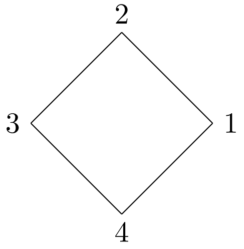
GROUP ACTIONS

KEITH CONRAD

1. Introduction

The symmetric group S_n behaves, by its very definition, as permutations of the set $\{1, 2, \ldots, n\}$. Dihedral groups D_n for $n \geq 3$, while interpreted geometrically as certain motions of the plane preserving a regular n-gon, can be considered as a group of permutations just of the n vertices; rigid motions of the vertices determine where the rest of the n-gon goes. If we label the n vertices in a definite manner by the numbers from 1 to n then we can view D_n as a subgroup of S_n.

Example 1.1. The labeling of the square below lets us regard the 90 degree counterclockwise rotation r in D_4 as the 4-cycle (1234) and the reflection s across the horizontal line bisecting the square as the transposition (24). The rest of the elements of D_4, as permutations of the vertices, are in the table below the square.



1	r	r^2	r^3	s	rs	r^2s	r^3s
(1)	(1234)	(13)(24)	(1432)	(24)	(12)(34)	(13)	(14)(23)

If we label the vertices in a different way (*e.g.*, swap the labels 1 and 2), then we turn D_4 into a different subgroup of S_4 (*e.g.*, swapping 1 and 2 turns r into $(2134) = (1342)$).

More abstractly, if we are given a set X (not necessarily the set of vertices of a square), then the set $\mathrm{Sym}(X)$ of all permutations of X is a group under composition, and the subgroup $\mathrm{Alt}(X)$ of even permutations of X is a group under composition. If we list the elements of X in a definite order, say as $X = \{x_1, \ldots, x_n\}$, then we can think about $\mathrm{Sym}(X)$ as S_n and $\mathrm{Alt}(X)$ as A_n, but a listing in a different order leads to different identifications of $\mathrm{Sym}(X)$ with S_n and $\mathrm{Alt}(X)$ with A_n.[1]

The "abstract" symmetric groups $\mathrm{Sym}(X)$ really do arise naturally:

Theorem 1.2 (Cayley). *Every finite group G can be embedded in a symmetric group.*

Proof. To each $g \in G$, define the left multiplication function $\ell_g \colon G \to G$, where $\ell_g(x) = gx$ for $x \in G$. Each ℓ_g is a permutation of G as a set, with inverse $\ell_{g^{-1}}$. So ℓ_g belongs to $\mathrm{Sym}(G)$. Since $\ell_{g_1} \circ \ell_{g_2} = \ell_{g_1 g_2}$ (that is, $g_1(g_2x) = (g_1g_2)x$ for all $x \in G$), associating to g

[1] When $X = \emptyset$, consider $\mathrm{Sym}(X)$ and $\mathrm{Alt}(X)$ to be trivial groups. The number of permutations of a set of size 0 is $0! = 1$.

the mapping ℓ_g gives a homomorphism of groups, $G \to \mathrm{Sym}(G)$. This homomorphism is one-to-one since ℓ_g determines g (after all, $\ell_g(e) = g$). Therefore the correspondence $g \mapsto \ell_g$ is an embedding of G as a subgroup of $\mathrm{Sym}(G)$. □

Allowing a group to behave as a permutations of a set, as in the proof of Cayley's theorem, is a very useful idea, and when this happens we say the group is acting on the set.

Definition 1.3. An *action* of a group G on a set X is the choice, for each $g \in G$, of a permutation $\pi_g \colon X \to X$ such that the following two conditions hold:

- π_e is the identity: $\pi_e(x) = x$ for each $x \in X$,
- for every g_1 and g_2 in G, $\pi_{g_1} \circ \pi_{g_2} = \pi_{g_1 g_2}$.

Example 1.4. The group S_n acts on $X = \{1, 2, \dots, n\}$ in the usual way: $\pi_\sigma(i) = \sigma(i)$ for all i. Then $\pi_1(i) = i$ for all $i \in X$ and $\pi_\sigma(\pi_{\sigma'}(i)) = \pi_\sigma(\sigma'(i)) = \sigma(\sigma'(i)) = (\sigma\sigma')(i) = \pi_{\sigma\sigma'}(i)$.

