equal.*

Proof.

$$\begin{aligned} N \lhd G \Leftrightarrow&\quad g^{-1}ng \in N \quad \forall n \in N \quad \forall g \in G \\ \Leftrightarrow&\quad g^{-1}Ng \subseteq N \quad \forall g \in G \\ \text{but replacing } g \text{ by } g^{-1}, \quad \Leftrightarrow&\quad gNg^{-1} \subseteq N \quad \forall g \in G \\ \Leftrightarrow&\quad N \subseteq g^{-1}Ng \end{aligned}$$

By showing both inclusions it follows that $N = g^{-1}Ng$ i.e. $gN = Ng$, as required. □

Since $N = g^{-1}Ng$, N is a union of complete conjugacy classes, $|N|$ must not only divide $|G|$ (from Lagrange's theorem), but also be the sum of orders of conjugacy classes. The conjugacy classes of S_4 are given in Table 16.3.8.

Cycle shape	1^4	$1^2 2$	1.3	4	2^2
$\lvert \mathrm{Cl}_G(x) \rvert$	1	$\binom{4}{2} = 6$	$4.2 = 8$	$3! = 6$	$\frac{1}{2}\binom{4}{2} = 3$

Table 3: Conjugacy classes for S_4

Any subgroup must contain the identity, indeed any group has the two normal subgroups of itself and $\langle e \rangle$. The only odd numbers available are 7 and 9, nether of which divide 24, so the 3 cycles of shape 2^2 must be included. $4\,|\,24$ so one normal subgroup is

$$N_1 = \left\{e, \begin{pmatrix}1 & 3\end{pmatrix}\begin{pmatrix}2 & 4\end{pmatrix}, \begin{pmatrix}1 & 3\end{pmatrix}\begin{pmatrix}2 & 4\end{pmatrix}, \begin{pmatrix}1 & 4\end{pmatrix}\begin{pmatrix}2 & 3\end{pmatrix}\right\}$$

This is in fact the Klein 4-group, V_4. Other possibilities for the size of sets are 4, 8, and 12. Only 12 divides 24 so it follows that the other normal subgroup of S_4 contains also the cycles of shape 1.3.

As another example consider A_5. Now, in any permutation group conjugate elements have the same cycle shape. Unfortunately it does not follow that all elements of the same cycle shape are conjugate; this is the case in A_5. Clearly there are $4! = 24$ 5-cycles in A_5 but $24 \nmid 60$.

Cycle shape	1^5	$1^2 3$	1.2^2	5	5
$\lvert \mathrm{Cl}_G(x) \rvert$	1	$2\binom{5}{2} = 20$	$5\frac{1}{2}\binom{4}{2} = 15$	12	12

Table 4: Conjugacy classes for A_5.

It is deduced that there are two conjugacy classes with 12 elements with permutations with cycle shape 5 by using the orbit-stabiliser theorem i.e. $|G| = |\mathrm{Cl}_G(x)|\,|C_G(x)|$. From equation (54), the order of the centraliser of a 5-cycle must be 5. $60 \div 5 = 12$, hence the result. Other than the trivial normal subgroups, it is evident that A_5 has no normal subgroups.

A group with only the trivial normal subgroups is called simple, although in practise these can be the most difficult groups to work with. In particular the Monster group with over 10^{55} elements has only recently been shown to be simple, although the proof is not fully complete yet.

Theorem 57 *Any subgroup of an abelian group is normal.*

Proof. Observe that $g^{-1}ng = ng^{-1}g = n$ which is certainly in N. □

Theorem 58 *If $H \leqslant G$ and $|[G : H]| = 2$ then H is normal.*

Proof. Well, if $g \in H$ then $Hg = gH$ and the result is proved.
If $g \notin H$ then since there are only two cosets, one of which must be H itself it follows that

$$G = H \cup Hg = H \cup gH$$

hence it must be the case that $Hg = gH$ and the remaining case is proved. □

(16.3.9) Factor Groups

The property of a normal subgroup that left and right cosets are equal allows them to form a group. Let $N \triangleleft G$ and consider $[G : N]$, the cosets of N in G. Claim that $[G : N]$ is a group under the operation '$*$' defined as $Ng * Nh = N(gh)$.

