From Geometry to Groups

David Mond

June 20, 2003

1 Introduction

These notes are the text of the course "From Geometry to Groups". Lectures will cover (I hope) all of the material here, but the division of the text into sections does not necessarily correspond to its division into lectures: most sections will occupy more than one lecture.

2 Definition and Examples

You will recall from *Foundations* that a group is a non-empty set together with a binary operation (denoted here by juxtaposition) satisfying the following axioms:

A1 Closure: For all $a, b \in G$, we have $ab \in G$.

A2 Associativity: For all $a, b, c \in G$, $(ab)c = a(bc)$

A3 Neutral element: There exists a (unique) element $e \in G$ (the *neutral* element, or *identity* element) such that for all $a \in G$, $ae = ea = a$

A4 Inverses: For all $a \in G$, there exists a (unique) element $a^{-1} \in G$ such that $aa^{-1} = a^{-1}a = e$.

But why should one study such objects? What can persuade our brains to allocate sufficient RAM (or mental modules, depending on the theory of mind you subscribe to) to process this material adequately? Will-power is rarely enough! For some people, the structures on their own are sufficiently fascinating for them to want to plunge straight in; others need to see the ideas in action, in interesting and important examples, to appreciate their power and significance. I belong to the seond category.

We will try to provide some reasons by looking at examples, mostly from geometry; appreciation for the definitions will come (I hope) through the interplay between the concrete and the abstract. The geometry we will do in this course is mostly of the no-holds-barred variety - provided it is convincing, we're not going to be particularly concerned about rigour. After all, in geometry you can often see whether something is true, even when you can't immediately lay your hands on a formal proof. On the other hand, in algebra one has no way of knowing except proof, and so we have to adopt a greater degree of formality.

Most of the examples you studied in Foundations fall into two classes:

I. Groups whose elements are "numbers", such as $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$; here the binary operations (addition, multiplication) arise essentially from the process of counting: the *number* of elements in the union of two disjoint sets is the *sum* of the number of elements in the two.

II. Groups whose elements are certain bijections (maps which are both injective and surjective) of some set X to itself, where the binary operation is *composition*.

The first example here is of course the set of *all* bijections $X \to X$. This is called $\mathrm{Bij}(X)$; if $X = \{1, 2, \cdots, n\}$, we call it S_n. In this case (and in general when X is a finite set), we usually refer to the elements of $\mathrm{Bij}(X)$ as *permutations* of X. It is easy to check that $\mathrm{Bij}(X)$ is a group: the composite of any two bijections $X \to X$ is another, so A1 holds; the two ways of bracketing the maps f, g and h, namely $f \circ (g \circ h)$ and $(f \circ g) \circ h$, are easily seen to do the same thing to any element of X, and thus are the same map, so associativity holds; the identity map $X \to X$ is the neutral element, and finally every bijection has an inverse, also a bijection, so A4 holds.

Some important examples for us fall in the second class: the groups $I(\mathbb{R}^2)$ and $I(\mathbb{R}^3)$ of *isometries* of $\mathbb{R}^2$ and $\mathbb{R}^3$. A map is an isometry if it is bijective, and for all points p, q, the distance $d(f(p), f(q))$ between $f(p)$ and $f(q)$ is the same as the distance $d(p, q)$ between p and q.

Example 2.1 Examples of isometries of the plane:

- Translation by a fixed vector v
- Reflection r_L in a line L
- Rotation $R_{P,\theta}$ through an angle θ about a point P.

In fact we lose nothing by dealing with the general case $\mathbb{R}^n$, where the distance between two points $a = (a_1, \cdots, a_n)$ and $b = (b_1, \cdots, b_n)$ is defined to be

$$((a_1 - b_1)^2 + \cdots + (a_n - b_n)^2)^{1/2}.$$

How does one show that $I(\mathbb{R}^n)$ is a group (under composition, of course)? The first thing one has to do is to check closure - that the composition of two isometries really is an isometry. But if $f(p)$ and $f(q)$ are as far apart as p and q (f is an isometry), and $g(f(p))$ and $g(f(q))$ are as far apart as $f(p)$ and $f(q)$ (g is an isometry) then of course $g(f(p))$ and $g(f(q))$ are as far apart as p and q. And we've already agreed that the composition of bijections is another bijection. So the closure axiom holds. You'll notice that here the fact that we are talking about isometries of $\mathbb{R}^n$ is completely irrelevant.

How about Associativity? We don't need to do anything at all here, because we have seen that composition (of maps of any kind at all) is an associative operation. Associativity is *inherited* from $\mathrm{Bij}(X)$.

Neutral element: the identity map evidently preserves distances!

Inverses: every isometry is a bijection, by definition, and hence it has an inverse map. To see that f^{-1} is an isometry, we reason as follows: $d(f^{-1}(p), f^{-1}(q)) = d(f(f^{-1}(p)), f(f^{-1}(q)))$ (as f is an isometry), and by the definition of inverse $f(f^{-1}(p)) = p, f(f^{-1}(q)) = q$, so $d(f^{-1}(p), f^{-1}(q)) = d(p, q)$ and f^{-1} is an isometry.[1]

We conclude that for each n, $I(\mathbb{R}^n)$ is a group. In fact, at no expense we can say a little more: for every n, $I(\mathbb{R}^n)$ is a *subgroup* of $\mathrm{Bij}(X)$. Recall that a sub*set* H of a group G is a sub*group* if it is a group in its own right, with respect to the group operation of G. Recall also from Foundations the following useful proposition:

Proposition 2.2 *Let G be a group and H a non-empty subset. Then H is a subgroup of G if it is closed under the binary operation of G, and if for every $g \in H$, the multiplicative inverse g^{-1} also lies in H.*

Proof Suppose that H is closed under the binary operation and contains the inverse of each of its elements. To show it is a group, it only remains to show associativity and the existence of a neutral element. But associativity is inherited from G, and as for the neutral element, it must lie in H, for if h is any element of H, then so is h^{-1}, and it follows that $e = hh^{-1} \in H$. □

Symmetries of a geometric figure

Let $X \subset \mathbb{R}^n$, and let $I(X)$ be the set of all isometries of $\mathbb{R}^n$ mapping X to itself:

$$I(X) = \{f \in I(\mathbb{R}^n) : f(X) = X\}.$$

Proposition 2.3 *If $X \subset \mathbb{R}^n$ then $I(X)$ is a subgroup of $I(\mathbb{R}^n)$.*

Proof Note that associativity is once again inherited. It is evident that if $f, g \in I(X)$ then $f \circ g \in I(X)$, and that the identity map id is in $I(X)$, so we need to show only that if f is in $I(X)$ then so is f^{-1}. In fact this too is clear: we already know that f^{-1} is an isometry, and it's now an easy exercise in set theory to show that if $f(X) = X$ then $f^{-1}(X) = X$ also. □

Example 2.4 We consider isometries of the square. We take it as evident that any isometry maps vertices to vertices and edges to edges. (Of course, an argument is easily supplied). It also maps opposite vertices to opposite vertices, and so fixes the centre, where the diagonals meet (this is of course the centre of mass referred to above). It must thus be a rotation R_θ (through an angle θ in an anticlockwise direction) about this centre, or a reflection r_L in a line L through the centre (by Corollary 2.7 below).

[1] In fact the assumption that $f : \mathbb{R}^n \to \mathbb{R}^n$ preserves distance implies that it is bijective, so in our definition of isometry we needn't have *assumed* bijectivity. You might like to try to prove this (as an exercise in geometry rather than group theory). Injectivity is easy; the hard part is surjectivity.

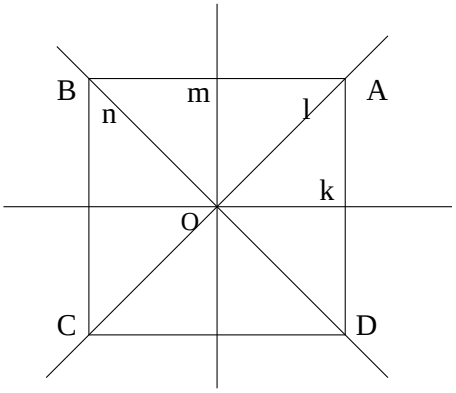


Figure 1: *The square and its lines of symmetry*

It's not hard to convince oneself that

$$I(Sq) = \{id, R_{\pi/2}, R_{\pi}, R_{3\pi/2}, r_K, r_L, r_M, r_N\}.$$

To justify this, and later calculations, we need a little geometry. The results we need are stated here without proof; proofs are consigned to the appendix to this section, and are optional as far as this course is concerned.

Proposition 2.5 *Let p, q and r be three non-collinear points in $\mathbb{R}^2$. Then any point x is uniquely determined by its distance from p, from q and from r.* □

If f is any map, we say f *fixes* p if $f(p) = p$.

Corollary 2.6 The 3-point rule
(i) Any isometry of the plane fixing three non-collinear points is the identity map.
(ii) If the plane isometries f and g agree on three non-collinear points then $f = g$. □

Corollary 2.7 *Any plane isometry f fixing a point O is either a rotation about O or a reflection in a line through O.* □

The composite of two elements of $I(Sq)$ is another element of $I(Sq)$. How to work out which? A simple way is to use the three-point rule: if two elements of $I(Sq)$ do the same thing to the vertices, then they are the same. This follows from 2.6(ii), applied to any 3 of the vertices.

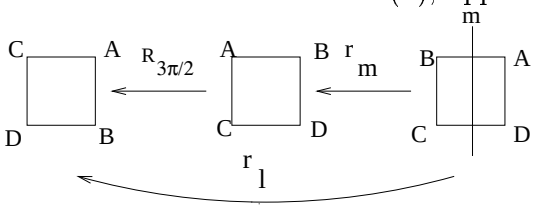


Figure 2: *The effect of $R_{3\pi/2} \circ r_M$ on the vertices is the same as the effect of r_L. Hence $R_{3\pi/2} \circ r_M = r_L$.*

	1	R	R^2	R^3	r_K	r_L	r_M	r_N
1	1	R	R^2	R^3	r_K	r_L	r_M	r_N
R	R	R^2	R^3	1	r_L	r_M	r_N	r_K
R^2	R^2	R^3	1	R	r_M	r_N	r_K	r_L
R^3	R^3	1	R	R^2	r_N	r_K	r_L	r_M
r_K	r_K	r_N	r_M	r_L	1	R^3	R^2	R
r_L	r_L	r_K	r_N	r_M	R	1	R^3	R^2
r_M	r_M	r_L	r_K	r_N	R^2	R	1	R^3
r_N	r_N	r_M	r_L	r_K	R^3	R^2	R	1

Composition table for $I(Sq)$.

In the table, and from now on, we write R for $R_{\pi/2}$ and R^2 and R^3 for R_π and $R_{3\pi/2}$ respectively.

The following proposition shows that the study of isometries can be seen as part of Linear Algebra: recall that a map $\mathbb{R}^n \to \mathbb{R}^n$ is *linear* if $T(v_1 + v_2) = T(v_1) + T(v_2)$, for all $v_1, v_2 \in \mathbb{R}^n$, and $T(\lambda v) = \lambda T(v)$ for all $v \in \mathbb{R}^n$ and all $\lambda \in \mathbb{R}$.

Proposition 2.8 *Suppose f is an isometry of $\mathbb{R}^n$ and $f(0) = 0$. Then f is a linear map.*

Proof Because f is an isometry, it maps straight lines to straight lines; for a straight line can be characterised in terms of distance - as the shortest route between any two of its points. In more detail: if p and q are any three collinear points, with q between p and r, then

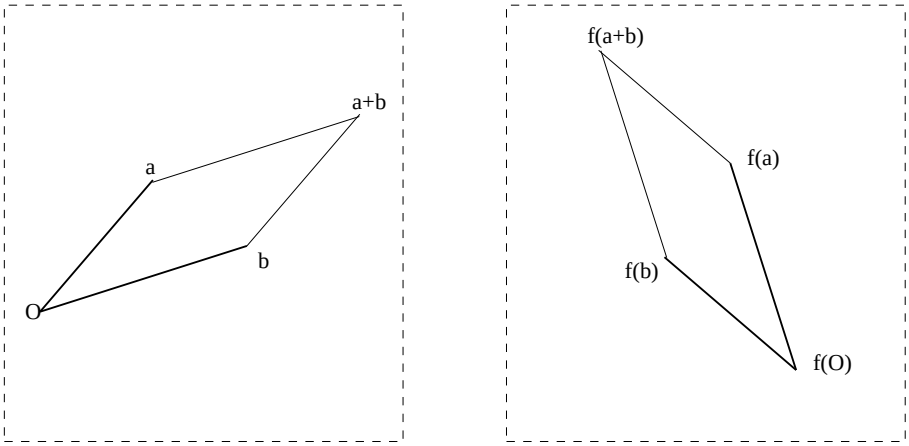
$$d(p, r) = d(p, q) + d(q, r).$$

Hence, $d(f(p), f(r)) = d(f(p), f(q)) + d(f(q), f(r))$. This means that $f(p), f(q)$ and $f(r)$ are collinear - if they were not, then we would have

$$d(f(p), f(r)) < d(f(p), f(q)) + d(f(q), f(r))$$

(think of the triangle with vertices $f(p), f(q)$ and $f(r)$.)

Next we show that for any two points $a, b \in \mathbb{R}^n$, $f(a + b) = f(a) + f(b)$. This is now easy: remember that addition of vectors obeys the parallelogram law.



Since parallelism is defined in terms of distance, f maps parallel lines to parallel lines, and so maps parallelograms to parallelograms. Thus, (see diagram) the quadrilateral with vertices $0 = f(0), f(a), f(b)$ and $f(a+b)$, being the image of a parallologram, is itself a parallelogram. It follows, by the paralleogram law, that the vector $f(a+b)$ is the sum of the vectors $f(a)$ and $f(b)$.

To see that $f(\lambda v) = \lambda f(v)$, note that (unless $v = 0$) the set $L := \{\lambda v : \lambda \in \mathbb{R}\}$ is a straight line, so its image $f(L) = \{f(\lambda v) : \lambda \in \mathbb{R}\}$ is a straight line also. We have

$$d(0, f(\lambda v)) = d(0, \lambda v) = |\lambda| \cdot d(0, v) = |\lambda| \cdot d(0, f(v)).$$

As $f(L)$ is a straight line, it contains only two points whose distance from 0 is equal to $|\lambda| \cdot d(0, f(v))$, namely $\lambda f(v)$ and $-\lambda f(v)$. So $f(\lambda v)$ must be one of these two. To rule out the possibility that $f(\lambda v) = -\lambda f(v)$, we compare $d(v, \lambda v)$ and $d(f(v), -\lambda f(v))$. The former is equal to $|\lambda - 1| d(0, v)$, and the latter is equal to $|\lambda + 1| \cdot d(0, f(v)) = |\lambda + 1| \cdot d(0, v)$. Unless $v = 0$, these two cannot be equal, and so we cannot have $f(\lambda v) = -\lambda f(v)$. The only possibility left is that $f(\lambda v) = \lambda f(v)$.

If $v = 0$ then $-\lambda f(v)$ and $\lambda f(v)$ are equal and in any case $f(\lambda v) = \lambda f(v)$. Our proof is complete. □

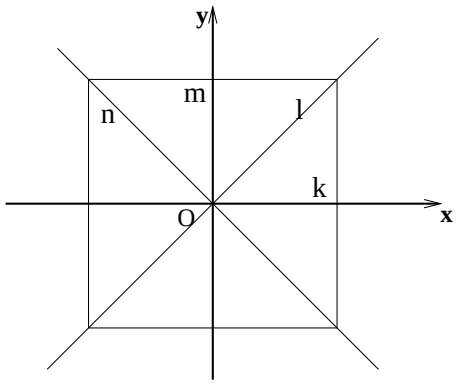
Corollary 2.9 *Any isometry of $\mathbb{R}^n$ is the composition of a linear map and a translation.*

Proof Let f be an isometry, and let $v = f(0) - 0$. Let T_v denote translation by v: $T(u) = u + v$, for every u. Since a translation is clearly an isometry, the map $T_{-v} \circ f$ is an isometry. Now $T_{-v} \circ f(0) = T_{-v}(f(0)) = T_{-v}(v) = 0$; since $T_{-v} \circ f$ fixes 0, it is therefore linear, by 2.8. Now f is the composite of T_v and the linear map $T_{-v} \circ f$. □

From this corollary lots of other things follow; for example, that an isometry of $\mathbb{R}^n$ sends planes to planes. Moreover, one can learn a lot about isometries by using linear algebra and matrices.

The group $I(X)$ is usually called the group of *symmetries* of X; its size is a measure of the symmetry of X, but there is much more to it than that - its algebraic structure reflects the geometry of X in subtle and intricate ways. We shall explore this with some examples. Before doing so, let me just point out a connection between the symmetry groups of subsets of $\mathbb{R}^n$, and groups of $n \times n$ matrices. If $X \subset \mathbb{R}^n$ is a bounded subset (i.e. contained in some ball of finite radius), then there is a point in $\mathbb{R}^n$ (the "centre of mass" of X) which is left fixed by every isometry of X. If we move X so that this point becomes the origin 0, then every isometry of X fixes 0, and is thus a linear map, by the corollary to Proposition 2.3. It can thus be represented by an $n \times n$ matrix. Composition of linear maps (and in particular of isometries fixing 0) corresponds to multiplication of matrices, and so there is a bijection between $I(X)$ and a subgroup of the group of $n \times n$ invertible matrices, which, in a sense to be made precise later, respects the algebraic structure.

Example 2.10 Let S be a square in the plane. Choose coordinates so that the origin is at its centre. Every isometry of the plane fixing 0 is then a linear map, by Corollary 2.9. All isometries of the square fix its centre, and thus are linear. We will write down the matrices of the 8 isometries in $I(\mathrm{Sq})$. To make this as easy as possible, situate the coordinate axes along the lines k and m.



Recall that if $L : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear map, then the matrix of L (with respect to the usual basis $(1,0),(0,1)$), which we will denote by $[L]$, is defined to be

$$\begin{pmatrix} [L(1,0)] & [L(0,1)] \end{pmatrix}$$

where $[L(1,0)]$ and $[L(0,1)]$ are the images of $(1,0)$ and $(0,1)$, written as column vectors.

For example, we take $L = R$. Clearly, we have $R(1,0) = (0,1)$, $R(0,1) = (-1,0)$, and thus

$$[R] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Similarly, $r_\ell(1,0) = (0,1), r_\ell(0,1) = (1,0)$, so

$$[r_\ell] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Exercise Complete the list of matrices representing elements of $I(\mathrm{Sq})$.

Having written the elements of $I(\mathrm{Sq})$ in the form of matrices, we have another means of calculating the composite of two elements of $I(\mathrm{Sq})$; for matrix multiplication is defined in the way it is precisely in order that the matrix of the composite of linear maps should be the product of the two matrices: symbolically,

$$[L_1 \circ L_2] = [L_1] \times [L_2].$$

Thus, for example

$$[R \circ r_\ell] = [R] \times [r_\ell] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

which we recognise to be the matrix of r_m.

Product groups

If G_1 and G_2 are groups, there is a straightforward way of making the Cartesian product $G_1 \times G_2$ into a group: we define the binary operation by

$$(g_1, g_2)(h_1, h_2) = (g_1 h_1, g_2 h_2).$$

It is easy to verify that all the group axioms are satisfied. For example, here is the group table for $\mathbb{Z}_2 \times \mathbb{Z}_2$.