Example 1.5. Each group G acts on itself ($X = G$) by left multiplication functions. That is, we set $\pi_g \colon G \to G$ by $\pi_g(h) = gh$ for all $g \in G$ and $h \in G$. Then the conditions for being a group action are $eh = h$ for all $h \in G$ and $g_1(g_2 h) = (g_1 g_2)h$ for all $g_1, g_2, h \in G$, which are both true since e is an identity and multiplication in G is associative. (This is the idea behind Cayley's theorem.)

In practice, one dispenses with the notation π_g and writes $\pi_g(x)$ simply as $g \cdot x$ or gx. This is *not* meant to be an actual multiplication of elements from two possibly different sets G and X. It is just the notation for the effect of g (really, the permutation g is associated to) on the element x. In this notation, the axioms for a group action take the following form:

- for each $x \in X$, $e \cdot x = x$.
- for every $g_1, g_2 \in G$ and $x \in X$, $g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x$.

The basic idea in a group action is that the elements of a group are viewed as permutations of a set in such a way that composition of the corresponding permutations matches multiplication in the original group.

There are various types of notation that are used to express the idea "G acts on X", such as $G \circlearrowright X$ and $G \curvearrowright X$, but we will not use these here.

To get used to the notation, let's prove a simple result.

Theorem 1.6. *Let a group G act on the set X. If $x \in X$, $g \in G$, and $y = g \cdot x$, then $x = g^{-1} \cdot y$. If $x \neq x'$ then $g \cdot x \neq g \cdot x'$.*

Proof. From $y = g \cdot x$ we get $g^{-1} \cdot y = g^{-1} \cdot (g \cdot x) = (g^{-1}g) \cdot x = e \cdot x = x$. To show $x \neq x' \Longrightarrow gx \neq gx'$, we show the contrapositive: if $g \cdot x = g \cdot x'$ then applying g^{-1} to both sides gives $g^{-1} \cdot (g \cdot x) = g^{-1} \cdot (g \cdot x')$, so $(g^{-1}g) \cdot x = (g^{-1}g) \cdot x'$, so $x = x'$. □

Another way to think about an action of a group on a set is that it is a certain homomorphism. Here are the details.

Theorem 1.7. *Actions of the group G on the set X are the same as group homomorphisms from G to $\mathrm{Sym}(X)$, the group of permutations of X.*

Proof. Suppose we have an action of G on X. We view $g \cdot x$ as a function of x (with g fixed). That is, for each $g \in G$ we have a function $\pi_g \colon X \to X$ by $\pi_g(x) = g \cdot x$. The axiom

$$e \cdot x = x$$

says π_e is the identity function on X. The axiom

$$g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x$$

says $\pi_{g_1} \circ \pi_{g_2} = \pi_{g_1 g_2}$, so composition of functions on X corresponds to multiplication in G. Moreover, π_g is an invertible function since $\pi_{g^{-1}}$ is an inverse: the composite of π_g and $\pi_{g^{-1}}$ is π_e, which is the identity function on X. Therefore $\pi_g \in \mathrm{Sym}(X)$ and $g \mapsto \pi_g$ is a homomorphism $G \to \mathrm{Sym}(X)$.

Conversely, suppose we have a homomorphism $f\colon G \to \mathrm{Sym}(X)$. For each $g \in G$, we have a permutation $f(g)$ on X, and $f(g_1 g_2) = f(g_1) \circ f(g_2)$. Setting $g \cdot x = f(g)(x)$ defines a group action of G on X, since the homomorphism properties of f yield the defining properties of a group action. □

From this viewpoint, the set of $g \in G$ that act trivially ($g \cdot x = x$ for all $x \in X$) is simply the kernel of the homomorphism $G \to \mathrm{Sym}(X)$ associated to the action. Therefore those g that act trivially on X are said to lie in the *kernel* of the action.