First of all, this binary operation must be shown to be well-defined. For sets A and B,

$$AB \stackrel{\text{def}}{=} \{ab \mid a \in A \quad b \in B\}$$

So considering Ng and Nh as subsets of G,

$$\begin{aligned}(Ng)(Nh) &= (gN)(Nh)\\ &= g(NN)h\\ &= (gN)h\\ &= Ngh\end{aligned}$$

But since N is a subgroup $NN = N$, so '$*$' is a good definition. Now, there is generally at least one $k \in G$ such that $Ng = Nk$ i.e. some cosets are the same. It is therefore necessary to check that the definition of '$*$' does not depend on the choice of element to make the coset.
Suppose $Nh = Nh'$ and $Ng = Ng'$, so there exists n_1 and n_2 in N with $g' = n_1 g$ and $h' = n_2 h$. Hence

$$\begin{aligned}Ng'h' &= Nn_1gn_2h\\ &= Ngn_2h \quad \text{since } n_1 \in N \text{ gives } Nn_1 = N\\ &= Ngn_2g^{-1}gh\\ &= Nn_3gh \quad \text{since } gn_2g^{-1} \text{ is a conjugate of } n_2 \text{ and so is in } N\\ &= Ngh\end{aligned}$$

Hence the binary operation on $[G : N]$ is well-defined.

To show that $[G : N]$ is a group, the group axioms are verified.

closure: Clearly $Ng * Nh = Ngh \in [G : N]$.

associative: $(Ng * Nh) * Nk = Ngh * Nk = Nghk = Ng * Nhk = Ng * (Nh * Nk)$

identity: Claim the identity is $Ne = N$ so,

$$\begin{aligned} N * Ng &= Ne * Ng = Neg = Ng \\ Ng * N &= Ng * Ne = Nge = Ng \end{aligned}$$

inverse: Claim $(Ng)^{-1} = Ng^{-1}$

$$\begin{aligned} Ng * Ng^{-1} &= Ngg^{-1} = Ne = N \\ Ng^{-1} * Ng &= Ng^{-1}g = Ne = N \end{aligned}$$

So the group axioms hold. The group $[G : N]$ defined in this way is called the factor group of G by N and is written $\frac{G}{N}$.

Theorem 59 *A subgroup N of a group G is normal $\Leftrightarrow$ N is the kernel of a homomorphism.*

Proof. Suppose that $N \lhd G$ and define $\phi\colon G \to \frac{G}{N}$ by $\phi\colon g \mapsto Ng$. Certainly ϕ is well-defined and onto.

$$(gh)^\phi = Ngh = Ng * Nh = g^\phi h^\phi$$

hence ϕ is a homomorphism. Consider the kernel of ϕ,

$$\begin{aligned} \ker\phi &= \{g \in G \mid g^\phi = N\} \\ &= \{g \in G \mid Ng = N\} \end{aligned}$$

but $Ng = N \Leftrightarrow g \in N$, hence

$$\ker\phi = N$$

Hence the result. □

Theorem 60 (First Isomorphism Theorem) *Let $\phi\colon G \to H$ be a homomorphism, then* $\dfrac{G}{\ker\phi} \cong G^\phi$.

Proof. To prove this result let $K = \ker\phi$ and define $\psi\colon \frac{G}{K} \to G^\phi$ by $\psi\colon Kg \mapsto g^\phi$.
ψ is certainly onto, furthermore

$$(Kg * Kh)^\psi = (Kgh)^\psi = (gh)^\phi = g^\phi h^\phi = (Kg)^\psi (Kh)^\psi$$

hence ψ is a homomorphism. Now to complete the proof,

$$\begin{aligned} (Kg)^\psi = (Kh)^\psi \Leftrightarrow &\quad g^\phi = h^\phi \\ \Leftrightarrow &\left(gh^{-1}\right)^\phi = e_H \\ \Leftrightarrow &\quad gh^{-1} \in K \\ \Leftrightarrow &\quad Kgh^{-1} = K \\ \Leftrightarrow &\quad Kh = Kg \end{aligned}$$

Hence ψ is an isomorphism and the theorem is proved. □