	0	1
0	0	1
1	1	0

	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	(0,0)	(0,1)	(1,0)	(1,1)
(0,1)	(0,1)	(0,0)	(1,1)	(1,0)
(1,0)	(1,0)	(1,1)	(0,0)	(0,1)
(1,1)	(1,1)	(1,0)	(0,1)	(0,0)

Multiplication tables for $\mathbb{Z}_2$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$

The group $G_1 \times G_2$ is called the *product* of G_1 and G_2, or sometimes the *direct product* of G_1 and G_2, or even the *exterior direct product* of G_1 and G_2.

There is no reason to stop at the product of just two groups; one can define the product of any number of groups, by essentially the same procedure.

Terminology and Notation

Definition 2.11 An **abelian** group is one in which $ab = ba$ for every pair of elements a and b.

When I wrote down the axioms for groups I used multiplicative notation: the binary operation is denoted by simple juxtaposition, as in the expressions "ab" or "g_1g_2", and the neutral element is denoted e or sometimes 1. Later on, when have occasion to talk about two different groups G_1 and G_2 in the same sentence, we will use e_1 to denote the neutral element of G_1, and e_2 for that of G_2.

The inverse of an element g is denoted g^{-1}. However, when talking about specific examples, we may use other notation;

- in the case of the (additive) groups $\mathbb{Z}, \mathbb{Z}_n, \mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$ we use the additive notation, $a + b$, $n + m$, etc., with the (additive) inverse of a denoted $-a$ and the neutral element, of course, denoted by 0. In general if G is an abelian group, we denote the neutral element by 0; of course, this does not mean that it is the same as the number 0 in $\mathbb{Z}$.

- when discussing the symmetry group of a geometric figure X, whose elements are mappings of X to itself, we often use the "composition" symbol (a little circle): the composition of maps f and g is $f \circ g$. The neutral element is denoted id or sometimes id_X. The multiplicative notation is also often used in this context. In the group S_n of permutations of a set of n elements, the multiplicative notation is commonest.

- For some reason, the additive notation is almost always used when we discuss abelian groups.

When using the multiplicative notation, we write g^n to mean $g \times g \times \cdots \times g$ (n times); in the additive notation $g + g + \cdots + g$ (n times) is written ng. Caution! This does not mean that we are thinking of n as an element of the group! We are simply writing ng as an abbreviation for $g + \cdots + g$.

To revert to multiplicative notation: it's also worth noting that $(ab)^n$ is not usually equal to a^nb^n. For $(ab)^n = abab\cdots ab$ whereas $a^nb^n = aa\cdots abb\cdots b$; the two are only equal if we can interchange the order of the a's and b's, i.e. if $ab = ba$.

A word of clarification concerning $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$: here there are two operations, $+$ and $\times$, so one might expect some confusion unless the operation is specified. However, the element 0 never has a multiplicative inverse, so none of $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ is a group under multiplication. If we remove 0, we do get a group under multiplication; the resulting groups are referred to as $\mathbb{Q}^\times, \mathbb{R}^\times$ and $\mathbb{C}^\times$. In consequence, $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ will *always* denote the additive groups.

Appendix

In this appendix we prove some of the geometrical facts about isometries that we have used. The proofs use Euclidean geometry, although at the end we give an alternative proof, using linear algebra, for the "three point rule" Corollary 2.6.

Proposition 2.5 Let p, q and r be three non-collinear points in $\mathbb{R}^2$. Then any point x is uniquely determined by its distance from p, from q and from r.

Proof Suppose $d(x,p) = d(y,p)$ and $d(x,q) = d(y,q)$. Then the triangles pxq and pyq are congruent (SSS), and so either $x = y$ or x is the reflection of y in the line through p and q. If also $d(x,r) = d(y,r)$, then the triangles pxr and pyr are congruent, and so if $x \neq y$ then x is also the reflection of y in the line through p and r. But then the line through p and r must coincide with the line through p and q, and this means that p, q and r are collinear. This is a contradiction, so we must have $x = y$. □

Corollary 2.6 (i) Any isometry of the plane fixing three non-collinear points is the identity map.

(ii) If the plane isometries f and g agree on three non-collinear points then $f = g$.

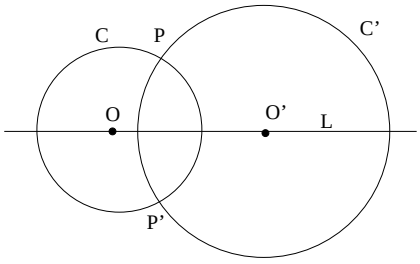
Proof (i) If f fixes p, q and r, then since for any x we have $d(f(x), p) = d(f(x), f(p)) = d(x,p)$, $d(f(x), q) = d(f(x), f(q)) = d(x,q)$ and $d(f(x), r) = d(f(x), f(r)) = d(x,r)$, it follows by the proposition that $f(x) = x$.
(ii) Apply (i) to $f \circ g^{-1}$. □

Corollary 2.7 Any isometry f fixing a point O is either a rotation about O or a reflection in a line through O.

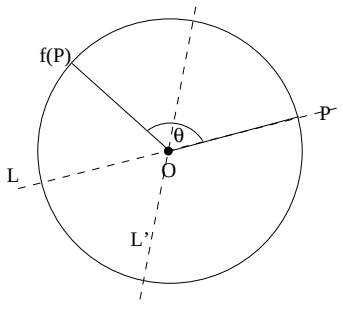
Proof There are two cases: f fixes only O, and f fixes some other point also.

Suppose first that f fixes also some point $O' \neq O$. Let p be some point not on the line L through O and O'. If $f(p) = p$, then f fixes three non-collinear points, so f is the identity map, by 2.6. So suppose instead that $f(p) \neq p$. Let C and C' be the circles centred at O and O' and passing through p. They meet at p and at the point $p' = r_L(p)$. As C and C' are mapped to themselves by f, $f(p) \in C \cap C'$. As $f(p) \neq p$, $f(p) = p'$, i.e. $f(p) = r_L(p)$.

We now have $f(O) = r_L(O)$ and $f(O') = r_L(O')$ (since both isometries fix O and O') and $f(p) = r_L(p)$. It follows by 2.6 that $f = r_L$.



The case where f fixes O and O'



The case where f fixes only O

Now we deal with the case where f fixes only O. Let $P \neq O$. Let θ be the angle $P\hat{O}f(p)$. Then $f(p) = R_\theta(p)$. Hence $R_{-\theta} \circ f(p) = p$. Also $R_{-\theta} \circ f(O) = O$, and so by what we have shown, $R_{-\theta} \circ f$ is either the identity, or a reflection in the line L through O and p. In the second case, $f = R_\theta \circ r_L$; but this fixes points on the line L' bisecting the angle $pOf(p)$, and f fixes only O, by assumption. So the second case can't happen, and thus $R_{-\theta} \circ f$ is the identity and $f = R_\theta$. □

Algebraic proof of Corollary 2.6

If the isometry f fixes three non-collinear points, take one of them to be the origin of coordinates, O. Call the other two points P and Q. Because O, P and Q are not collinear, the vectors $p := OP$ and $q := OQ$ are linearly independent, and therefore a basis for $\mathbb{R}^2$. If v is any other vector in the plane we can write v as a linear combination $\alpha \cdot p + \beta \cdot q$. Then

$$f(v) = f(\alpha \cdot p + \beta \cdot q) = \alpha \cdot f(p) + \beta \cdot f(q),$$

the last equality by linearity of f. As f fixes all O, P and Q, $f(p) = p$ and $f(q) = q$, so

$$f(v) = \alpha \cdot p + \beta \cdot q = v,$$

and f is the identity map. □

3 Finding Subgroups

In this section we look at three different ways of obtaining subgroups of a group. Later, when we introduce the notion of *homomorphism*, we will find another, in some ways the most important of all.

I. Symmetry Breaking

One way of manufacturing subgroups of $I(Sq)$ is to look at ways of *breaking the symmetry of the square.* For example, a square is a special case of a rectangle, and so among its symmetries must be those of a rectangle. Since the symmetries of a rectangle form a group under composition, they must form a subgroup of $I(\mathrm{Sq})$. One way of isolating this subgroup is to remove the other symmetries ("symmetry breaking"). We can think of the square as the "limiting position" of the rectangle S_t with base of length $1 + t$ and height $1 - t$, as $t \to 0$. As soon as $t \neq 0$, the rectangle becomes less symmetrical, and this is reflected in the fact that some of the symmetries (elements of $I(Sq)$) are lost. In fact, we lose $R_{\pi/2}, R_{3\pi/2}, r_L$ and r_N, since these exchange the vertical and horizontal, and the rectangle S_t has vertical and horizontal sides of different length. An easy calculation now shows that for $t \neq 0$,

$$I(S_t) = \{I, R_\pi, r_K, r_M\}.$$

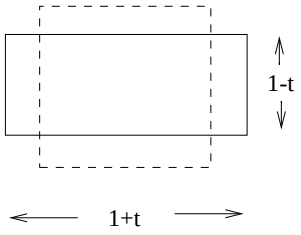


Figure 3:*Breaking the symmetry of the square*

Note that if we were careless and tilted the rectangle while we were doing this, or moved its centre, then it wouldn't be symmetrical in the very same lines K and M as the square. We have to be a little careful to make sure that the symmetry group of our deformed square really is contained in $I(Sq)$.

However, the diagram below (once you complete it) shows that *every* subgroup of $I(Sq)$ can be obtained in this way.

Of the three methods for finding subgroups, this is perhaps the least important from a purely algebraic point of view; however, it provides us with an interesting and useful collection of examples.

Exercise: Is there a quadrilateral with the same symmetry group as the figure in the middle of the second row in the diagram below?

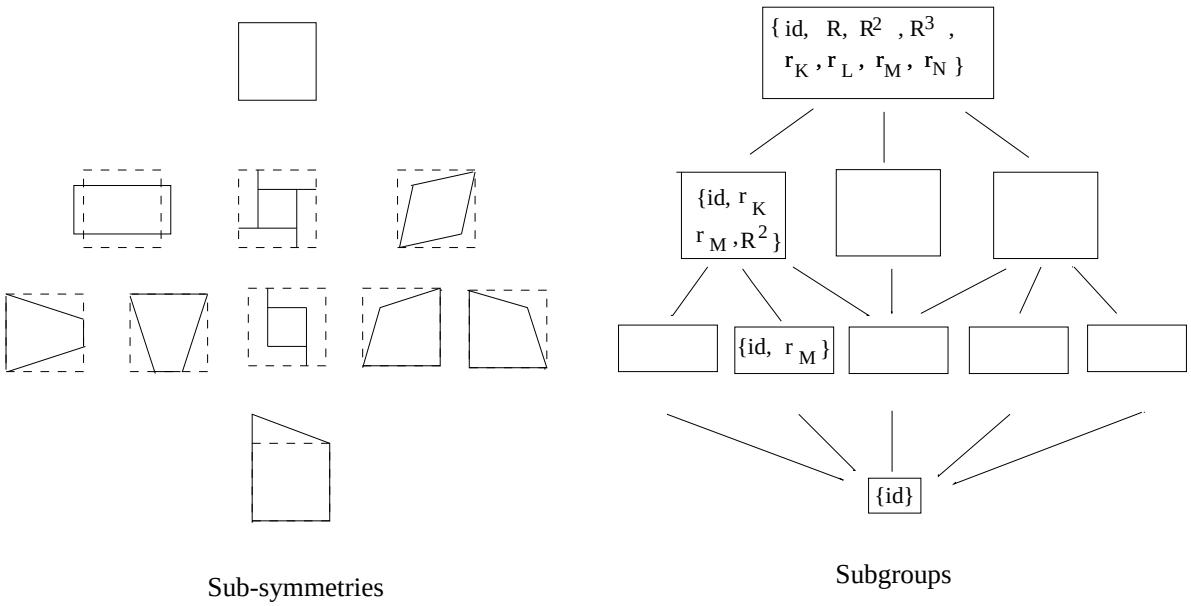


Figure 4: The hierarchy of symmetries of the square: fill in the empty boxes on the right with the lists of symmetries of the corresponding figure on the left.

II.

The second method for manufacturing subgroups is by considering subsets consisting of el-

ements with some property in common. In order that we get a subgroup, the property in question should be preserved by composition and by inversion.

Example 3.1 1. $Sl(n,\mathbb{R})$ is the set of $n \times n$ matrices with determinant 1. It's evidently contained in the group $Gl(n,\mathbb{R})$ of all invertible $n \times n$ matrices with real entries — if $\det(A)$ is non-zero then A is invertible. Moreover, the fact that $\det(AB) = \det(A)\det(B)$ shows that if A and B each have determinant 1 then so does AB - i.e. if $A, B \in Sl(n,\mathbb{R})$ then $AB \in Sl(n,\mathbb{R})$ - and that if $\det(A) = 1$ then $\det(A^{-1}) = (det(A))^{-1} = 1$.

2. On the other hand, the set of matrices with determinant a non-zero integer is *not* a subgroup of $Gl(n,\mathbb{R})$. For although it is closed under multiplication, the inverse of a matrix with determinant n has determinant $1/n$, not an integer in general.

3. In $Gl(n,\mathbb{R})$, consider the set $O(n)$ of all matrices A such that $AA^t = I$ (where A^t denotes the transpose of A and I is the identity matrix).
Exercise $O(n)$ is a subgroup of $Gl(n,\mathbb{R})$. (You need to know here that the transpose of AB is B^tA^t.)
$O(n)$ is called the *orthogonal group.*
Exercise(i) Let A be an $n \times n$ matrix with real entries and let $a_1, \ldots, a_n$ be its columns, thought of as vectors in $\mathbb{R}^n$. Show that $A \in O(n)$ if and only if

$$a_i \cdot a_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j, \end{cases}$$

i.e. if and only if the columns of A form an orthonormal basis for $\mathbb{R}^n$.
(ii) Show that if $A \in O(n)$ then for any vectors $u, v \in \mathbb{R}^n$, we have

$$Au \cdot Av = u \cdot v.$$

Hint: if we write vectors as columns rather than rows, then the scalar product of vectors $x, y \in \mathbb{R}^n$ can be written in terms of matrix multiplication as $x^t y$, where x^t is the transpose of x.

4. In $Gl(n,\mathbb{C})$, consider the set $U(n)$ of matrices A such that $AA^* = I$. Here A^* is the *conjugate transpose* of A — the matrix you get by transposing A and taking the complex conjugate of each of its entries.
Exercise $U(n)$ is a subgroup of $Gl(n,\mathbb{C})$.
$U(n)$ is called the *unitary group.* The groups $O(n)$ and $U(n)$ are important in physics as well as in many branches of mathematics.

5. The set of $n \times n$ matrices A such that multiplication by A induces an isometry of $\mathbb{R}^n$ is a subgroup of $Gl(n,\mathbb{R})$.
Exercise Show that this subgroup is in fact $O(n)$.
Hint: in view of the exercise concerning $O(n)$ above, it is necessary only to show that multiplication by A induces an isometry of $\mathbb{R}^n$ if and only if for any two vectors $u, v \in \mathbb{R}^n$, $Au \cdot Av = u \cdot v$.

6. If G is any group, its *centre*, $Z(G)$, is the set of all elements which commute with every element in the group:

$$Z(G) = \{g \in G : gh = hg \text{ for all } h \in G\}.$$

(Of course, if G is abelian, $Z(G)$ is all of G.
Exercise: Show that $Z(G)$ is a subgroup of G.

- $Z(S_3) = \{\text{id}\}$.
- **Exercise:** Find $Z(I(Sq))$.
- In the group $Gl(n, \mathbb{R})$ of invertible $n \times n$ matrices with real entries, the centre is the set of "scalar matrices" - matrices which are non-zero scalar multiples of the identity matrix. It's easy to see that scalar matrices do commute with every matrix, and so lie in $Z(Gl(n, \mathbb{R}))$. To see that *only* the scalar matrices are in $Z(Gl(n, \mathbb{R}))$ requires more thought, and is left as a guided exercise (see Exercises I).

7. The set of rational numbers with odd denominator is an additive subgroup of $\mathbb{Q}$. In fact, for any set $\mathcal{S}$ of primes, the set of all rational numbers whose denominators have all their prime factors in $\mathcal{S}$ is again a subgroup of $\mathbb{Q}$ (**Exercise**).

   Is the set of rationals with even denominator a subgroup?

8. Let A be one of the vertices of the square. The set of symmetries of the square fixing A is a subgroup of $I(Sq)$.

9. If $f \in I(Sq)$, then either it interchanges the diagonals, or it leaves them where they are. The set of f which leave them where they are is a subgroup.

10. More generally, if $Y \subset X$, then $I(X;Y) = \{f \in I(X) : f(Y) = Y\}$ and

$$I_Y(X) = \{f \in I(X) : f(y) = y \text{ for all } y \in Y\}$$

    are subgroups of $I(X)$.

11. The subset $\{0, 2, 4\} \subset \mathbb{Z}_6$ is a subgroup; $\{0, 3\} \subset \mathbb{Z}_6$ is a subgroup.

    Exercise: How many subgroups of $\mathbb{Z}_{12}$ can you find? Perhaps the next method for finding subgroups will help here.

III

The third way of getting subgroups is by specifying a *subset* S of G and then taking the set $\langle S \rangle$ of all elements of G that you can get by multiplying together members of this subset and their inverses: if $S = \{s_1, s_2, \cdots\}$ then $\langle S \rangle$ contains the elements $s_1, s_1^{-1}, s_2, s_2^{-1}, s_1s_2, s_2s_1$, $s_1^{-1}s_2, s_1s_2s_3,\ s_1^{-1}s_3s_2s_2, \cdots$ and so on. It looks as though there are infinitely many elements in this list, but there will in general be lots of collapsing; consider for example in $I(Sq)$,

$$Rr_K = r_K RRR, \quad r_K r_K = r_L r_L = r_M r_M = r_N r_N = \text{id},$$

and

$$r_K r_L = r_L r_M = r_M r_N = r_N r_K = R.$$

etc.. There has to be collapsing, of course - $\langle S \rangle$ is contained in G, and so can have no more elements than G does.

It's fairly obvious that $\langle S \rangle$ is closed under multiplication: if $s_{i_1}^{\pm 1} \cdots s_{i_r}^{\pm 1}$ and $s_{j_1}^{\pm 1} \ldots s_{j_t}^{\pm 1}$ are two elements of $\langle S \rangle$ then so is their product $s_{i_1}^{\pm 1} \ldots s_{i_r}^{\pm 1} s_{j_1}^{\pm 1} \ldots s_{j_t}^{\pm 1}$.

Similarly, $\langle S \rangle$ contains the inverse of each of its members:

$$(s_1^{\pm 1} s_2^{\pm 1} \cdots s_n^{\pm 1})^{-1} = s_n^{\mp 1} s_{n-1}^{\mp 1} \cdots s_1^{\mp 1}.$$

So $\langle S \rangle$ is a subgroup of G. It is referred to as the subgroup *generated by* S, and the members of S are its *generators*.

In the special case where $\langle S \rangle$ is all of G, we say that the elements of S *generate* G.

Example 3.2 1. The element $1 \in \mathbb{Z}$ generates the whole group: $\langle 1 \rangle = \mathbb{Z}$. On the other hand $\langle 2 \rangle$ is a proper subgroup of $\mathbb{Z}$, consisting of the even numbers.