We will not often use the interpretation of Theorem 1.7 before Section 6. Until then we take the more concrete viewpoint of a group action as a kind of product $g \cdot x$ of g with x, taking values in X subject to the properties $e \cdot x = x$ and $g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x$.

Here is an outline of later sections. Section 2 describes several concrete examples of group actions and also some general actions available for all groups. Section 3 describes the important orbit-stabilizer formula. The short Section 4 isolates an important fixed-point congruence for actions of p-groups. Sections 5 and 6 give applications of group actions to group theory. In Appendix A, group actions are used to derive three classical congruences from number theory.

2. Examples

Example 2.1. We can make $\mathbf{R}^n$ act on itself by translations: for $\mathbf{v} \in \mathbf{R}^n$, let $T_{\mathbf{v}}\colon \mathbf{R}^n \to \mathbf{R}^n$ by $T_{\mathbf{v}}(\mathbf{w}) = \mathbf{w} + \mathbf{v}$. The axioms for a group action are: $T_{\mathbf{0}}(\mathbf{w}) = \mathbf{w}$ and $T_{\mathbf{v}_1}(T_{\mathbf{v}_2}(\mathbf{w})) = T_{\mathbf{v}_1+\mathbf{v}_2}(\mathbf{w})$. These are true from properties of vector addition:

$$\mathbf{w} + \mathbf{0} = \mathbf{w}, \quad (\mathbf{w} + \mathbf{v}_2) + \mathbf{v}_1 = \mathbf{w} + (\mathbf{v}_1 + \mathbf{v}_2).$$

(This is a special case of Example 1.5 using additive notation.)

Example 2.2. Let G be the group of Rubik's cube: all sequences of motions on the cube (keeping center colors in fixed locations). This group acts on two different sets: the 8 corner cubelets and 12 edge cubelets. Or we could let G act on the set of all 20 non-centerface cubelets together.

Example 2.3. For $n \geq 3$, D_n acts on a regular n-gon as rigid motions. We can also view D_n as acting just on the n vertices of a regular n-gon. This does not lose information, since knowing where vertices go under a rigid motion determines where everything else goes. By focusing on the action of D_n on the n vertices, and labelling them by $1, 2, \ldots, n$ in some way, we make D_n act on $\{1, 2, \ldots, n\}$ (the case $n = 4$ is Example 1.1).

We can also make D_n act on the set of diagonals of the regular n-gon, since a rigid motion sends diagonals to diagonals.

Example 2.4. The group $\mathrm{GL}_n(\mathbf{R})$ acts on vectors in $\mathbf{R}^n$ in the usual way that a matrix can be multiplied with a (column) vector: $A \cdot \mathbf{v} = A\mathbf{v}$. In this action, the origin $\mathbf{0}$ is fixed by every A while other vectors get moved around (as A varies). The axioms of a group action are properties of matrix-vector multiplication: $I_n\mathbf{v} = \mathbf{v}$ and $A(B\mathbf{v}) = (AB)\mathbf{v}$.

Example 2.5. The group of affine-linear transformations $f : \mathbf{R}^n \to \mathbf{R}^n$, where $f(\mathbf{v}) = A\mathbf{v} + \mathbf{b}$ with $A \in \mathrm{GL}_n(\mathbf{R})$ and $\mathbf{b} \in \mathbf{R}^n$ (an invertible linear map plus a translation), acts on $\mathbf{R}^n$ by $f \cdot \mathbf{v} = f(\mathbf{v}) = A\mathbf{v} + \mathbf{b}$ for all $\mathbf{v} \in \mathbf{R}^n$. In this action, the origin $\mathbf{0}$ is moved to $\mathbf{b}$. The affine-linear mapping f knows what the A and $\mathbf{b}$ in its formula are: $\mathbf{b}$ is $f(\mathbf{0})$ and the columns of A are $f(\mathbf{e}_i) - f(\mathbf{0})$ where $1 \leq i \leq n$. Writing f as $f_{A,\mathbf{b}}$, the identity in the group is $f_{I_n,\mathbf{0}}$, multiplication is $f_{A,\mathbf{b}} \circ f_{A',\mathbf{b}'} = f_{AA',A\mathbf{b}'+\mathbf{b}}$ (compare both sides at each $\mathbf{v}$), and $f_{A,\mathbf{b}}^{-1} = f_{A^{-1},-A^{-1}\mathbf{b}}$.