2. Still in $\mathbb{Z}$, we have $\langle 2, 3 \rangle = \mathbb{Z}$. This example is interesting, because neither 2 nor 3 will generate $\mathbb{Z}$ on its own. Thus $2, 3$ is a *minimal* set of generators - as soon as you take away any of its members, it stops generating all of $\mathbb{Z}$. So, two distinct minimal sets of generators of a group (in this example, $\{1\}$ and $\{2, 3\}$ in $\mathbb{Z}$) may have different numbers of elements. This contrasts with the situation in linear algebra, where two minimal sets of generators of the same vector space always have the same number of elements (and this number is of course the dimension of the space). Although a vector space is a group under addition, we do not have a contradiction here, since in linear algebra we use the term "generator" in a somewhat different way from in group theory. In linear algebra, you are allowed to multiply the generators by scalars from the field, rather than just adding or subtracting them. In fact, $\mathbb{R}$ and $\mathbb{Q}$ are not finitely generated (as groups) at all.

3. Those of you who took "Introduction to Geometry" proved that $I(\mathbb{R}^2)$ is generated by reflections: every isometry of $\mathbb{R}^2$ can be expressed as the composite of up to three reflections.

4. In $\mathbb{Z}_3$ we have $2 + 2 = 1$, so $0, 1$ and 2 are all in $\langle 2 \rangle$, i.e. $\langle 2 \rangle = \mathbb{Z}_3$.

5. $I(Sq) = \langle r_K, R \rangle$. One can see this by noting that since R^2 and R^3 are obviously in $\langle r_K, R \rangle$, we only have to find r_L, r_M and r_N. Now one checks that $R^3 r_K R = r_{R(K)} = r_L$ (so that $r_L \in \langle r_K, R \rangle$) and then, proceeding cumulatively, $R^3 r_L R = r_M, R^3 r_M R = r_N$, so that r_M and r_N also lie in $\langle r_K, R \rangle$. The same argument shows that in fact R together with any one of the reflections r_K, r_L, r_M or r_N will generate $I(Sq)$.

6. $I(Sq) = \langle r_K, r_L \rangle$. For $r_L r_K = R$, so $R \in \langle r_K, r_L \rangle$, and therefore $I(Sq) = \langle r_K, R \rangle \subseteq \langle r_K, r_L \rangle$. In fact the only property of K, L that we're using here is that they are separated by an angle of only $\pi/4$; L and M or M and N would have done just as well. For the composite of two reflections (in lines passing through O) is a rotation through twice the angle from the first to the second (**Exercise**), and thus is L_1, L_2 are any two lines separated by an angle of $\pi/4$, then $R \in \langle r_{L_1}, r_{L_2} \rangle$.

7. On the other hand, $\langle r_K, r_M \rangle$ does not contain R, so is a proper subgroup of $I(Sq)$.

8. If G is a group, any element of the form $a^{-1}b^{-1}ab$ is called a *commutator*. It is equal to e if and only if $ab = ba$ (i.e. if a and b commute). The *commutator subgroup*, or *derived subgroup*, of G is the subgroup generated by all commutators $a^{-1}b^{-1}ab$. It is usually denoted $[G, G]$.

   It is hard to characterise its elements as "the set of all $g \in G$ such that ...", unless the property is "that can be written as a product of commutators".

   Exercise (i) $[G, G] = \{1\}$ if and only if G is abelian.

   (ii) Find the commutator subgroup of $I(Sq)$. (This seems just to be a question of calculation. Later we'll see a way of showing (easily) that the commutator subgroup of $I(Sq)$ is not *all* of $I(Sq)$.)

Definition 3.3 *A group G is finitely generated if there is a finite subset $S \subset G$ such that $\langle S \rangle = G$.*

Proposition 3.4 *(i) $\mathbb{R}$ is not a finitely generated group. (ii) $\mathbb{Q}$ is not a finitely generated group.*

Proof (i) $\mathbb{R}$ is uncountable, but if $S \subset \mathbb{R}$ is finite with, say, r elements, then $\langle S \rangle$ is at most countably infinite. For since $\mathbb{R}$ is abelian, every element in $\langle S \rangle$ can be written in the form

$$n_1 s_1 + n_2 s_2 + \cdots n_r s_r,$$

where the n_i are integers (and of course $n_i s_i$ means $s_i + s_i + \cdots s_i$ (n times)). It is not hard to show that only countably many elements can be obtained in this way (**Exercise**- with hints in Exercises I).

(ii) Suppose that $S = \{s_1, \cdots, s_n\}$ is a finite subset of $\mathbb{Q}$. The denominators of all of the s_i have only a finite number of distinct prime factors; if $q \in \langle S \rangle$ then *its* denominator has only these, and no other, prime factors. Since the number of primes is infinite (- do you recall a proof of this?), and any one can appear as the denominator of a rational number, this shows that $\langle S \rangle$ cannot be all of $\mathbb{Q}$. □

An alternative way of thinking of the subgroup of a group G generated by a subset S is as follows: the set of *all* subgroups of G that contain S is not empty, since G itself is one. The intersection of any collection of subgroups of G is itself a subgroup (**Exercise**); thus in particular the intersection of *all subgroups of G containing S* is a subgroup of G, and evidently contains S. Since it is contained in every subgroup of G which contains S, it is *the smallest subgroup of G containing S*. We denote it for the moment by $\langle\langle S \rangle\rangle$. Fortunately, we have

Proposition 3.5 $\langle\langle S \rangle\rangle = \langle S \rangle$.

Proof Any subgroup containing S must in particular be closed under multiplication and inversion. Thus it must contain every element obtained by multiplying together elements of S and their inverses. That is, every subgroup containing S must contain $\langle S \rangle$. It follows that $\langle\langle S \rangle\rangle$ contains $\langle S \rangle$.

Conversely, $\langle S\rangle$ is an example of a subgroup of G containing S; hence it contains the smallest, $\langle\langle S\rangle\rangle$. □

These two ways of thinking of the same thing, $\langle S\rangle$ and $\langle\langle S\rangle\rangle$, are both useful. The first works from the inside, building up what has to be in the subgroup, and the second works from the outside, stripping away what is unnecessary and keeping only what has to be there.

Subgroups of a product

Let G_1 and G_2 be groups, and denote their neutral elements by e_1 and e_2 respectively. The product group $G_1 \times G_2$ has a subgroup $G_1 \times \{e_2\}$ which is pretty much the same as G_1: $G_1 \times \{e_2\} = \{(g_1, g_2) \in G_1 \times G_2 : g_2 = e_2\}$. There is also a subgroup $\{e_1\} \times G_2$ defined analogously. In fact whenever H_1 is a subgroup of G_1 and H_2 a subgroup of G_2 then $H_1 \times H_2$ is a subgroup of $G_1 \times G_2$. On the other hand, not every subgroup of $G_1 \times G_2$ is of this form. For example, in $\mathbb{R} \times \mathbb{R}$, for each non-zero vector (a, b) the line $\{\lambda(a, b) : \lambda \in \mathbb{R}\}$ is an additive subgroup, but is not equal to the Cartesian product of a subgroup of the first copy of $\mathbb{R}$ and a subgroup of the second copy of $\mathbb{R}$, unless the line is parallel to one of the co-ordinate axes.
Exercise Find all subgroups of $\mathbb{Z}_2 \times \mathbb{Z}_2$.

4 Cyclic Groups

Definition 4.1 *(i) A group G is* cyclic *if there is an element $g \in G$ such that $G = \langle g\rangle$.*

(ii) The order *$o(g)$ of an element $g \in G$ is the least positive integer n such that $g^n = e$; if g^n is never e for positive n (e.g. $g \neq 0$ in $\mathbb{Z}$ - of course here one would use the additive notation and write $ng \neq 0$ rather that $g^n \neq e$) then $o(g)$ is defined to be ∞.*

Exercise Show that if $g^n = e$ then n is a multiple of $o(g)$.

Let G be a group, let $g \in G$ and suppose that $o(g) = n < \infty$. Then the subgroup $\langle g\rangle$ is equal to $\{e, g, g^2, \cdots, g^{n-1}\}$, and so has $o(g)$ elements. (Thus, if G is a finite group, the element g generates the whole group G if and only if $o(g) = |G|$.) Now recall *Lagrange's Theorem:*

Theorem 4.2 *Let H be a subgroup of the finite group G. Then $|H|$ divides $|G|$.* □

We will go over the proof later in the course, to develop further some of the ideas it involves.

Corollary 4.3 *If g is any element of the finite group G, then $o(g)$ divides $|G|$.*

Proof $o(g) = |\langle g\rangle|$ must divide $|G|$, by Lagrange's Theorem. □

Proposition 4.4 *Every cyclic group is abelian.*

Proof $g^n g^m = g^{n+m} = g^m g^n$. □

Example 4.5 Examples of Cyclic Groups

1. The rotation subgroup $\{\mathrm{id}, R, R^2, R^3\}$ of $I(Sq)$ is generated by R, but $I(Sq)$ itself is not cyclic. Every non-trivial element has order either 2 (all the reflections, and R^2) or 4 (R and R^3), and thus there is no element of order 8.

2. The group $\mathbb{Z}$ is generated by 1, and also by -1: $\mathbb{Z} = \langle 1 \rangle = \langle -1 \rangle$.

3. The group $\mathbb{Z}_n$ of integers modulo n (under addition) is generated by 1: $\mathbb{Z}_n = \langle 1 \rangle$.

4. Let G_n be the multiplicative group of *n-th roots of unity*,

$$G_n = \{z \in \mathbb{C} : z^n = 1\} = \{e^{2\pi i q/n} : q = 0, \cdots, n-1\}.$$

   Then $G_n = \langle e^{2\pi i/n} \rangle$ is cyclic.

5. The group $\mathbb{Z}_2 \times \mathbb{Z}_2$ is *not* cyclic. Since it has only four elements, it's very easy to check this! It turns out that every element except the neutral element $(0,0)$ has order 2.

6. In fact, $\mathbb{Z}_n \times \mathbb{Z}_n$ is never cyclic - every element except $(0,0)$ has order n, for $(a,b) + \cdots + (a,b)(n \text{ times})$ is equal to $(a + \cdots + a, b + \cdots + b) = (0,0)$. Whereas the order of the group is n^2.

7. By contrast, $\mathbb{Z}_2 \times \mathbb{Z}_3$ is cyclic! It is easily checked that it is generated by $(1,1)$ (**Exercise**).

8. In any group whatsoever, the subgroup generated by any single element is of course cyclic.

Proposition 4.6 *The element $m \in \mathbb{Z}_n$ generates it if and only if $(m,n) = 1$. (Here (m,n) means the highest common factor of m and n).*

Proof "Only if" is easy: suppose that m and n have common factor $p > 1$. Then $m = pq$ for some q, and so $0 = nq = (n/p)pq = (n/p)m$. That is, the order of m is no bigger that (n/p), and so $\langle m \rangle$ has fewer than n elements.

The converse is a little more complicated, and we use the Euclidean algorithm in the following form: since n and m are coprime, there exist integers a, b such that $am + bn = 1$. This means that $am = 1$ in $\mathbb{Z}_n$, and thus that $1 \in \langle m \rangle$. In case you feel unhappy at the possibility of a being a negative integer (after all, one of a and b has to be negative), note that you can replace a by $a + n$ or $a + 2n$ or a plus any multiple of n, since we're working modulo n.

Anyway, it follows immediately from the fact that $1 \in \langle m \rangle$ that $\langle m \rangle = \mathbb{Z}_n$. □

Note that when talking about $\mathbb{Z}_n$ we use the additive notation. In the next proposition we revert to the multiplicative notation to talk about the "general" cyclic group.

Proposition 4.7 *Let G be a cyclic group of order n, and suppose that q divides n. Then G has a subgroup of order q (i.e. with q elements).*

Proof **Exercise**: suppose that g generates G; you have to choose p appropriately so that g^p has order q. It's not hard! □

Proposition 4.8 *Every group of prime order is cyclic.*

Proof This is a consequence of Lagrange's theorem: if H is a subgroup of G then $|H|$ divides $|G|$. So if $g \in G$, then $o(g) = |\langle g \rangle|$ must divide $|G|$. But $|G|$ is prime, so unless $g = 1$, $o(g)$ must be equal to p, and $\langle g \rangle = G$. □

Observe that this proof shows that if G has prime order and $g \neq 1$, then g generates G.

Proposition 4.9 *Every subgroup of a cyclic group is itself cyclic.*

Proof Let H be a subgroup of the cyclic group G. Suppose g generates G. Then every element of H is equal to some power of g. Let $h_0 = g^n$ be the least positive power of g in H. In other words, n is a positive integer, $g^n = h_0 \in H$, and if $g^m \in H$ with $m > 0$ then $m \geq n$.

I claim that $H = \langle h_0 \rangle$. For let $h = g^m \in H$, with $m > 0$. Dividing m by n, we can write

$$m = qn + r$$

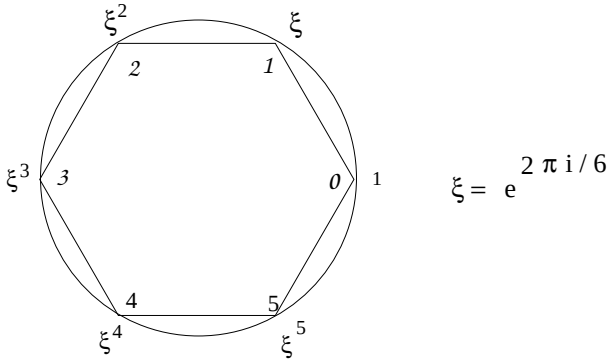
with $0 \leq r < n$. Then $g^r = ((h_0)^{-1})^q h$ is in H. If $r \neq 0$, then we have found a positive power of g lower than n in H. This contradicts the choice of g^n as the *lowest* positive power of g in H. Hence $r = 0$, $m = qn$, and $h = (h_0)^q$.

We can do the same thing with negative powers of g in H, and show that each must be some power of h_0^{-1}.

Thus, $H = \langle h_0 \rangle$ is cyclic. □

5 Isomorphism

You may be feeling that all the examples of cyclic groups that I gave are pretty much the same - each is a thinly disguised version of the other. For example, the group G_6 of n-th roots of unity is really the same as $\mathbb{Z}_6$:



In the diagram the circle is the set S^1 of complex numbers with unit modulus. The number *outside* the circumference next to a marked point is its value as complex number. The number *inside* the circumference next to the same point is the power to which ξ is raised to get this point. Thus, round the outside of the circle we get a list of the elements of G_6, while inside we get a list of the members of $\mathbb{Z}_6$. This is more than just a *bijection* between G_6 and

$\mathbb{Z}_6$: elements of $\mathbb{Z}_6$ and G_6 which correspond under this bijection, behave in the same way algebraically. What I mean is just that the well known "law of indices"

$$\xi^{a+b} = \xi^a \xi^b$$

relates the (additive) structure of $\mathbb{Z}_6$ with the (multiplicative) structure of G_6. Via the correspondence

$$m \mapsto \xi^m,$$

the groups $\mathbb{Z}_6$ and G_6 are "the same".

Definition 5.1 *Let G and H be two groups. A bijection $\phi : G \to H$ is an* isomorphism *if for all g_1, g_2 in G,*

$$\phi(g_1 g_2) = \phi(g_1)\phi(g_2).$$

Note that in this last expression, the product on the right hand side is the product in H, while the product on the left is the product in G. The definition says that the bijection ϕ is an isomorphism if the product in G and the product in H correspond via ϕ. If there is an isomorphism from G to H, we say that G and H are *isomorphic*.

Example 5.2 One familiar and important example is *logarithm*. In the terms we have just introduced, the map

$$\log : \mathbb{R}^\times_{>0} \to \mathbb{R}$$

is an isomorphism from the (multiplicative) group $\mathbb{R}^\times_{>0}$ to the additive group $\mathbb{R}$. So to an algebraist, multiplication in $\mathbb{R}^\times_{>0}$ and addition in $\mathbb{R}$ are the same thing!

Its inverse $\exp : \mathbb{R} \to \mathbb{R}^\times_{>0}$ is also an isomorphism - as is the inverse of any isomorphism (**Exercise**).

The fact that log is an isomorphism from the multiplicative group $\mathbb{R}^\times_{>0}$ to the additive group $\mathbb{R}$ was the basis for an important practical aid to the carrying out of long and difficult multiplications, in the days before electronic calculators were invented. In order to multiply together two or more complicated numbers (e.g. $7,3492.161$ and 1.8555), one looked up their logarithms in a book of tables (every schoolchild had one), added them together (adding is easier than multiplication), and then looked up the antilogarithm (exponential) of the result. In this case we have

$$\log 73,492.161 = 11.2048; \quad \log 1.8555 = 0.6179$$

(to four decimal places); the sum of the logarithms is 11.8227, and the antilogarithm (i.e. the exponential) of this is 135990.0 (to five significant figures). That is, the method gives

$$73,492.161 \times 1.8555 = 135990.0$$

(to five significant figures). Although the logarithms and antilogarithms given in the tables were necessarily approximations, the results were sufficiently accurate to provide the basis for many engineering calculations.

Example 5.3 Some isomorphisms are so natural that one hardly notices them. As you saw in Linear Algebra (Theorem 7.1 in Derek Holt's Lecture Notes), once you choose a basis for an n dimensional vector space V, then each linear map $V \to V$ can be represented by an $n \times n$ matrix, (its matrix with respect to the chosen basis), and every matrix induces a linear map $V \to V$. A choice of basis for V thus determines a bijection

$$m : \mathrm{L}(V) \to \mathrm{M}(n, \mathbb{R}),$$

where $\mathrm{L}(V)$ is the set of all linear maps $V \to V$, and $\mathrm{M}(n, \mathbb{R})$ is the set of all $n \times n$ real matrices, invertible or not. The peculiar rule for multiplying matrices is devised precisely in order that the matrix representing the *composite* of the linear maps S and T is the *product* of the matrix representing S and the matrix representing T. That is,

$$m(S \circ T) = m(S)m(T).$$

It is precisely this that makes matrices useful in studying linear maps. Of course, $(\mathrm{L}(V), \circ)$ and $(\mathrm{M}(n, \mathbb{R}), \times)$ are not groups, since in neither cases does every element have an inverse. The subset of $\mathrm{L}(V)$ consisting of isomorphisms is denoted $\mathrm{Aut}(V)$ (for "automorphisms of V"), and the subset of $\mathrm{M}(n, \mathbb{R})$ consisting of invertible matrices is denoted $\mathrm{Gl}(n, \mathbb{R})$. Both of these are groups, with respect to the operations of composition and matrix multiplication, and the bijection of sets

$$m : \mathrm{L}(V, V) \to \mathrm{M}(n, \mathbb{R})$$

in fact gives an isomorphism of groups

$$(\mathrm{Aut}(V), \circ) \simeq (\mathrm{Gl}(n, \mathbb{R}), \text{matrix multiplication}).$$

Example 5.4 The map $T_a : \mathbb{R} \to \mathbb{R}$ defined by $T_a(x) = x + a$ is a bijection, with inverse T_{-a}. But if $a \neq 0$ it is not an isomorphism of the (additive) group $\mathbb{R}$. For it is *not* true that $T_a(x + y) = T_a(x) + T_a(y)$: the left hand side is equal to $x + y + a$, while the right hand side is equal to $x + y + 2a$. On the other hand, *multiplication* M_a by a (non-zero) number a is an isomorphism of $\mathbb{R}$. However, it is not an isomorphism of the multiplicative group $\mathbb{R}^\times$, although it is a bijection.

Proposition 5.5 *Let $\phi : G_1 \to G_2$ be an isomorphism of groups. Then*

1. *$\phi(e_1) = e_2$;*
2. *for all $g \in G_1$, $\phi(g^{-1}) = (\phi(g))^{-1}$*
3. *for all $g_1, \cdots, g_r \in G_1$, $\phi(g_1 g_2 \cdots g_r) = \phi(g_1)\phi(g_2)\cdots\phi(g_r)$;*
4. *for all $g \in G_1$, $o(\phi(g)) = o(g)$.*

Proof (1): $\phi(e_1 e_1) = \phi(e_1)\phi(e_1)$ by definition of isomorphism. The left hand side is just $\phi(e_1)$, and hence we have $\phi(e_1) = \phi(e_1)\phi(e_1)$. As G_2 is a group, we can multiply both sides of this equation by $(\phi(e_1))^{-1}$ (which must exist, even if we don't know what it is yet). In other words, we can cancel $\phi(e_1)$ on each side. We get $e_2 = \phi(e_1)$, as required.

(2): $\phi(g)\phi(g^{-1}) = \phi(gg^{-1}) = \phi(e_1)$ by the defining property of isomorphism; by (1), $\phi(e_1) = e_2$. Thus, $\phi(g)\phi(g^{-1}) = e_2$, and this shows that $\phi(g^{-1})$ is the inverse of $\phi(g)$, i.e. that $\phi(g^{-1}) = (\phi(g))^{-1}$.

(3):

$$\phi(g_1 g_2 \cdots g_r) = \phi((g_1 \cdots g_{r-1})g_r) = \phi(g_1 \cdots g_{r-1})\phi(g_r)$$

(applying the defintion of isomorphism to the two elements $g_1 \cdots, g_{r-1}$ and g_r). We can assume by induction on r that

$$\phi(g_1 \cdots g_{r-1}) = \phi(g_1) \cdots \phi(g_{r-1});$$

thus

$$\phi(g_1 \cdots g_r) = \phi(g_1) \cdots \phi(g_r).$$

(4): $(\phi(g))^{o(g)} = \phi(g^{o(g)}) = \phi(e_G) = e_H$ (the first equality by (3)); hence $o(\phi(g))$ divides $o(g)$ (why?). If $o(\phi(g)) < o(g)$, then we have $g^{o(\phi(g))} \neq e_G$ but $\phi(g^{o(\phi(g))}) = \phi(g)^{o(\phi(g))} = e_H$. As $\phi(e_G)$ is also equal to e_H, this contradicts the fact that ϕ is 1-1. Hence, $o(\phi(g))$ is not less that $o(g)$. Since it divides $o(g)$, the two must be equal. □

I reiterate that when two group are isomorphic, they are the same in *every* algebraic respect. An isomorphism $\phi : G \to H$ maps *every* algebraic structure in G to an identical structure in H. In Proposition 5.5 we have listed some of these, but it is important to appreciate that it is true for every algebraic stucture. After all, a group is nothing but a set together with a binary operation; if $\phi : G \to H$ is an isomorphism, then

- the fact that ϕ is a bijection means that as sets G and H are the same (as far as we are concerned), and
- the fact that $\phi(g_1 g_2) = \phi(g_1)\phi(g_2)$ means the binary operation in H is the same as the binary operation in G.

Proposition 5.6 *The composition of two isomorphisms is an isomorphism.* □

Example 5.7

1. The additive group $\mathbb{Z}_2 = \{0, 1\}$ is isomorphic to the multiplicative subgroup $\{1, -1\}$ of $\mathbb{Q}^\times$. Note that the isomorphism sends the neutral element 0 of the additive group $\{0, 1\}$ to the neutral element 1 of the multiplicative group $\{1, -1\}$. Check the details yourself. From now on we refer to the multiplicative group as $\mathbb{Z}_2$ also. The two sets, with their binary operations, are both incarnations of the same abstract structure.

2. Complex conjugation defines an isomorphism $\mathbb{C} \to \mathbb{C}$ (for it's evidently a bijection, and $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$) and an isomorphism $\mathbb{C}^\times \to \mathbb{C}^\times$, since (**Exercise**) $\overline{z_1 z_2} = \overline{z}_1 \overline{z_2}$. Apart from the identity, this is the only bijection $\mathbb{C} \to \mathbb{C}$ giving rise to isomorphisms $\mathbb{C} \to \mathbb{C}$ and $\mathbb{C}^\times \to \mathbb{C}^\times$.

3. We've already commented on the isomorphism $\mathbb{Z}_n \to G_n$ sending m to ξ^m. In fact, the suspicion that all our examples of cyclic groups of a given order are really thinly disguised versions of one another, is correct: any two are indeed isomorphic. For suppose G is generated by g and H is generated by h, and that $o(g) = o(h) = n$. We define a map $\phi : G \to H$ by $\phi(g^j) = h^j$, for $j = 0, \cdots, n-1$. Note that every element of G is of the

form g^j for some unique j between 0 and $n-1$, so our map really is defined on all of G; as each element of H can be written uniquely as h^j for some j between 0 and $n-1$, ϕ is a bijection. Also,

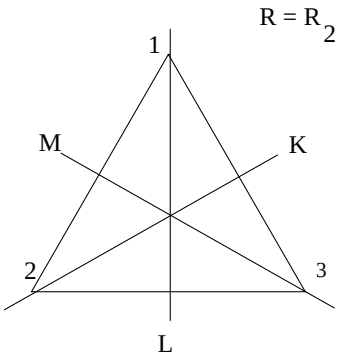
$$\phi(g^j g^k) = \phi(g^{j+k}) = h^{j+k} = h^j h^k = \phi(g^j)\phi(g^k),$$

and thus ϕ is an isomorphism.

Essentially contained in our description of the isomorphism $\phi : G \to H$ is an isomorphism $\mathbb{Z}_n \to G$, simply sending $j \in \mathbb{Z}_n$ to $g^j \in G$. So in some sense $\mathbb{Z}_n$ is "the only" cyclic group of order n. To an algebraist, isomorphic groups really are "the same group" - the pure algebraist strives to uncover the abstract structure, rather than its incarnation as a group of symmetries of a geometric figure, or a group of matrices, or whatever.

4. Let T be an equilateral triangle, and let $I(T)$ be its symmetry group. We define a map ϕ from $I(T)$ to the permutation group S_3 as follows: each map $f \in I(T)$ maps vertices to vertices, and thus induces a permutation of the vertices. The group of permutations of the vertices can be identified with the group S_3 of permutations of $\{1,2,3\}$ by numbering the vertices in some way; then to each element $f \in I(T)$ we can associate a permutation $\phi(f) \in S_3$.

   We've already seen that f is determined by the permutation it induces (cf. 2.6), and so our map is 1-1. It follows that it is surjective also, since $|S_3| = |I(T)| = 6$.



f	id	R	R^2	r_K	r_L	r_M
ϕ (f)	$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}$	$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$	$\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$	$\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$	$\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}$	$\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$

$$\phi(f) = \textit{permutation of the vertices induced by } f$$

Does ϕ satisfy the multiplication rule

$$\phi(f)\phi(g) = \phi(f \circ g)?$$

Yes: for really all ϕ does is *restrict* maps $f \in I(T)$ to a subset of T, namely the subset consisting of the three vertices. That is, we have an inclusion $\{1,2,3\} \hookrightarrow T$, and any f in $I(T)$ maps vertices to vertices, so there is a diagram

$$\begin{array}{ccc} T & \xrightarrow{f} & T \\ \uparrow & & \uparrow \\ \{1,2,3\} & \xrightarrow{\phi(f)} & \{1,2,3\} \end{array}$$

in which the vertical arrows are just inclusions. Now it is evident that

the composite of the restrictions is the restriction of the composite

in other words that $\phi(f) \circ \phi(g) = \phi(f \circ g)$, as required.

Important Note The isomorphism ϕ is not unique; its definition depends on the (arbitrary) labelling of the vertices as $1, 2$ and 3. If we choose a different labelling of the vertices, we get a different isomorphism $I(T) \to S_3$.
Exercise How are the two isomorphisms related? There is a neat answer, involving the permutation of $1, 2$ and 3 corresponding to the re-labelling.

5. If P_n is a regular plane polygon with n vertices, there is a map $I(P_n) \to S_n$, defined in exactly the way we defined ϕ in the previous example. It is always injective, by 2.6, but surjective only when $n = 3$. For example, $|I(Sq)| = 8$, but the group of permutations of the four vertices is S_4, with 24 elements.

6. The symmetry group $I(\text{Rect})$ of the rectangle consists of four elements: the identity, two reflections and a rotation. So far we have met two other groups of order 4: the cyclic group $\mathbb{Z}_4$ and the product group $\mathbb{Z}_2 \times \mathbb{Z}_2$. Is $I(\text{Rect})$ isomorphic to either of these?

   We can rule out $\mathbb{Z}_4$ straight away: it has an element of order 4 (in fact it has two of them) whereas every element in $I(\text{Rect})$ has order 2.

   This leaves $\mathbb{Z}_2 \times \mathbb{Z}_2$. **Exercise:** Is $I(\text{Rect})$ isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$?

7. The map $M : \text{Aut}(\mathbb{R}^n) \to Gl(n, \mathbb{R})$ sending a linear map to its matrix restricts to an isomorphism $I_0(\mathbb{R}^n) \to O(n)$, where $I_0(\mathbb{R}^n)$ is the set of isometries of $\mathbb{R}^n$ fixing 0.

8. **Exercise** By Corollary 2.9, an isometry of $\mathbb{R}^n$ is the composite of a translation and an isometry fixing 0.
   (i) Show that the subgroup of $I(\mathbb{R}^n)$ consisting of translations is isomorphic to $\mathbb{R}^n$.
   (ii) Is $I(\mathbb{R}^n)$ isomorphic to $\mathbb{R}^n \times O(n)$?

9. In Section 2 I mentioned that the product $G_1 \times G_2$ has a subgroup "much the same" as G_1, namely the subgroup $G_1 \times \{e_2\}$. With the notion of isomorphism, we can make "much the same" into a precise statement: the map $G_1 \to G_1 \times \{e_2\}$ sending g_1 to (g_1, e_2) is an isomorphism.

5.1 Classification

Algebraists tend not to distinguish between isomorphic groups, and part of our work now is to classify the finite groups we have already obtained, up to isomorphism. Ordering by size, we have met

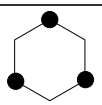
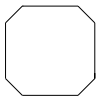
$$\mathbb{Z}_1 = \{0\}, \mathbb{Z}_2,\ \mathbb{Z}_3,\ \mathbb{Z}_4,\ \mathbb{Z}_2 \times \mathbb{Z}_2,\ \mathbb{Z}_5,\ \mathbb{Z}_6 \simeq\ \mathbb{Z}_2 \times \mathbb{Z}_3,\ S_3$$

as well as $I(Sq)$, which we still do not have an abstract name for - we know it only as group of isometries of the square - and the matrix groups

$$Gl(n, \mathbb{R}),\ Sl(n, \mathbb{R}),\ O(n),\ U(n)$$

.

Exercise: The symmetry group of each of the following plane figures is isomorphic to a group in the list just given. Complete the table by adding the name of that group.

X	A	B	C	X	+			Z	R	
I(X)										

Symmetry groups of some plane figures

5.2 Conjugation

Up to now we have looked at isomorphisms $\phi : G \to H$ where G and H are different groups. However, it is important also to consider isomorphisms from a group to itself (often known as *automorphisms*). For example, consider $I(Sq)$. In Figure 4 we list all of its subgroups; each arises as the group of symmetries of some deformation of the square. Some of the degenerations pictured look the same - one might be obtained from the other by applying a rotation, or a reflection. In this case one might expect that the corresponding subgroups of $I(Sq)$ should "sit inside $I(Sq)$ in the same way" - they may not be equal, but some automorphism of $I(Sq)$ should transform one into the other.

Similarly, we have spoken of $I(Sq)$ as if the square were unique; but there are many different squares one can draw in the plane. It seems reasonable that they all have the "same" symmetry group, in the sense that their symmetry groups are isomorphic. Indeed, if there is an isometry sending one square to the other, then the relation between their symmetry groups is straightforward, as we shall see.

The key idea is *conjugation*. Let G be any group and fix some element $g \in G$. We define a map $c_g : G \to G$ (*conjugation by* g) by $c_g(g_1) = g^{-1}g_1g$.

Proposition 5.8 *For each fixed $g \in G$, c_g is an isomorphism.*

Proof For any $g_1, g_2 \in G$,

$$c_g(g_1g_2) = g^{-1}g_1g_2g = g^{-1}g_1gg^{-1}g_2g;$$

(we insert $g^{-1}g = e$ between g_1 and g_2; this does not alter the value of the expression). Since $g^{-1}g_1gg^{-1}g_2g = c_g(g_1)c_g(g_2)$, we have shown that $c_g(g_1g_2) = c_g(g_1)c_g(g_2)$.

(The argument can be run in the opposite direction, of course:

$$c_g(g_1)c_g(g_2) = g^{-1}g_1gg^{-1}g_2g = g^{-1}g_1g_2g = c_g(g_1g_2).$$

This proves just the same thing, that $c_g(g_1g_2) = c_g(g_1)c_g(g_2)$, and avoids the apparently unmotivated step of inserting gg^{-1} between g_1 and g_2. Sometimes the proof that two expressions are equal is easier when begun at one end than at the other.)

Checking that c_g is injective is straightforward: if $g^{-1}g_1g = g^{-1}g_2g$ then multiplying on the left by g and on the right by g^{-1} we deduce that $g_1 = g_2$. Surjectivity is equally easy: for

any element $g_1 \in G$, $c_g(gg_1g^-1) = g^{-1}gg_1gg^{-1} = g_1$, and thus c_g is surjective. Thus c_g is a bijection, and this completes the proof that it is an isomorphism. □

Exercise: What is the inverse of c_g?

Now we can make precise the relation between the isometry groups of isometric figures.

Proposition 5.9 *Suppose that $X, Y \subset \mathbb{R}^n$ and that there exists an isometry $g \in I(\mathbb{R}^n)$ such that $Y = g(X)$. Then $I(X) = c_g(I(Y))$ (i.e. c_g maps $I(Y)$ to $I(X)$).*

Proof Suppose $f \in I(Y)$, and consider the diagram

$$\begin{array}{ccc} Y & \stackrel{f}{\longrightarrow} & Y \\ g\uparrow & & \downarrow g^{-1} \\ X & \stackrel{g^{-1}fg}{\longrightarrow} & X. \end{array}$$

Evidently $g^{-1}fg$ maps X to X. As the composite of three isometries, it is an isometry. That is, if $f \in I(Y)$ then $g^{-1}fg \in I(X)$, or in other words $c_g(f) \in I(X)$. This holds for every $f \in I(Y)$, so $c_g(I(Y)) \subseteq I(X)$.

To see that $c_g(I(Y))$ is *all* of $I(X)$, suppose that $h \in I(X)$. Then by the argument we used above, $ghg^{-1} \in I(Y)$ (for g^{-1} maps X to Y); now $c_g(ghg^{-1}) = h$, so h is in the image of c_g, and $c_g(I(Y))$ is all of $I(X)$. □

Exercise What if X and Y are not isometric, but merely *similar* (that is, related by the composite of an isometry and a dilation)? Example 3.1 6, concerning the centre of the group $Gl(n, \mathbb{R})$, plays a role here.

There are a number of other related geometrical situations in which conjugacy plays an important role.

Proposition 5.10 *(i) Let $f, g \in Bij(X)$, and let $Fix(f)$ denote the fixed-point set of f: $Fix(f) = \{x \in X : f(x) = x\}$. Then*

$$Fix(f) = g(Fix(c_g(f))).$$

(ii) Let K, L be two lines in the plane, meeting at O, and suppose that R is a rotation about O taking K to L. Then $r_K = c_R(r_L)$.

Proof **Exercise.** □

Exercise Take another look at the exercise in Example 5.7 4.

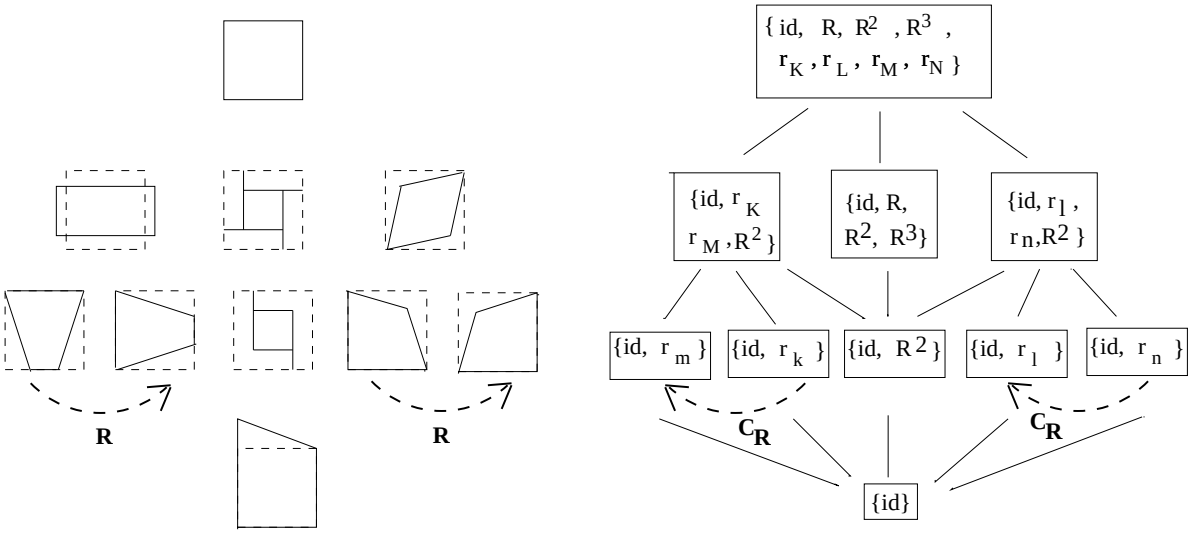
Let G be a group. The elements g_1 and g_2 are said to be *conjugate* if there exists $g \in G$ such that $c_g(g_1) = g_2$; similarly, two subgroups H_1 and H_2 of G are conjugate if there is a $g \in G$ such that $c_g(H_1) = H_2$.

Proposition 5.11 *Conjugacy (in either sense) is an equivalence relation.*

Proof **Exercise.** □

The set of all elements conjugate to a given element in a group (including the element itself) is its *conjugacy class*; this is of course the equivalence class of the element under the relation of conjugacy.

Example 5.12 Here we show again the lattice of subgroups of $I(\mathrm{Sq})$, each one corresponding to a deformation of the square. Certain of these deformations are congruent to one another *via isometries in* $I(Sq)$, and their isometry groups are in consequence conjugate subgroups of $I(\mathrm{Sq})$.



Change of basis

Let U be an n-dimensional real vector space. A choice of basis determines a bijection

$$\mathrm{L}(U) \simeq \mathrm{M}(n, \mathbb{R})$$

and an isomorphism

$$\mathrm{Aut}(U) \simeq \mathrm{Gl}(n, \mathbb{R}),$$

but since in general one basis is no better than another, there are many equally good bijections and isomorphisms. What is the relation between them all? The answer, as we shall see, is conjugacy. To sort this all out, we introduce some notation which may seem rather ponderous at first, but works like a Rolls Royce once you put it on the road.