The axioms of a group action in this setting say $f_{A,\mathbf{b}} \cdot (f_{A',\mathbf{b}'} \cdot \mathbf{v}) = (f_{A,\mathbf{b}} \circ f_{A',\mathbf{b}'}) \cdot \mathbf{v}$ for all $\mathbf{v} \in \mathbf{R}^n$, which the reader should check do hold.

Example 2.6. Let G be the group of affine-linear transformations on $(\mathbf{Z}/2\mathbf{Z})^2$, meaning maps $f : (\mathbf{Z}/2\mathbf{Z})^2 \to (\mathbf{Z}/2\mathbf{Z})^2$ where $f(\mathbf{v}) = A\mathbf{v}+\mathbf{b}$, with $A \in \mathrm{GL}_2(\mathbf{Z}/2\mathbf{Z})$ and $\mathbf{b} \in (\mathbf{Z}/2\mathbf{Z})^2$. Each f knows its A and $\mathbf{b}$: $\mathbf{b} = f(\mathbf{0})$ and A has columns $f(\binom{1}{0}) - \mathbf{b}$ and $f(\binom{0}{1}) - \mathbf{b}$. By counting A's and $\mathbf{b}$'s, $|G| = |\mathrm{GL}_2(\mathbf{Z}/2\mathbf{Z})||(\mathbf{Z}/2\mathbf{Z})^2| = (6)(4) = 24$. Is $G \cong S_4$?

Since each element of G is an invertible function $(\mathbf{Z}/2\mathbf{Z})^2 \to (\mathbf{Z}/2\mathbf{Z})^2$, G acts on $(\mathbf{Z}/2\mathbf{Z})^2$. (This is like the previous example, with $\mathbf{R}^n$ replaced by $(\mathbf{Z}/2\mathbf{Z})^2$.) Each element of G is completely determined by its effect on $(\mathbf{Z}/2\mathbf{Z})^2$, which has size 4, so the action of G on $(\mathbf{Z}/2\mathbf{Z})^2$ gives us an *injective* homomorphism $G \to \mathrm{Sym}((\mathbf{Z}/2\mathbf{Z})^2) \cong S_4$. Since $|G| = 24$, necessarily this homomorphism is an isomorphism, so $G \cong S_4$.

To interpret elements of G as permutations in S_4, let's label the elements of $(\mathbf{Z}/2\mathbf{Z})^2$ by $1, 2, 3, 4$, say $\binom{0}{0} \leftrightarrow 1$, $\binom{1}{0} \leftrightarrow 2$, $\binom{0}{1} \leftrightarrow 3$, and $\binom{1}{1} \leftrightarrow 4$. To realize the permutation (12) in G means finding an affine linear map f on $(\mathbf{Z}/2\mathbf{Z})^2$ that swaps $\binom{0}{0}$ with $\binom{1}{0}$ and fixes $\binom{0}{1}$ and $\binom{1}{1}$ (that f fixes one of $\binom{0}{1}$ and $\binom{1}{1}$ forces f to fix the other one since there's nowhere left for the other vector to go). Writing $f(\mathbf{v}) = A\mathbf{v} + \mathbf{b}$, we need $f\binom{0}{0} = \binom{1}{0}$ and $f\binom{1}{0} = \binom{0}{0}$, so $\mathbf{b} = \binom{1}{0}$ and $A\binom{1}{0} + \mathbf{b} = \binom{0}{0}$. Thus $A\binom{1}{0} = -\mathbf{b} = \mathbf{b} = \binom{1}{0}$: the first column of A is $\binom{1}{0}$. That $f\binom{0}{1} = \binom{0}{1}$ says $A\binom{0}{1} + \binom{1}{0} = \binom{0}{1}$, so $A\binom{0}{1} = \binom{0}{1} - \binom{1}{0} = \binom{1}{1}$: the second column of A is $\binom{1}{1}$. Thus $f(\mathbf{v}) = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)\mathbf{v} + \binom{1}{0}$.