Remember that a choice of basis gives us a way of converting vectors in our n-dimensional space U into n-tuples of numbers. Let $E = e_1, \ldots, e_n$ be a basis for U. If $u \in U$ is equal to

$\alpha_1 e_1 + \cdots + \alpha_n e_n$, then we write

$$[u]_E = \begin{bmatrix} \alpha_1 \\ \cdots \\ \cdots \\ \alpha_n \end{bmatrix}.$$

That is, $[u]_E$ *is the expression of* u *with respect to the basis* E, as a column vector.
Suppose that U and V are vector spaces with bases $E = e_1, \ldots, e_n$ and $F = f_1, \ldots, f_m$ respectively, and let $T : U \to V$ be a linear map. Let us denote the matrix of T with respect to the bases E in the source, U, and F in the target V by

$$[T]_F^E.$$

In case you've forgotten, and to try out our notation, recall that

$$[T]_F^E = \begin{bmatrix} [T(e_1)]_F & \cdots & [T(e_n)]_F \end{bmatrix}$$

— its columns are the expressions of the vectors $T(e_1), \ldots, T(e_n)$ with respect to the basis F. Multiplication of matrices is defined so that multiplication by $[T]_F^E$ converts the expression for a vector u with respect to the basis E, into the expression for $T(u)$ with respect to the basis F. That is,

$$[T(u)]_F = [T]_F^E [u]_E.$$

Once you know the rule for multiplying a column vector by a matrix, this last equality is actually obvious:

$$[T]_F^E \begin{bmatrix} \alpha_1 \\ \cdots \\ \cdots \\ \alpha_n \end{bmatrix} = \alpha_1 \cdot \text{column 1 of } [T]_F^E + \cdots + \alpha_n \cdot \text{column } n \text{ of } [T]_F^E$$

$$= \alpha_1 [T(e_1)]_F + \cdots + \alpha_n [T(e_n)]_F$$
$$= [\alpha_1 T(e_1) + \cdots + \alpha_n T(e_n)]_F$$
$$= [T(\alpha_1 e_1 + \cdots + \alpha_n e_n)]_F$$
$$= [T(u)]_F.$$

From this everything else follows. For example, suppose that $U \xrightarrow{S} V$ and $V \xrightarrow{T} W$ are linear maps, and let E be a basis for U, F a basis for V and G a basis for W. Then

$$[T]_G^F \, [S]_F^E = [T]_G^F \begin{bmatrix} [S(e_1)]_F & \cdots & [S(e_n)]_F \end{bmatrix};$$

since

$$[T]_G^F [v]_F = [T(v)]_G$$

for any $v \in V$, in particular

$$[T]_G^F [S(e_i)]_F = [T(S(e_i))]_G$$

and thus

$$[T]^F_G[S]^E_F = \left[\begin{array}{ccc} [T(S(e_1))]_G & \cdots & [T(S(e_n))]_G \end{array} \right]$$
$$= [T \circ S]^E_G.$$

The equality we have just deduced,

$$[T]^F_G[S]^E_F = [T \circ S]^E_G,$$

is the fundamental relation between matrices and linear transfomations; it figures as Theorem 7.3 in Derek Holt's Linear Algebra Lecture Notes.
The matrix for a change of basis can also be conveniently represented in this way. Suppose that E and F are now both bases for the same space U. Denote the identity map from U to U by I. Then of course $[I]^E_E$ and $[I]^F_F$ are both "identity matrices": matrices with 1's down the diagonal and zeros everywhere else: for example

$$[I]^E_E = \left[\begin{array}{ccc} [I(e_1)]_E & \cdots & [I(e_n)]_E \end{array} \right]$$
$$= \left[\begin{array}{ccc} [e_1]_E & \cdots & [e_n]_E \end{array} \right]$$
$$= \left[\left[\begin{array}{c} 1 \\ 0 \\ \cdots \\ 0 \end{array} \right] \quad \cdots \quad \left[\begin{array}{c} 0 \\ \cdots \\ 0 \\ 1 \end{array} \right] \right].$$

On the other hand, $[I]^E_F$ and $[I]^F_E$ are *not* identity matrices; for example

$$[I]^E_F = \left[\begin{array}{ccc} [e_1]_F & \cdots & [e_n]_F \end{array} \right]$$

and of course $[e_i]_F$ is not the i-th column of the identity matrix unless $e_i = f_i$.

However, $[I]^E_F$ and $[I]^F_E$ are mutually inverse:

$$[I]^E_F[I]^F_E = [I \circ I]^F_F = [I]^F_F$$

is the identity matrix. And in particular if $T : U \to U$ we have

$$[I]^F_E \; [T]^F_F \; [I]^E_F = [I \circ T]^F_E \; [I]^E_F$$
$$= [I \circ F \circ I]^E_E = [T]^E_E.$$

In other words,

$$[T]^E_E = \left(\; [I]^E_F \; \right)^{-1} \; [T]^F_F \; [I]^E_F,$$

and

$$\boxed{[T]^E_E \text{ is the conjugate of } [T]^F_F \text{ by } [I]^E_F.}$$

6 Homomorphisms

A map $\phi : G \to H$ is a *homomorphism* if, like an isomorphism, it respects the binary operation:

$$\phi(g_1 g_2) = \phi(g_1)\phi(g_2)$$

for all g_1 and g_2 in G. However, a homomorphism is *not required* to be injective nor surjective. While we're on the subject of long words with Greek etymology, how about

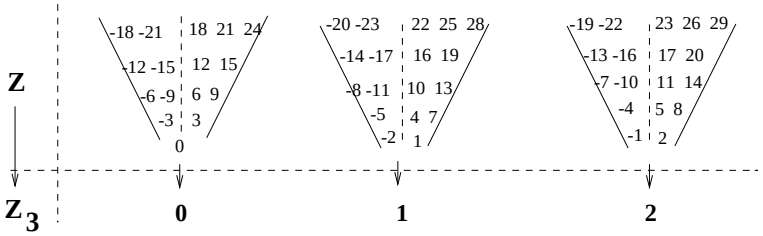
Definition 6.1 An *epimorphism* is a surjective homomorphism; a *monomorphism* is an injective homomorphism.

There are plenty of interesting examples; here are a few:

Example 6.2

1. Two minuses make a plus; so if we define a map $\text{sign} : \mathbb{R}^\times \to \{1, -1\} = \mathbb{Z}_2$ by
$$\text{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$
then we have a homomorphism.

2. Slightly less simple: for a fixed n, we map $\mathbb{Z}$ to $\mathbb{Z}_n$ by sending m to its remainder on division by n. We illustrate this for the case $n = 3$:



This is a homomorphism: suppose that

$$m_1 = q_1 n + r_1 \quad m_2 = q_2 n + r_2$$

with r_1 and r_2 between 0 and n (so $\phi(m_1) = r_1$, $\phi(m_2) = r_2$). Then

$$m_1 + m_2 = (q_1 + q_2)n + (r_1 + r_2)$$

and the remainder of $m_1 + m_2$ on division by n is the same as the remainder of $r_1 + r_2$ when it is divided by n. Since "adding together and then taking the remainder after division by n" is exactly how we add in $\mathbb{Z}_n$, we have shown that $\phi(m_1 + m_2) = \phi(m_1) + \phi(m_2)$, as required.

3. $\det : Gl(n, \mathbb{R}) \to \mathbb{R}^\times$ is a homomorphism. This accounts, for example, for the fact that the set of matrices with determinant 1 is a subgroup of $Gl(n, \mathbb{R})$. You were shown a proof that det is a homomorphism in Linear Algebra (though it was not stated in these terms); it is not a trivial result.

However, geometrically it is easy to understand. For if we ignore its sign, the determinant of a matrix A is *the factor by which the linear transformation* $A : \mathbb{R}^n \to \mathbb{R}^n$ *multiplies volume.* That is, if $R \subset \mathbb{R}^n$ is an (n-dimensional) cube, then $\mathrm{vol}(A(R)) = |\det(A)|\mathrm{vol}(R)$, and from this one can show, by approximating other subsets $X \subset \mathbb{R}^n$ by disjoint unions of cubes, that for every set X that has a volume one can measure, $\mathrm{vol}(A(X)) = |\det(A)|\mathrm{vol}(X)$. Now if A and B are two matrices then $\mathrm{vol}((AB)(X)) = \mathrm{vol}(A(B(X)) = |\det(A)|\mathrm{vol}(B(X)) = |\det(A)||\det(B)|\mathrm{vol}(X)$. Since $\mathrm{vol}((AB)(X)) = \det(AB)\mathrm{vol}(X)$, we divide through by $\mathrm{vol}(X)$ to conclude that $|\det(AB)| = |\det(A)||\det(B)|$.

This proof is an example of the principle that if one can find the *meaning* of a definition (in this case the geometric meaning of the determinant) then what was rather a complicated and formal result can become clear and simple. Of course, showing that det has this meaning itself requires some thought - it's not obvious - but it has a payoff, not just in this result but in many others.

Making mathematics meaningful is a matter of taste. For some people it will be via geometry; for others, algebraic structures alone suffice. You have to try it for yourself, to discover what works for you.

One of the most important hidden stumbling blocks in learning mathematics is failing to appreciate (often through no fault of your own) *why* a definition is the way it is, even though one "understands" what it says. Quite frequently the awful intellectual lethargy that derails one's efforts to study, can be traced to this particular kind of lack of understanding.

4. The (additive) group of rational numbers, $\mathbb{Q}$, is countable, and thus in bijection with $\mathbb{Z}$. Nevertheless, as *groups* $\mathbb{Q}$ and $\mathbb{Z}$ are very different. In fact,

Proposition 6.3 *The only homomorphism from* $\mathbb{Q}$ *to* $\mathbb{Z}$ *is the trivial homomorphism* $\phi(q) = 0$ *for all* $q \in \mathbb{Q}$. *In other words, if* $\phi : \mathbb{Q} \to \mathbb{Z}$ *is a homomorphism, then* $\phi(q) = 0$ *for all* $q \in \mathbb{Q}$.

Proof Suppose that $\phi : \mathbb{Q} \to \mathbb{Z}$ is a homomorphism and that $\phi(q) \neq 0$ for some element $q \in \mathbb{Q}$. By replacing q by $-q$ if necessary, we can suppose that $n > 0$. Since

$$\underbrace{q/n + q/n + \cdots + q/n}_{n \text{ times}} = q$$

in $\mathbb{Q}$, and since ϕ is a homomorphism, we have

$$\underbrace{\phi(q/n) + \phi(q/n) + \cdots + \phi(q/n)}_{n \text{ times}} = n$$

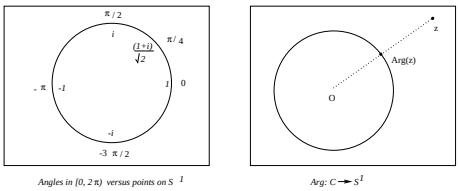
and thus $\phi(q/n) = 1$.

But then if m is any integer bigger than 1, what is $\phi(q/mn)$? By a similar argument, it must be some integer which, added to itself m times, gives 1. Of course, there is no such integer. □

The fundamental difference between $\mathbb{Q}$ and $\mathbb{Z}$, upon which our proof relies, is that in $\mathbb{Q}$ one can divide by integers, and in $\mathbb{Z}$ in general one cannot.

5. A complex number is determined by its *magnitude* or *modulus* $|z|$ and its *argument* $\arg(z)$. The map $\mathbb{C}^\times \to \mathbb{R}^\times_{>0}$ sending z to its modulus $|z|$ is easily seen to be a homomorphism: I leave to you to check it (i.e. that $|z_1 z_2| = |z_1||z_2|$).

   One is usually told that $\arg(z)$ is an angle in $[0, 2\pi)$, and that the argument of the *product* of two complex numbers is the *sum* of the arguments - so there is clearly a homomorphism somewhere here. Its domain of definition is $\mathbb{C}^\times$, but where does it go to? Before answering this, let us look at one slightly confusing point in this description. This is that if the sum of the two arguments is greater than or equal to 2π, you must subtract 2π from it, to force it back into the correct interval $[0, 2\pi)$. This seems a bit *ad hoc* and un-mathematical. There are two essentially equivalent ways of tidying this up. One is to say that instead of an angle in $[0, 2\pi)$, the argument should simply be a point on the circle S^1. After all, there is a quite natural way of identifying the interval $[0, 2\pi)$ and S^1, with the added benefit that when we go on round the circle we get back to 1 automatically, without having to use the *ad hoc* procedure of taking off 2π if the argument exceeds 2π. The circle is a more natural home for angles than $\mathbb{R}$, since the crucial fact about angles, that on reaching 2π one is back at 0, comes for free in the circle, whereas in $\mathbb{R}$ we have to impose it by saying explicitly that angles differing by a multiple of 2π are the same.



When we do regard the argument as a point on S^1, then the map $\arg : \mathbb{C}^* \to S^1$ takes on a very natural geometric description: it is simply

$$z \mapsto \frac{z}{|z|}.$$

And it is verey easy to see that arg, defined in this way, is a homomorphism (from $\mathbb{C}^*$ to S^1); for

$$\arg(z_1 z_2) = \frac{z_1 z_2}{|z_1 z_2|} = \frac{z_1}{|z_1|}\frac{z_2}{|z_2|}$$

(here we are using the fact that modulus $|\ | : \mathbb{C}^* \to \mathbb{R}_{>0}$ is a homomorphism). But notice that in this version, the argument of the product of two complex numbers is the *product*, not the sum, of their two arguments. Does this mean we are not on the right track? Wasn't the argument of the product of two complex numbers meant to be the *sum* of their arguments? We shall resolve this question later.

The other way of tidying up the definition of argument is to say that as we want the angle 0 to equal the angle 2π, we simply declare them equal.

That is, we introduce an equivalence relation in $\mathbb{R}$, defined by

$$a \sim b \text{ if } a - b \text{ is a multiple of } 2\pi;$$

after all, angles *are* the same if they differ by a multiple of 2π. We then define our map arg as going from $\mathbb{C}^*$ to the set $\mathbb{R}/\sim$ of equivalence classes of real numbers. This $\mathbb{R}/\sim$ is a group, for it turns out that we can add together equivalence classes in a perfectly sensible and natural way, and $\arg : \mathbb{C}^* \to \mathbb{R}/\sim$ then becomes a homomorphism; but we leave the rest of this story for a later chapter.

Making different things equal by means of an equivalence relation is an idea with a big future — some of which you will see in this course.

Another general point worth mentioning here is that **Exercise** Prove that $\mathbb{C}^\times$ is isomorphic to $\mathbb{R}_{>0} \times S^1$.

So here is a curious fact: our two favourite co-ordinate systems on the plane are both group isomorphisms. Cartesian co-ordinates give us an isomorphism

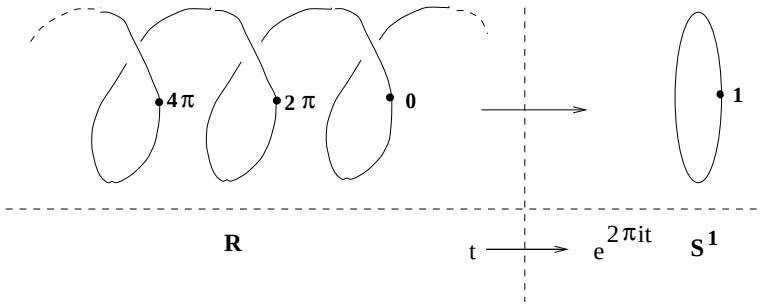
$$(\text{the plane, vector addition}) \simeq (\mathbb{R} \times \mathbb{R}, +)$$

while polar coordinates give us an isomorphism

$$(\text{the plane} \setminus \{0\}, \text{complex multiplication}) \simeq (\mathbb{R}_{>0} \times S^1, \times).$$

6. Another important homomorphism involving the unit circle $S^1 \subset \mathbb{C}$ is the map

$$\exp : \mathbb{R} \to S^1$$
$$\exp(t) = e^{2\pi i t}.$$



The exponential map from $\mathbb{R}$ to S^1

The picture shows the real line, bent into a spiral, and then simply projected onto the circle on the right. The fact that exp is a homomorphism is often called the law of indices: $\exp(t_1 + t_2) = \exp(t_1)\exp(t_2)$, or in other words $e^{2\pi i(t_1+t_2)} = e^{2\pi i t_1} e^{2\pi i t_2}$. If we

assume de Moivre's theorem that $\exp(t) = \cos\ 2\pi t + i \sin\ 2\pi t$, then the fact that exp is a homomorphism is equivalent to the two angle sum formulae

$$\cos(t_1 + t_2) = \cos\ t_1 \cos\ t_2 - \sin\ t_1 \sin\ t_2$$

and

$$\sin(t_1 + t_2) = \sin\ t_1 \cos\ t_2 + \cos\ t_1 \sin\ t_2.$$

Exercise Prove this. In other words, prove that *if* we assume de Moivre's theorem, then

(i) the two angle sum formulae follow from the fact that exp is a homomorphism, and

(ii) the two angle sum formulae together *imply* that exp is a homomorphism.

7. We define a homomorphism $\phi : I(\mathrm{Sq}) \to \mathbb{Z}_2$ as follows: in Example 3.1 we remarked that every element of $I(Sq)$ either interchanges the diagonals of the square, or leaves them both where they are. Define

$$\phi_1(f) = \begin{cases} 1 & \text{if } f \text{ leaves the diagonals where they are} \\ -1 & \text{if } f \text{ interchanges the diagonals} \end{cases}$$

You can easily check (and should) that this map really is a homomorphism.

8. **Exercise** Define another homomorphism $\phi_2 : I(\mathrm{Sq}) \to \mathbb{Z}_2$ by considering the lines K and M (in Figure 1, page 4) instead of the diagonals.

9. Another homomorphism $I(Sq) \to \mathbb{Z}_2$: we assume that the centre of the square is at 0, so that every isometry of the square must be a linear map (*cf 2.8*). Now define a map $I(\mathrm{Sq}) \to \mathbb{Z}_2$ by $f \mapsto \mathrm{sign}(\det(f))$. In fact it is unnecessary to write $\mathrm{sign}(\det(f))$; for any isometry f one has $\det(f) = \pm 1$ (**Exercise** — after all, what does an isometry do to volume?), and thus the map det itself maps $I(Sq)$ to $\mathbb{Z}_2$.
Exercise Show that the commutator subgroup of $I(Sq)$ is not all of $I(Sq)$. Hint: what is the determinant of a commutator $a^{-1}b^{-1}ab$, for $a, b \in I(Sq)$?

10. Any vector space is a group under addition; this is part of the structure of being a vector space. And any linear map $T : V \to W$ is a homomorphism of groups.

11. Let a be a fixed integer and define a map $\phi_a : \mathbb{Z} \to \mathbb{Z}$ by $\phi_m(m) = am$. Then ϕ_a is a homomorphism. When is it an isomorphism?

12. Similarly, for $a \in \mathbb{Z}_n$ we define $\phi_a : \mathbb{Z}_n \to \mathbb{Z}_n$ by $\phi_a(m) = am$. Here am means the remainder of am after division by n. When is ϕ_a an isomorphism? See Proposition 4.6 for guidance.

13. Every permutation in S_n can be written as a product of transpositions: for example $(2, 3, 5, 7) = (2, 3)(3, 5)(5, 7)(7, 2)$. The same permutation can be written in several different ways as a product of transpositions: for example the transposition $(2, 3)$ itself can be written $(1, 2)(1, 3)(1, 2)$ (check it!). A permutation is *even* if it can be written as the product of an even number of transpositions, and *odd* otherwise. It turns out (see Exercise 18 on Sheet 2) that if a permutation can be expressed as the product of an even

number of transpositions, then *every* expression of the same permutation as product of transpositions, consists of an even number: once even, always even. So the same goes for odd permutations.

Proposition 6.4 *The map $S_n \to \mathbb{Z}_2$ sending σ to 0 if σ is even, and to 1 if σ is odd, is a homomorphism.*

Proof If σ_1 is the product of k_1 transpositions and σ_2 is the product of k_2 transpositions, then $\sigma_1\sigma_2$ is the product of $k_1 + k_2$ transpositions. Thus the composite of two permutations is

even	if both are even
odd	if one is odd and one is even
even	if both are odd.

Since these three lines remain true if we replace

"composite of two permutations"	with	"product of two numbers"
"even"	with	0
"odd"	with	1

it follows that the map is a homomorphism.

□

14. The *projection* maps $p_1 : G_1 \times G_2 \to G_1$ and $p_2 : G_1 \times G_2 \to G_2$ given by $p_1((g_1, g_2)) = g_1$ and $p_2((g_1, g_2)) = g_2$ are homomorphisms. So are the maps $j_1 : G_1 \to G_1 \times G_2$ and $j_2 : G_2 \to G_1 \times G_2$ defined by $j_1(g_1) = (g_1, e_2)$ and $j_2(g_2) = (e_1, g_2)$.

Some simple fact about homomorphisms:

Proposition 6.5 *Let $\phi : G_1 \to G_2$ be a homomorphism of groups. Then*

1. *$\phi(e_1) = e_2$;*
2. *for all $g \in G_1$, $\phi(g^{-1}) = (\phi(g))^{-1}$*
3. *for all $g_1, \cdots, g_r \in G_1$, $\phi(g_1 g_2 \cdots g_r) = \phi(g_1)\phi(g_2)\cdots\phi(g_r)$;*
4. *for all $g \in G_1$, $o(\phi(g))$ divides $o(g)$.*

Proof The proofs are the same as the proofs in the (practically identical) Proposition 5.5, except for the last statement, which I leave as an **Exercise** □

Proposition 6.6 *The composite of two homomorphisms is itself a homomorphism.*

Proof **Exercise**. □

Associated to every group homomorphism $\phi : G \to H$ are two important subgroups, one of G and one of H:

Definition 6.7 Let $\phi : G \to H$ be a homomorphism of groups. Then

$$\ker(\phi) = \{g \in G : \phi(g) = e_H\}$$

and

$$\mathrm{Im}(\phi) = \{\phi(g) : g \in G\}.$$

$\mathrm{Im}(\phi)$ can also be written as $\{h \in H : \exists g \in G \text{ s.t. } \phi(g) = h\}$.

Proposition 6.8 *If $\phi : G \to H$ is a homomorphism of groups then $ker(\phi)$ is a subgroup of G.*

Proof First we check closure. If $g_1, g_2 \in \ker(\phi)$ we have to show that $g_1 g_2 \in \ker(\phi)$. But $\phi(g_1 g_2) = \phi(g_1)\phi(g_2) = e_H e_H = e_H$ and so $g_1 g_2$ is in $\ker(\phi)$. Thus, $\ker(\phi)$ satisfies the closure axiom.

Now we need only check the inverse axiom. Suppose $g \in \ker(\phi)$. We want to show that $g^{-1} \in \ker(\phi)$ also. But $\phi(g^{-1}) = \phi(g)^{-1} = e_H^{-1} = e_H$, so indeed $g^{-1} \in \ker(\phi)$. This completes the proof that $\ker(\phi)$ is a subgroup of G.

Proposition 6.9 *If $\phi : G \to H$ is a homomorphism of groups then $Im(\phi)$ is a subgroup of H.*

Proof **Exercise**. □

Exercise: Go back to Example 6.2 and find the kernel of each of the three homomorphisms $\mathrm{I(Sq)} \to \mathbb{Z}_2$ described there. According to Section 3.1, every subgroup of I(Sq) is the symmetry group of some deformation of the square; which are they in these three cases?

Every isometry has determinant equal to ± 1; those for which the determinant is $+1$ are called "direct isometries", or "orientation- preserving" isometries. For any subset X of $\mathbb{R}^n$, the direct isometries in $I(X)$ form a subgroup, the kernel of $\det : I(X) \to \mathbb{Z}_2$.

Part of the importance of the kernel of a homomorphism is due to the following property:

Lemma 6.10 *Suppose that $\phi : G \to H$ is a homomorphism of groups. Then ϕ is 1-1 if and only if $ker(\phi) = \{e_G\}$.*

Proof "Only if" is trivial; to prove "if", suppose $\phi(g_1) = \phi(g_2)$. Then $\phi(g_1 g_2^{-1}) = \phi(g_1)\phi(g_2^{-1}) = \phi(g_1)\phi(g_2)^{-1}$. As $\phi(g_1) = \phi(g_2)$, we have $\phi(g_1)\phi(g_2)^{-1} = e_H$, so $\phi(g_1 g_2^{-1}) = e_H$. But we are supposing that the only element of G sent to e_H by ϕ is e_G, so we must have $g_1 g_2^{-1} = e_G$, and thus $g_1 = g_2$. □

In fact one can make a stronger statement: if $\phi : G \to H$ is a homomorphism of groups, then every element in $\mathrm{Im}(\phi)$ has the same number of preimages in G, or in other (better) words, for any two elements h_1 and h_2 of $\mathrm{Im}(\phi)$, there is a bijection $\phi^{-1}(h_1) \to \phi^{-1}(h_2)$.

Warning: "$\phi^{-1}(h_i)$" means the preimage of h_i under ϕ, i.e. the set of all elements in G sent to to h_i by ϕ. It does *not* assume the existence of an inverse map ϕ^{-1}. Note that by definition $\ker(\phi) = \phi^{-1}(e_H)$.

I leave you to construct this bijection, following this recipe:

Exercise (i) Suppose that $\phi : G \to H$ is a homomorphism of groups, and that $g_0 \in G$. Denote $\phi(g_0)$ by h_0. Show that the map

$$\begin{aligned} \ker(\phi) &\to \phi^{-1}(h_0) \\ g &\mapsto gg_0 \end{aligned}$$

is a bijection.
(ii) Suppose $g_1, g_2 \in G$, and write $\phi(g_1) = h_1, \phi(g_2) = h_2$. Find an explicit bijection $\phi^{-1}(h_1) \to \phi^{-1}(h_2)$.

Whereas isomorphisms help us to see that there are rather fewer groups than at first one might think (we tend to regard two isomorphic groups as the same), homomorphisms help us to see the connections between things which we continue to regard as different. They are like the connections between the neurons making up a brain. Having lots of neurons is not on its own much use - the important things is that they should be multiply interconnected. So *cherchez l'homomorphisme!*

Of course, if in the course of constructing a homomorphism between two *a priori* unrelated groups, you find that it is an isomorphism, you will either have scored a triumph by showing that the two groups have the same structure, or you will have rung up another disappointment by finding that an apparently new object is actually a boring old one.

Commutative Diagrams

A diagram of sets and maps, such as

$$\begin{array}{ccc} A & \stackrel{f}{\to} & B \\ g\downarrow & & \downarrow i \\ D & \stackrel{h}{\to} & C \end{array}$$

is said to be *commutative* if an element always arrives at the same destination whichever of the different possible routes it follows. In this diagram, given $a \in A$, going across via f and then down via i sends it to $i(f(a))$, while going down via g and then across via h sends it to $h(g(a))$. We say that the diagram is commutative if these two are equal: $h(g(a)) = i(f(a))$, for all $a \in A$, or in other words if $i \circ f = h \circ g$.

Example 6.11 1. Consider the diagram

$$\begin{array}{ccc} & \mathbb{Z} & \\ f\swarrow & & \searrow g \\ \mathbb{Z}_6 & \stackrel{h}{\to} & \mathbb{Z}_2 \end{array}$$

in which the maps $f : \mathbb{Z} \to \mathbb{Z}_6$ and $g : \mathbb{Z} \to \mathbb{Z}_2$ are the standard homomorphisms described in Example 6.2(2), and the map $h : \mathbb{Z}_6 \to \mathbb{Z}_2$ is a variant of this: $h(m) =$ remainder of m on division by 2. This diagram is commutative.

2. Consider the diagram

$$\begin{array}{ccc} & I(T) & \\ \phi \swarrow & & \searrow \det \\ S_3 & \overset{\text{sign}}{\to} & \mathbb{Z}_2 \end{array}$$

in which $I(T)$ is the isometry group of an equilateral triangle and

$\phi : I(T) \to S_3$	is the isomorphism discussed in Example 5.7(3)
$\text{sign} : S_3 \to \mathbb{Z}_2 = \{\pm 1\}$	is the sign homomorphism described in Example 6.2 (11)
$\det : I(T) \to \mathbb{R}^\times$	necessarily falls in $\mathbb{Z}_2$.

I claim that this diagram is commutative: that $\text{sign}(\phi(f)) = \det(f)$ for every $f \in I(T)$. The proof of this claim is not trivial. It goes as follows: suppose that f is a reflection. Then $\det(f) = -1$. Now f interchanges two vertices and leaves one fixed; so $\phi(f)$ is a single transposition, and thus $\text{sign}(\phi(f)) = -1$. Hence $\det(f) = \text{sign}(\phi(f))$, as claimed. As a concrete example, let $f = r_L$ (notation as in the figure in Example 5.7(3)). Then f fixes vertex 1 and interchanges vertices 2 and 3. Hence $\phi(f)$ is the permutation $(2, 3)$, a transposition.

We have shown that $\text{sign} \circ \phi$ and det agree on all of the reflections in $I(T)$. Since

- reflections generate $I(T)$, and
- $\text{sign} \circ \phi$ and det are homomorphisms,

it follows (by an **Exercise**) that $\text{sign} \circ \phi$ and det agree on all of $I(T)$.

3. **Exercise:** Is the diagram

$$\begin{array}{ccc} & I(\text{Sq}) & \\ \phi \swarrow & & \searrow \det \\ S_4 & \overset{\text{sign}}{\to} & \mathbb{Z}_2 \end{array}$$

commutative? Here for each $f \in I(\text{Sq})$, $\phi(f)$ is once again the permutation of the vertices that it induces.

7 Normal Subgroups and Quotient Groups

In the last section we saw that the kernel of a homomorphism $\phi : G_1 \to G_2$ is a subgroup of G_1. Not every subgroup arises in this way, though - kernels belong to an extremely important special class of subgroup:

Definition 7.1 *Let G be a group and H a subgroup. Then H is* normal *if for all g in G and $h \in H$, we have $g^{-1}hg \in H$*

The defining property $g^{-1}hg \in G \ \forall \ g \in G, h \in H$ is often written simply

$$g^{-1}Hg \subset H \quad \text{for all} \quad g \in G.$$

Actually, under these circumstances $g^{-1}Hg = H$, but we do not care about the equality — it follows from the fact that $g^{-1}Hg \subset H$, and that's all that interests us here.

Proposition 7.2 *If $\phi : G_1 \to G_2$ is a homomorphism of groups then $ker(\phi)$ is a normal subgroup of G.*

Proof We already know that $\ker(\phi)$ is a subgroup. Suppose $g \in G$ and $h \in \ker(\phi)$. We have just to show that $g^{-1}hg \in \ker(\phi)$, or in other words that $\phi(g^{-1}hg) = e$. This is plain sailing:

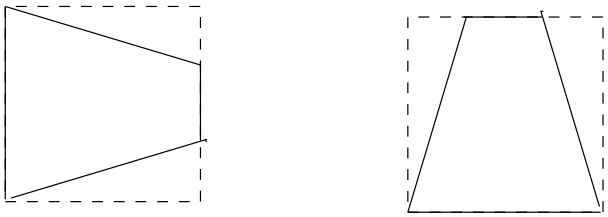
$$\begin{array}{rcll} \phi(g^{-1}hg) & = & \phi(g^{-1})\phi(h)\phi(g) & \text{since } \phi \text{ is a homomorphism} \\ & = & \phi(g^{-1})e\phi(g) & \text{since } h \in \ker(\phi) \\ & = & \phi(g)^{-1}\phi(g) & \text{since } \phi(g^{-1}) = \phi(g)^{-1}, \\ & = & e & \end{array}$$

□

In a disturbing failure of symmetry in the fundamental structure of the universe, the *image* of a group homomorphism $G \to H$ is *not* a normal subgroup of H, in general.

Example 7.3 1. Every subgroup of an abelian group is normal! This is easy: $g^{-1}hg = h$ for any g and any h, and thus $g^{-1}Hg = H$ and H is normal.

2. The subgroup $H = \{\mathrm{id}, r_K\}$ of $I(\mathrm{Sq})$ is not normal - for $R^{-1}r_K R = r_M \notin H$. In fact one can understand this geometrically also: H is the isometry group of the shape S' obtained from the square by shortening its right hand edge,



$I(S')$ is not a normal subgroup of $I(Sq)$

and if we apply R to this shape we get a different degeneration of the square, with different isometry group. By Proposition 5.9, $I(R(S')) = RI(S')R^{-1}$ so $I(S')$ cannot be normal.

3. On the other hand, the rotation subgroup $\{\mathrm{id}, R, R^2, R^3\}$ of $I(Sq)$ is normal; indeed, the rotation subgroup of $I(P_n)$ is normal for any n. This is easy to see geometrically: it is not hard to check that the conjugate of a rotation by another rotation is the first rotation (more formally, $(R^j)^{-1}R^kR^j = R^k$) and that the conjugate of a rotation by a reflection is also a rotation. Since every element of $I(P_n)$ is either a rotation or a reflection, then what we have said implies that the rotation subgroup is normal.

   However, there are two alternative proofs here.

First, if r_L is a reflection and $[r_L]$ is its matrix, then $\det([r_L]) = -1$. This is evident if L is the x-axis; for the general case, just choose a basis for $\mathbb{R}^2$ consisting of a vector in L and a vector orthogonal to L. The matrix of r_L with respect to this basis is

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

with determinant -1, and of course the determinant is independent of the choice of basis.

On the other hand the determinant of the matrix of a rotation is is equal to $+1$:

$$[R_{0,\theta}] = \begin{pmatrix} \cos\ \theta & -\sin\ \theta \\ \sin\ \theta & \cos\ \theta \end{pmatrix}.$$

It follows that the rotation subgroup of $I(P_n)$ is the kernel of the homomorphism $I(P_n) \to \mathbb{Z}_2$ defined by $f \mapsto \det[f]$, and hence is normal.

The second alternative proof involves *cosets*, and deserves a paragraph all of its own, so we postpone it for now.

Exercise Which subgroups of $I(Sq)$ are normal?

4. The centre $Z(G)$ of any group G (Example 3.1(3)) is a normal subgroup of G; so is the commutator subgroup $[G, G]$ (Example 3.2(4)). I leave the proof to you as an exercise. In fact these two subgroups have an even stronger property than normality: if $\phi : G \to G$ is *any* automorphism then $\phi(Z(G)) = Z(G)$ and $\phi([G, G]) = [G, G]$ (**Exercise**) The defining property of normality is only that the subgroup should be mapped to itself by automorphisms of a certain kind, namely those induced by conjugation by group elements. Subgroups mapped to themselves by every automorphism of the group are called *characteristic subgroups*. Thus, every characteristic subgroup is normal.

5. In S_n, the set A_n of even permutations (those which can be expressed as the product of an even number of transpositions - see Proposition 6.4 and Exercise II 17) is a normal subgroup. It is very easy to prove this directly, but more interesting to obtain it as a consequence of the fact that A_n is the kernel of the homomorphism sign : $S_n \to \mathbb{Z}_2$. defined by

$$\text{sign } (\sigma) = \begin{cases} +1 & \text{if } \sigma \text{ is the product of an even number of transpositions} \\ -1 & \text{if } \sigma \text{ is the product of an odd number of transpositions} \end{cases}$$

(cf. Proposition 6.4). Normality of A_n now follows by 7.2. A_n is called the *alternating group on n objects*.

Cosets

Let G be a group and H a subgroup. In Foundations you looked at the *cosets* of H in G, whose definition I now recall:

Definition 7.4
The *left coset* gH is the set $\{gh : h \in H\}$;
the *right coset* Hg is the set $\{hg : h \in H\}$.

As revision in these ideas, let's run through the proof of Lagrange's Theorem, which you also met in Foundations.

Theorem 7.5 *Let G be a finite group and H a subgroup. Then $|H|$ divides $|G|$.*

Proof
Step 1 *The left cosets of H partition G.*
In other words, each element of G lies in one and only one left coset of H. To see this, we suppose that g lies in the coset g_1H and in the coset g_2H, and show that it must follow that g_1H and g_2H are *the same coset*.

For

$$\begin{aligned} g \in g_1H &\Rightarrow g = g_1h_1 \text{ for some } h_1 \in H \\ g \in g_2H &\Rightarrow g = g_2h_2 \text{ for some } h_2 \in H \end{aligned}$$

Hence

$$g_1h_1 = g_2h_2$$

and so

$$g_1 = g_2h_2h_1^{-1}.$$

If h is any element of h, we have $g_1h = g_2h_2h_1^{-1}h \in g_2H$; that is,

$$g_1H \subset g_2H.$$

By the symmetry of the hypothesis, we also have

$$g_2H \subset g_1H,$$

and the two cosets are equal.

(Of course, this does not mean that g_1 and g_2 coincide. We can think of g_1 and g_2 as different labels for the same coset.)

Step 2 *All the left cosets of H have the same number of elements as H.*
This is very easy: if g_1 is any element of G then we define a map $H \to g_1H$ by $h \mapsto g_1h$. It's obvious that this is a *surjection*, but it's also an *injection*, for if $g_1h = g_1h'$, then cancelling g_1 we get $h = h'$. So the map is a bijection, and thus $|g_1H| = |H|$.

Step 3 From Steps 1 and 2 it follows immediately that

$$\begin{aligned} |G| &= \text{the number of distinct left cosets of } H \times \text{the number of elements in each coset} \\ &= \text{the number of distinct left cosets of } H \ \times |H|. \end{aligned}$$

Thus, $|G|$ is divisble by H. □

The proof would work just the same with "right coset" in place of "left coset".

For future use, we note

Lemma 7.6

$$g_1H = g_2H \iff g_1^{-1}g_2 \in H; \qquad Hg_1 = Hg_2 \iff g_1g_2^{-1} \in H.$$

Proof **Exercise** □

It is slightly irritating to have these two distinct collections of cosets; however, under some circumstances left cosets and right cosets coincide.

Lemma 7.7 *Suppose that N is a normal subgroup of G. Then for each $g \in G$, $gN = Ng$.*

Proof Since N is normal, $g^{-1}Ng = N$; left-multiplying both sides of this equation by g, we get $Ng = gN$. In case this notation is a bit disconcerting, (multiplying subgroups and elements together as it does), I spell the argument out in more detail:

$$g^{-1}Ng = N$$

means

$$\{g^{-1}ng : n \in N\} = \{n : n \in N\};$$

multiplying this equality by g on the left means left-multiplying every element in the two equal sets by g, giving

$$\{ng : n \in N\} = \{gn : n \in N\},$$

which we immediately recognise as the statement that

$$Ng = gN.$$

□

Exercise Prove the converse of this proposition: if $gN = Ng$ for every $g \in G$, then N is normal.

Quotient Groups

Let N be a normal subgroup of G. We now intoduce the symbol G/N for the set of cosets of N. Miraculously, normality of N means that on G/N we can define a group structure, in a perfectly natural way. With this operation, G/N is called the *quotient group* of G by N. This is the last of the important concepts we will study in this course, and crucial to almost all later developments in algebra.

The group operation is simply

$$g_1N \ \times \ g_2N \ = \ (g_1g_2)N.$$

On the face of it, it's not clear what the normality of N is contributing to this. We could write down the same rule even if N were not normal. The point, though, is that if N were not normal, the rule would not be well-defined — despite appearances, it would not enable us to calculate the product of two cosets. The problem is that the same coset "has many different labels" — two cosets gN and $g'N$ can be the same, even though g and g' are different. Suppose

that $g_1'N = g_1N$ and $g_2'N = g_2N$. In order that our rule make sense, we would need to be sure that under these circumstances

$$(g_1'g_2')N \;=\; (g_1g_2)N.$$

And it turns out that this is the case only when N is a normal subgroup:

Proposition 7.8 *Suppose that N is a normal subgroup of G, and suppose that $g_1'N = g_1N$ and $g_2'N = g_2N$. Then*

$$(g_1'g_2')N \;=\; (g_1g_2)N.$$

Conversely, if N is a subgroup of G and $(g_1'g_2')N \;=\; (g_1g_2)N$ whenever $g_1'N = g_1N$ and $g_2'N = g_2N$, then N is normal.

Proof As $g_1'N = g_1N$, it follows from 7.6 that $g_1^{-1}g_1' \in N$. Again by 7.6, to show that $(g_1g_2)N = (g_1'g_2')N$, it's enough to show that $(g_1g_2)^{-1}(g_1'g_2') \in N$. But $(g_1g_2)^{-1}(g_1'g_2') = g_2^{-1}g_1^{-1}g_1'g_2'$. If g_2' and g_2 were equal, this last expression would just be the conjugate $c_{g_2}(g_1^{-1}g_1')$ of $g_1^{-1}g_1'$ (which lies in N) by g_2, and by normality of N this lies in N. However, since we do not assume $g_2' = g_2$, we write

$$g_2^{-1}g_1^{-1}g_1'g_2' = g_2^{-1}g_1^{-1}g_1'(g_2g_2^{-1})g_2' = c_{g_2}(g_1^{-1}g_1') \times (g_2^{-1}g_2').$$

Now both $c_{g_2}(g_1^{-1}g_1')$ and $g_2^{-1}g_2'$ lie in N (the latter because $g_2N = g_2'N$), and therefore so does their product, and so we have indeed shown that $(g_1g_2)^{-1}(g_1'g_2') \in N$. This completes the proof that $(g_1g_2)N = (g_1'g_2')N$.

For the converse, suppose $g \in G$ and $n \in N$. We have to line up our data to suit the form of our hypothesis, and after a little thought I use the hypothesis in the following form: I take

$$g_1 = g, g_1' = gn, g_2 = g^{-1}, g_2' = g^{-1}.$$

Note that $g_1N = g_1'N$, by 7.6, since $g_1^{-1}g_1' = n \in N$, and trivially $g_2N = g_2'N$, since $g_2 = g_2'$. Our hypothesis (that together $g_1N = g_1'N$ and $g_2N = g_2'N$ imply $(g_1g_2)N = (g_1'g_2')N$) now gives us

$$(gg^{-1})N = (gng^{-1})N,$$

i.e.

$$eN = g^{-1}ngN.$$

By 7.6, this means that $g^{-1}ng \in N$. We conclude that N is normal. □

To summarise,

The operation $g_1N \cdot g_2N = (g_1g_2)N$ is well-defined if and only if N is normal.

From here it is an easy step to the stronger statement that this operation makes the set of left cosets into a group:

Associativity

$$(g_1N \cdot g_2N) \cdot g_3N = (g_1g_2)N \cdot g_3N = (g_1g_2)g_3N;$$

by associativity in G, $(g_1g_2)g_3 = g_1(g_2g_3)$ and so $(g_1g_2)g_3N = g_1(g_2g_3)N$; and $g_1(g_2g_3)N = g_1N \cdot (g_2N \cdot g_3N)$.

Neutral Element: eN (i.e. N itself) is the neutral element.

Inverses Evidently

$$g^{-1}N \ \cdot \ gN = g^{-1}gN = eN$$

so the inverse of gN is $g^{-1}N$.

Thus, we have proved

Theorem 7.9

If N is normal then the operation $g_1N \ \cdot \ g_2N = (g_1g_2)N$ makes G/N into a group.

In abelian groups we usually denote the binary operation with a "+" rather than by juxtaposition. So if A is an abelian group and H a subgroup, the coset of an element $a \in A$ is denoted $a + H$ rather than aH. And the group operation in A/H becomes $(a_1 + H) + (a_2 + H) = (a_1 + a_2) + H$, rather than $(a_1H)(a_2H) = (a_1a_2)H$.

Examples of Quotient Groups

Example 7.10 The simplest examples to deal with are where G is an abelian group — in this case every subgroup is of course normal.

1. Earlier (in Example 6.2 (4)), we discussed the problem of adding angles, and the fact that angles differing by a multiple of 2π are really the same. I suggested that one way round the difficulty was simply to declare that any two angles differing by an integer multiple of 2π are equal. Formally, this is done by introducing an equivalence relation on $\mathbb{R}$:

   $$a_1 \sim a_2 \quad \text{if } a_1 - a_2 \text{ is an integer multiple of } 2\pi.$$

   In other words,

   $$a_1 \sim a_2 \quad \text{if } a_1 - a_2 \in 2\pi\mathbb{Z}.$$

   In view of Lemma 7.6, and the fact that $2\pi\mathbb{Z}$ is a subgroup of $\mathbb{R}$, the equivalence relation is just the relation of being in the same coset of $2\pi\mathbb{Z}$:

   $$a_1 \sim a_2 \quad \text{if } a_1 + 2\pi\mathbb{Z} = a_2 + 2\pi\mathbb{Z}.$$

   And the equivalence classes of $\sim$ are just the cosets of $2\pi\mathbb{Z}$ in $\mathbb{R}$. At the time I introduced this equivalence relation, I said that later we would see that one could add equivalence classes in a natural way. With our definition of quotient groups, in this case $\mathbb{R}/2\pi\mathbb{Z}$, we have now achieved this.

2. The fact that the relation $\sim$ introduced in the previous paragraph is an equivalence relation (i.e. is symmetric, reflexive and transitive) is intimately connected with the fact that $2\pi\mathbb{Z}$ is a *subgroup* of $\mathbb{R}$:

   Exercise 7.11 Let A be an abelian group and H a subgroup. Define a relation $\sim$ on the elements of A, by

   $$g_1 \sim g_2 \quad \text{if} g_1 - g_2 \in H.$$

   Show that $\sim$ is an *equivalence* relation.

   In fact, each of the two properties that ensure that H is a subgroup (being closed under the binary operation, and containing the inverse of easch of its elements), gives rise in a rather charming way to one of the defining properties of an equivalence relation.

3. For G we take the group $\mathbb{Z}$, and for N the subgroup of even integers, which for obvious reasons we denote by $2\mathbb{Z}$. What is G/N? First, how many distinct cosets of N are there? Recall from 7.6 (now written additively) that

   $$n_1 + N = n_2 + N \iff n_1 - n_2 \in N.$$

   That is, $n_1 + N = n_2 + N$ if and only if $n_1 - n_2$ is even. But that means there are just two cosets:

   $2\mathbb{Z}$ itself i.e. all the even integers, and
   $1 + 2\mathbb{Z}$ i.e. all the odd integers.

   For the difference between any two even integers is even, and so is the difference between any two odd numbers. To determine the group structure in $\mathbb{Z}/2\mathbb{Z}$, we observe that

   $$\begin{array}{rcl} \text{even} + \text{even} & = & \text{even} \\ \text{odd} + \text{even} & = & \text{odd} \\ \text{odd} + \text{odd} & = & \text{even} \end{array}$$

   and so $\mathbb{Z}/2\mathbb{Z}$ is just our old acquaintance $\mathbb{Z}_2$.

4. What about $\mathbb{Z}/n\mathbb{Z}$?
   First observation: $n_1 - n_2 \in n\mathbb{Z}$ if and only if n_1 and n_2 have the same remainder on division by n. So the distinct cosets of $n\mathbb{Z}$ in $\mathbb{Z}$ are

   $$n\mathbb{Z}, \quad 1 + n\mathbb{Z}, \quad 2 + n\mathbb{Z}, \quad \ldots, (n-1) + n\mathbb{Z};$$

   for

   $$\begin{array}{lcl} n\mathbb{Z} & = & \text{all integers with remainder 0 on division by } n \\ 1 + n\mathbb{Z} & = & \text{all integers with remainder 1 on division by } n \\ \ldots & \ldots & \ldots\ldots\ldots\ldots \\ \ldots & \ldots & \ldots\ldots\ldots\ldots \\ (n-1) + n\mathbb{Z} & = & \text{all integers with remainder } n-1 \text{ on division by } n \end{array}$$

   Thus, the elements of $\mathbb{Z}/n\mathbb{Z}$ are just the same as the elements of $\mathbb{Z}_n$.

Second observation: $(m_1+n\mathbb{Z})+(m_2+n\mathbb{Z}) = (m_1+m_2)+n\mathbb{Z}$, but if we want to view this coset in the list $n\mathbb{Z}, 1+n\mathbb{Z}, \ldots, (n-1)+n\mathbb{Z}$, we may have to subtract n from m_1+m_2; in any event, the coset $m+n\mathbb{Z}$ is determined by the remainder of m on division by n, and so we can summarise the addition procedure by

$$\boxed{(m_1+n\mathbb{Z})+(m_2+n\mathbb{Z}) = m+n\mathbb{Z}, \text{ where } m \text{ is the remainder of } m_1+m_2 \text{ after division by } n.}$$

So $\mathbb{Z}/n\mathbb{Z}$ is just the group $\mathbb{Z}_n$ of integers modulo n.

5. Now for a non-abelian example. In Section 3 II, I defined the *commutator subgroup* or *derived subgroup* $[G,G]$ of a group G, as the subgroup generated by all the commutators $a^{-1}b^{-1}ab$ of elements $a, b \in G$. I left you to show that it is a normal subgroup.

   Proposition 7.12 *The quotient $G/[G,G]$ is an abelian group!*

   Proof For brevity, we denote $[G,G]$ by N. We have to show that if aN and bN are any two cosets, then
   $$aN \cdot bN = bN \cdot aN.$$
   That is, we have to show that $abN = baN$. By 7.6, this is equivalent to having $(ba)^{-1}ab \in N$. But $(ba)^{-1}ab$ is just the commutator $a^{-1}b^{-1}ab$, and of course this is in N. □

   This example is rather important. In some ways it illustrates the same basic idea as the first of our examples of quotient groups: we can make different things equal one another, by decree. In this case, in the (non-abelian) group G we have (in general) $ba \neq ab$, or, in other words, $a^{-1}b^{-1}ab \neq e$. To make $a^{-1}b^{-1}ab$ equal e, we simply declare it to be so: we put all the commutators into a subgroup, and on taking the quotient of G by this subgroup, we obtain a group in which the coset of a times the coset of b is equal to the coset of b times the coset of a. In other words, by putting all the commutators $a^{-1}b^{-1}ab$ into a subgroup and taking the quotient of G by this subgroup, we kill the difference between ab and ba, for every a and b.

A question of notation: the symbol "aN" (or "$a+N$" in the abelian case) is clumsy and makes things seem more difficult than they are. It is often more comfortable and more revealing to write $\bar{a}$ for aN, or even $[a]$. For example, the fact, seen in the last example, that the quotient of a group G by its commutator subgroup is abelian, appears as

$$\bar{a}\bar{b} = \bar{b}\bar{a}.$$

I will sometimes use this notation, where it seems to me that it simplifies things.

Exercise 7.13 The assumption in Exercise 7.11 that the group A be abelian is actually completely irrelevant. In fact if G is *any* group and H any subgroup, one can define a relation $\sim$ in either of two different ways: by

$$g_1 \sim g_2 \quad \text{if } g_1^{-1}g_2 \in H$$

or by

$$g_1 \sim g_2 \quad \text{if } g_1 g_2^{-1} \in H.$$

Show that each of these is an equivalence relation.

Exercise 7.14 Show that the converse statement is also true: if S is a sub*set* of a group G, and if the relation $\sim$ defined by

$$g_1 \sim g_2 \quad \text{if } g_1^{-1} g_2 \in S$$

turns out to be an equivalence relation, then S is a subgroup of G.

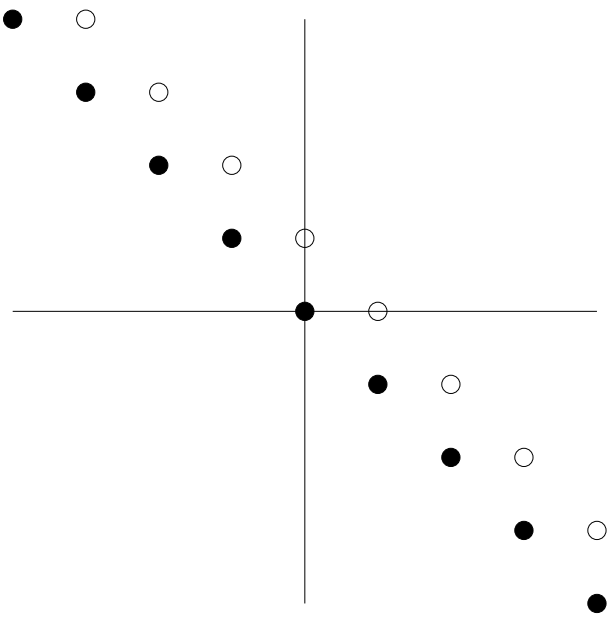
Exercise 7.15 Show that if G is a group and H a subgroup then the left cosets of H are equivalence classes under the relation

$$g_1 \sim g_2 \quad \text{if } g_1^{-1} g_2 \in H,$$

and the right cosets are equivalence classes under the relation

$$g_1 \sim g_2 \quad \text{if } g_1 g_2^{-1} \in H.$$

Example 7.16 1. Consider the quotient of $\mathbb{Z} \times \mathbb{Z}$ by the subgroup H generated by the element $(1, -1)$.



$H = \langle (1,-1) \rangle$ and the coset $(1,0) + H$.

The figure shows the subgroup H and just one of its cosets, $(1,0)+H$. As is obvious from the definition, $(0,1) \in (1,0)+H$, since $(1,0)-(0,1) \in H$; that is, $(1,0)+H = (0,1)+H$.

Now since $\mathbb{Z}\times\mathbb{Z}$ is generated by $(1,0)$ and $(0,1)$, it follows that $(\mathbb{Z}\times\mathbb{Z})/H$ is generated by $(1,0)+H$ and $(0,1)+H$; for since an arbitrary element (m,n) of $\mathbb{Z}\times\mathbb{Z}$ can be obtained as a sum

$$(m,n) = m\cdot(1,0) + n\cdot(0,1)$$

(where $m\cdot(1,0)$ means $(1,0)+\cdots+(1,0)$ (m times)), it follows that

$$(m,n)+H = (m\cdot(1,0)+n\cdot(0,1))+H$$

$$= m\cdot((1,0)+H))+n\cdots((0,1)+H)\,.$$

However, since $(1,0)+H$ and $(0,1)+H$ are the same, the last expression can be rewritten as

$$(m,n)+H = (m+n)\cdot(1,0)+H,$$

or equally well as

$$(m,n)+H = (m+n)\cdot(0,1)+H.$$

Thus, $(1,0)+H$ on its own generates $(\mathbb{Z}\times\mathbb{Z})/H$ (and so does $(0,1)+H$); $(\mathbb{Z}\times\mathbb{Z})/H$ is a cyclic group.

This fact can be appreciated visually; it is clear in the figure that the translates of H to the right and left by multiples of $(1,0)$ fill out the whole integer lattice. The translate by $(1,0)$ is of course the coset of $(1,0)$, shown in the figure.

Being cyclic, $(\mathbb{Z}\times\mathbb{Z})/H$ is isomorphic either to $\mathbb{Z}_n$ for some n or to $\mathbb{Z}$ itself. Which is it?

Exercise:

(i) Draw the cosets of $(2,0), (3,0)$ and $(-1,0)$.

(ii) Draw the coset which is the sum of the cosets

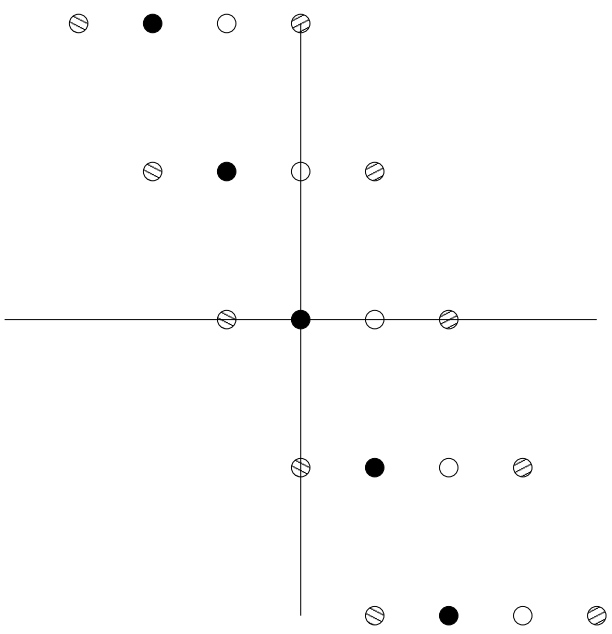
(a) $(1,0)+H$ and $(1,0)+H$;

(b) $(2,0)+H$ and $(3,0)+H$;

(c) $(2,0)+H$ and $(-2,0)+H$.

(iii) Construct an isomorphism $\mathbb{Z}\to(\mathbb{Z}\times\mathbb{Z})/H$.

2. $G=\mathbb{Z}\times\mathbb{Z}, H=\langle(1,-2)\rangle$.



$H = \langle(1,-2)\rangle$ and the cosets of $(1,0)$, of $(2,0)$ and of $(-1,0)$.

From the figure, it is clear that H and its translates by multiples of $(1,0)$ (i.e. the cosets $(n,0)+H$) do not fill out the integer lattice. That is, the quotient $(\mathbb{Z}\times\mathbb{Z})/H$ is not generated by $(1,0)+H$. Does this mean that it is not cyclic?

Exercise Show that $(\mathbb{Z}\times\mathbb{Z})/\langle(1,-2)\rangle$ *is* cyclic, and construct an explicit isomorphism to it from a "standard" cyclic group ($\mathbb{Z}$ or $\mathbb{Z}_n$ for some n).

3. **Exercise:** If $G=\mathbb{Z}\times\mathbb{Z},\ H=\langle(2,0)\rangle$, is G/H cyclic? If not, what is it?
4. **Exercise** Ditto for $H=\langle(2,0),(0,2)\rangle$.
5. **Exercise** Ditto for $H=\langle(2,0),(0,3)\rangle$.
6. **Exercise:** Suppose $G=\mathbb{Z}\times\mathbb{Z},\ H=\langle(2,2)\rangle$.

   (i) Make a drawing showing H and the cosets of $(1,0)$, $(2,0)$ and $(0,1)$.

   (ii) Construct an explicit isomorphism $\mathbb{Z}\times\mathbb{Z}_n\to G/H$ for some n.
7. **Exercise:** Same as part (ii) of previous exercise except with $(2,n)$ in place of $(2,2)$.
8. **Exercise:** Same as previous exercise except with (m,n) in place of $(2,n)$.
9. **Exercise:** Let $H\subset\mathbb{Z}\times\mathbb{Z}$ be generated by $(-2,1)$ and $(1,-2)$.

   (i) Draw H and the coset of $(1,0)$ on the integer lattice. Be careful! The coset of $(1,0)$ is not confined to a single line as in the previous examples.

   (ii) Find an explicit isomorphism between $(\mathbb{Z}\times\mathbb{Z})/H$ and one of the standard cyclic groups.

10. **Exercise:** (Harder) The quotient $(\mathbb{Z}\times\mathbb{Z}\times\mathbb{Z})/H$, where H is the subgroup generated by $(-2,1,0),(1,-2,1)$ and $(0,1,-2)$, is cyclic. Prove this, and identify the quotient group.

11. **Exercise:** Make up your own exercise along the lines of some of the previous ones.

The index of a subgroup

Definition 7.17 *Let G be a group and H a subgroup. We define the* index *of H in G, which is denoted by $[G:H]$, to be the number of distinct left cosets of H in G, i.e. the number of elements of G/H.*

When H is not normal, it is not obvious, *a priori*, that the number of left cosets of H in G is equal to the number of right cosets, since the left and right cosets are different. However:

Proposition 7.18 *There is a natural bijection between the set of left cosets and the set of right cosets of any subgroup.*

Proof Given a left coset gH, consider the set

$$S = \{x^{-1} : x \in gH\}.$$

Since $gH = \{gh : h \in H\}$, it follows that $S = \{(gh)^{-1} : h \in H\} = \{h^{-1}g^{-1} : h \in H\}$. Now the set $\{h^{-1} : h \in H\}$ is equal to H itself (inversion defines a bijection $H \to H$); thus S is the right coset Hg^{-1}.

This suggests that we define a map ϕ from the set of left cosets to the set of right cosets, by $\phi(gH) = Hg^{-1}$. I leave as an **Exercise** the proof that this map is well defined, is injective and is surjective. □

Corollary 7.19 $[G:H]$ = *the number of left cosets of H in G = the number of right cosets of H in G.*

Example 7.20 When $|G|$ is finite, Lagrange's Theorem tells us that $[G:H] = |G|/|H|$, but $[G:H]$ may be finite even when $|G|$ is not.

1. $G = Gl(n,\mathbb{R})$, $H = \{A \in Gl(n,\mathbb{R}) : \det A > 0\}$. Then H has index 2. For if A_1 and A_2 both have negative determinant, then $A_1^{-1}A_2$ has positive determinant, i.e. is in H. That is, any two elements not in H are equivalent under the relation $\sim$, and thus lie in the same left coset of H. Hence, there are just two cosets, H and $Gl(n,\mathbb{R}) \setminus H$.

2. **Exercise** What is the index of the subgroup $n\mathbb{Z}$ in $\mathbb{Z}$?

3. Recall that $Sl(n,\mathbb{R})$ is the subgroup of $Gl(n,\mathbb{R})$ consisting of matrices with determinant 1. What is $[Gl(n,\mathbb{R}) : Sl(n,\mathbb{R})]$? Is it finite? I claim that it is not. This is not hard to see: the elements A and B in $Gl(n,\mathbb{R})$ have the same coset if and only if $A^{-1}B \in Sl(n,\mathbb{R})$. But
$$A^{-1}B \in Sl(n,\mathbb{R}) \Longleftrightarrow \det A^{-1}B = 1 \Longleftrightarrow \det A = \det B;$$
as there are matrices in $Gl(n,\mathbb{R})$ with determinant of every non-zero real value, there is in fact a *bijection*
$$Gl(n,\mathbb{R})/Sl(n,\mathbb{R}) \to \mathbb{R}^{\times}.$$
Is $Sl(n,\mathbb{R})$ a normal subgroup of $Gl(n,\mathbb{R})$?

We now return to Example 7.3(3).

Proposition 7.21 *If $[G : H] = 2$ then H is normal in G.*

Proof There are two left cosets, H and gH (where g is any element not in H). Together they make up all of G. As they are disjoint, it follows that for $g \notin H$, $gH = G \setminus H$; by a similar argument $Hg = G \setminus H$. Thus for $g \notin H$, $gH = Hg$. If $g \in H$, then $gH = H = Hg$. In either case $gH = Hg$, so H is normal. □

Exercise Suppose H is a subgroup of G and K is a subgroup of H. Then K is a subgroup of G; show that $[G : K] = [G : H][H : K]$.

8 The First Isomorphism Theorem

In Example 6.2(2), we looked at the homomorphism $\mathbb{Z} \to \mathbb{Z}_n$ defined by sending m to its remainder on division by n. In terms of the notation we have since adopted, this map can be described as

$$\phi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$$
$$\phi(m) = m + n\mathbb{Z}.$$

Given a group G and a normal subgroup N, such a homomorphism always exists:

$$q : G \to G/N$$
$$q(g) = gN.$$

The kernel of q is N itself, and q is obviously surjective. I leave it to you to check that it is a homomorphism. We use the letter q because this map is the "passage to the quotient".

Under what circumstances can we construct a homomorphism *from* G/N to some other group? A partial answer is given by

Theorem 8.1 *(*The First Isomorphism Theorem*) Suppose that $\phi : G_1 \to G_2$ is a homomorphism of groups. Let $N = ker(\phi)$. Then ϕ gives rise to an isomorphism*

$$\bar{\phi} : G_1/N \ \to Im(\phi),$$

defined by

$$\bar{\phi}(gN) = \phi(g).$$

Proof By 7.2, N is a normal subgroup of G, so at least G/N is a group. As usual with quotient groups, the first thing to check is that the map $\bar{\phi}$ is well-defined. In view of the putative definition given in the statement of the theorem, this amounts to checking that if $g_1N = g_2N$ then $\phi(g_1) = \phi(g_2)$. But

$$g_1N = g_2N \iff g_1^{-1}g_2 \in N \iff \phi(g_1^{-1}g_2) = e_{G_2} \iff \phi(g_1^{-1})\phi(g_2) = e_{G_2}$$

$$\iff \phi(g_1)^{-1}\phi(g_2) = e_{G_2} \iff \phi(g_2) = \phi(g_1),$$

and so indeed $\bar{\phi}$ is well defined. In fact this reasoning has also shown that $\bar{\phi}$ is injective (just follow the backward-pointing halves of the equivalence signs), so it remains only to show that

$\bar{\phi}$ is surjective. But this is completely obvious: if $g_2 \in \mathrm{Im}(\phi)$, say $g_2 = \phi(g_1)$, and then $g_2 = \bar{\phi}(g_1 N) \in \mathrm{Im}(\bar{\phi})$.

So $\bar{\phi}$ is a well-defined bijection. We have to show that it is also a homomorphism. But this is very easy:

$$\bar{\phi}(g_1 N \cdot g_2 N) = \bar{\phi}((g_1 g_2)N) = \phi(g_1 g_2) = \phi(g_1)\phi(g_2) = \bar{\phi}(g_1 N)\bar{\phi}(g_2 N).$$

□

Exercise Write out this argument again, using the notation $\overline{g_1}$ and $\overline{g_2}$ in place of $g_1 N$ and $g_2 N$. You might find it easier to remember in this form.

Example 8.2 1. In Example 7.20 we calculated $[Gl(n,\mathbb{R}) : Sl(n,\mathbb{R})]$ and decided that it was infinite. We can see this even more clearly now: the (evidently surjective) homomorphism $\det : Gl(n,\mathbb{R}) \to \mathbb{R}^\times$ passes to the quotient to give an isomorphism (and in particular a bijection)

$$Gl(n,\mathbb{R})/Sl(n,\mathbb{R}) \to \mathbb{R}^\times.$$

2. In Example 6.2(5) we looked at the (surjective) homomorphism $\exp : \mathbb{R} \to S^1$, defined by $\exp(t) = e^{it}$. Now kerexp is the subgroup $2\pi\mathbb{Z} = \{0, \pm 2\pi, \pm 4\pi, \cdots\}$ of $\mathbb{R}$, and the First Isomorphism Theorem tells us that exp passes to the quotient to define an isomorphism

$$\mathbb{R}/2\pi\mathbb{Z} \to S^1.$$

This casts more light on an example that may have been slightly confusing. Recall from Example 6.2(4) that we suggested two ways of thinking of the argument of a non-zero complex number:

(i) $\arg(z) = z/|z| \in S^1$, and

(ii) $\arg(z) \in \mathbb{R}/\sim$, where the relation $\sim$ is defined by $a \sim b$ if $a - b \in 2\pi\mathbb{Z}$.

Using the first interpretation of arg, we reached the slightly disconcerting conclusion that

$$\arg(z_1 z_2) = \arg(z_1)\arg(z_2)$$

instead of the more usual version that the argument of the product of two complex numbers is the *sum* of the arguments of the two numbers. The second interpretation should now be clearer than when it was introduced: $\mathbb{R}/\sim$ is just the quotient of $\mathbb{R}$ by the subgroup $2\pi\mathbb{Z}$. I reiterate that this is a natural habitat for *angles*, since angles differing by a multiple of 2π are the same. Moreover, as a quotient of the group $\mathbb{R}$ by a subgroup (which is automatically normal because $\mathbb{R}$ is abelian), $\mathbb{R}/2\pi\mathbb{Z}$ is itself a group, and we can add angles. Needless to say, in this interpretation of arg, we *do* have $\arg(z_1 z_2) = \arg(z_1) + \arg(z_2)$. The two versions of arg are easily reconciled using the isomorphism

$$\overline{\exp} : \mathbb{R}/2\pi\mathbb{Z} \to S^1$$

from the group $\mathbb{R}/2\pi\mathbb{Z}$ where the binary operation is $+$ to the group S^1, where the binary operation is $\times$. We have a commutative diagram

$$\begin{array}{ccc} & \mathbb{C}^\times & \\ \swarrow & & \searrow \\ \mathbb{R}/2\pi\mathbb{Z} & \xrightarrow{\overline{exp}} & S^1 \end{array}$$

in which the two downward arrows are the two versions of the map arg.

Appendix: Equivalence Relations in other parts of Mathematics

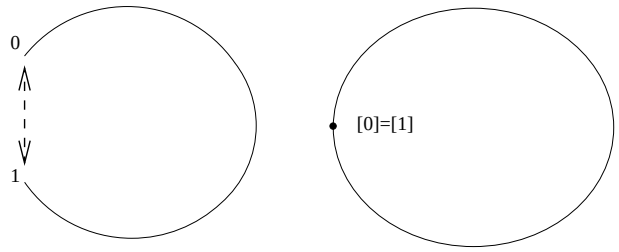
The idea of taking the quotient of a set by an equivalence relation is widely used not just in group theory, but in practically all other branches of mathematics. In this appendix, we look at one of them, in the hope that this will help to cement the notion of the quotient of a set by an equivalence relation. If this material, which is explained in a necessarily imprecise way, only seems to increase your burden, ignore it.

Topologists use equivalence relations to make mathematical the idea of gluing things together. For example, take the unit interval $[0,1]$ in $\mathbb{R}$ and define an equivalence relation on it by

$$x_1 \sim x_2 \text{ if } \begin{cases} x_1 = x_2 & \text{or} \\ x_1 = 1 \text{ and } x_2 = 0 & \text{or} \\ x_1 = 0 \text{ and } x_2 = 1. \end{cases}$$

In other words, the end points of the interval are declared equivalent to one another, and all the interior points are equivalent only to themselves.

In the quotient $[0,1]/\sim$, equivalent points are identified with one another (regarded as the same), and thus we can think of $[0,1]/\sim$ as a circle, or perhaps better as a loop, since it has no particular shape:



Gluing its endpoints, the interval becomes a loop

(ii) Take the unit square $X = [0,1] \times [0,1]$ in the plane, and define an equivalence relation on it by

$$(x_1, y_1) \sim (x_2, y_2) \text{ if } \begin{cases} (x_1, y_1) = (x_2, y_2) & \text{or} \\ x_1 = 0, x_2 = 1 \text{ and } y_1 = y_2 & \text{or} \\ x_2 = 0, x_1 = 1 \text{ and } y_1 = y_2. \end{cases}$$

In other words, points on the left and right edges are declared equivalent to one another if they have the same y coordinate, and all other points are equivalent only to themselves.

Exercise: Sketch the quotient space $X/\sim$.

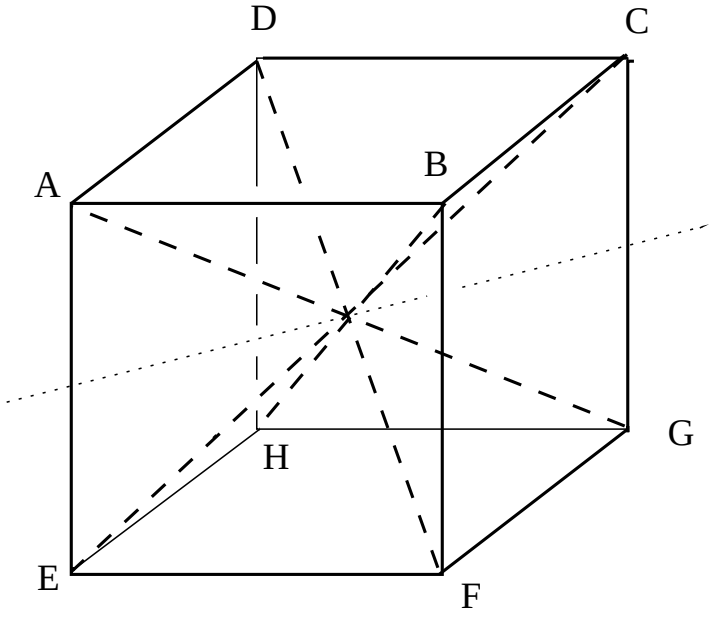
(iii) X is the unit square as before, and the equivalence relation of (ii) is augmented by declaring points on the upper and lower edges which have the same x coordinate to be equivalent to one another.

Exercise: Sketch the quotient space $X/\sim$.

Exercise: Consider the additive group $\mathbb{R} \times \mathbb{R}$ and its subgroup $\mathbb{Z} \times \mathbb{Z}$. The quotient $\mathbb{R} \times \mathbb{R}/\mathbb{Z} \times \mathbb{Z}$ can be sketched. What does it look like?

9 Isometries of the cube

This section is concerned with a wonderful example of many of the ideas we've been exploring.



The cube and its four diagonals

Theorem 9.1 *The group of isometries of the cube has 48 elements. The set of direct isometries (those with determinant $+1$) is a subgroup of index 2, and is isomorphic to the permutation group S_4.*

Proof Place the centre of the cube at $0 \in \mathbb{R}^3$; then the isometries of the cube are linear maps (cf. 2.8). Each has determinant equal to ± 1, and there are isometries of each type: for example, reflection in the plane midway between the faces $ABCD$ and $EFGH$ has determinant -1; the identity has determinant 1. Thus

$$\det : I(\text{Cube}) \to \mathbb{Z}_2$$

is a surjective homomorphism. Its kernel is a subgroup of index 2 (since $I(\text{Cube})/\ker(\det)$ is isomorphic to $\mathbb{Z}_2$, by the First Isomorphism Theorem 8.1).

Lemma 9.2 *Every isometry of $\mathbb{R}^3$ fixing 0, and with determinant $+1$, is a rotation about some axis (on which every point is fixed).*

Proof Let A be the matrix of such an isometry.

Step 1 A has an eigenvalue equal to $+1$.

For the characteristic equation

$$\det(A - \lambda I) = 0$$

is a cubic, and so has at least one real root. That is, A has at least one real eigenvalue λ. This eigenvalue must be ± 1, since by definition there is a non-zero vector v such that $Av = \lambda v$, and as A is an isometry it cannot change the length of v.

Suppose that A has an eigenvalue equal to -1. The product of the three (in general complex) roots of the characteristic equation is equal to $\det A$, i.e. to 1. The remaining two roots are either both real, or a pair of conjugate complex numbers (**Exercise**). In the latter case, their product is real and positive, and now multiplying by the eigenvalue we already had, -1, we get that the product of the three roots is negative, a contradiction. Hence, if there is an eigenvalue of -1, the remaining two roots of the characteristic equation must both be real. Each is equal to ± 1; they cannot both be equal to -1 as then once again the product of all three roots would be negative, contradicting the fact that $\det A = 1$. Hence, one of them must be equal to $+1$.

This completes the proof of Step 1.

Step 2 From Step 1, we know that A has a (pointwise) fixed axis L (the line generated by an eigenvector with associated eigenvalue 1). Let $L^\perp$ be the plane through 0 orthogonal to L. As A is an isometry, it must map $L^\perp$ to itself. The restriction of A to $L^\perp$ is an isometry, and thus (Corollary 2.7) is either a rotation or a reflection. In the latter case, with respect to a basis for $\mathbb{R}^3$ consisting of a vector e_1 in L, a vector e_2 in the line in $L^\perp$ fixed by the reflection, and a vector e_3 in $L^\perp$ at right angles to e_2, the matrix of A is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix};$$

and but then $\det A = -1$, a contradiction. Hence, A must be a rotation about the axis L, as claimed. □

Exercise Is every isometry of the cube with determinant -1 a reflection?

We now take a close look at ker(det), which in the light of the lemma we refer to as the "rotation subgroup" of $I(\text{Cube})$; we denote it by $I^+(\text{Cube})$.

The cube has eight vertices— four pairs of opposite vertices joined by diagonals passing through 0. Any isometry of the cube maps a pair of opposite vertices to some other pair of opposite vertices; it thus permutes the four diagonals. By labelling the four diagonals with the numbers $1, 2, 3, 4$, (I have not not shown this in the picture) we can identify these permutations with elements of S_4, and so obtain a map $I(\text{Cube}) \to S_4$.

Claim This map defines an isomorphism $\phi : I^+(\text{Cube}) \to S_4$.

That ϕ is a homomorphism is clear: as in Example5.7(3), the composite of the restrictions is the restriction of the composite. We have just to show that it is both 1-1 and onto.

To prove that it is onto, note that S_4 is generated by transpositions. If for each pair of diagonals we can find a rotation of the cube interchanging them, and mapping each of the other diagonals to itself, then we will have shown that all of the transpositions in S_4 lie in the image of ϕ. As the image of ϕ is a subgroup, it must be all of S_4, since S_4 is generated by the transpositions.

We pick the pair AG and EC of diagonals. In the picture there is a fine dotted line through the midpoints of the edges AE and CG. Rotation by π about this line interchanges AG and EC, and maps each of the other diagonals to itself.

There is nothing special about the pair we have considered — the same idea shows that we can interchange any pair of diagonals, while leaving all the rest where they are.

Thus, ϕ is onto.

To see that ϕ is 1-1, let's suppose that f is a rotation of the cube which maps each diagonal to itself — i.e. $f \in \ker(\phi)$. Consider the diagonal AG. Evidently, either f interchanges ts endpoints A and G, or it leaves them fixed.

Suppose first that f leaves A where it is. Then it must also leave fixed the endpoints of the other diagonals. For example, it must leave E fixed; if not, it would have to map E to the other endpoint of the diagonal EC, namely C, and this cannot happen as $AE < AC$.

It follows that f is the identity.

Suppose second that f maps A to G. Then by the argument of the previous paragraph, it can't leave *any* of the endpoints of the other diagonals fixed. So it must map E to C, D to F and H to B. There is only one isometry which does this: the map $-\mathrm{id}$ sending each point x to $-x$. But $-\mathrm{id}$ is not in $I^+(\mathrm{Cube})$, since it has determinant -1.

So the kernel of $\phi : I^+(\mathrm{Cube}) \to S_4$ is just $\{\mathrm{id}\}$, and ϕ is injective.

This completes the proof that $I^+(\mathrm{Cube})$ is isomorphic to S_4.

Finally, since $I^+(\mathrm{Cube})$ has index 2 in $I(\mathrm{Cube})$, the latter must have 48 elements. □

Exercise: Is $I(\mathrm{Cube})$ isomorphic to $S_4 \times \mathbb{Z}_2$? Both groups have 48 elements.

The following exercise is not really group theory, but geometry. Nevertheless I feel impelled to end these notes with it.

Exercise Draw a deformation (one might better say, a decoration) of the cube having $I^+(\mathrm{cube})$ as its full group of isometries. Recall from Section3(I) that the corresponding deformation in the case of the square was the figure on the left below. In the present context the "oriented square" drawn on the right, which has the same isometry group (equal to the group of rotations of the square) as the figure on the left, might be more helpful.