Check you can derive on your own that the affine-linear maps realizing (13) and (14) in G are $f(\mathbf{v}) = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)\mathbf{v} + \binom{0}{1}$ and $f(\mathbf{v}) = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\mathbf{v} + \binom{1}{1}$. Don't merely check those work: be able to derive these expressions for f in each case.

Example 2.7. The laws of motion in physics to an observer should be the same at every location, at every time, in every direction, and when traveling in a fixed direction at a fixed speed. All of these conditions under which the laws of physics should not change can be described by the action on $\mathbf{R}^4$ (spacetime) of a 10-dimensional group, both in relativistic and non-relativistic settings. See Appendix B for more details.

Example 2.8. The group S_n acts on polynomials $f(T_1, \ldots, T_n)$, by permuting the variables:

$$(\sigma \cdot f)(T_1, \ldots, T_n) = f(T_{\sigma(1)}, \ldots, T_{\sigma(n)}). \tag{2.1}$$

The effect is to replace T_i everywhere in $f(T_1, \ldots, T_n)$ by $T_{\sigma(i)}$. For example, $(12)(23) = (123)$ in S_3 and $(12)\cdot((23)\cdot(T_2+T_3^2)) = (12)\cdot(T_3+T_2^2) = T_3+T_1^2$ and $(123)\cdot(T_2+T_3^2) = T_3+T_1^2$, giving the same result both ways.

It's obvious that $(1) \cdot f = f$. To check that $\sigma \cdot (\sigma' \cdot f) = (\sigma\sigma') \cdot f$ for all σ and σ' in S_n, so (2.1) is a group action of S_n on polynomials in n variables, $\sigma' \cdot f$ replaces each T_i in f with $T_{\sigma'(i)}$. Applying σ to a polynomial replaces each T_j in it with $T_{\sigma(j)}$, so it replaces each

$T_{\sigma'(i)}$ with $T_{\sigma(\sigma'(i))}$. Therefore applying σ' and then σ has the effect

$$f(T_1,\dots,T_n) \overset{\sigma'}{\mapsto} f(T_{\sigma'(1)},\dots,T_{\sigma'(n)}) \overset{\sigma}{\mapsto} f(T_{\sigma(\sigma'(1))},\dots,T_{\sigma(\sigma'(n))}).$$

The last expression is $f(T_{(\sigma\sigma')(1)},\dots,T_{(\sigma\sigma')(n)})$, which is $\sigma\sigma'\cdot f$, so $\sigma\cdot(\sigma'\cdot f) = (\sigma\sigma')\cdot f$.

Since f and $\sigma\cdot f$ have the same degree, and if f is homogeneous then $\sigma\cdot f$ is homogeneous, this action of S_n on polynomials in n variables can be restricted to the polynomials in n variables with a fixed degree or the homogeneous polynomials in n variables with a fixed degree. An example is S_n acting on homogeneous linear polynomials $\{a_1T_1+\cdots+a_nT_n\}$, where

$$\sigma\cdot(c_1T_1+\cdots+c_nT_n) = c_1T_{\sigma(1)}+\cdots+c_nT_{\sigma(n)} = c_{\sigma^{-1}(1)}T_1+\cdots+c_{\sigma^{-1}(n)}T_n. \tag{2.2}$$

Lagrange's study of the group action in Example 2.8 (*ca.* 1770) marked the first systematic use of symmetric groups in algebra. Lagrange wanted to understand why nobody had found an analogue of the quadratic formula for roots of a polynomial of degree greater than four. He was not completely successful, although he found in this group action that there are some different features in the cases $n \le 4$ and $n = 5$.

Example 2.9. Here is a tricky example, so pay attention. Let S_n act on $\mathbf{R}^n$ by permuting the coordinates: for $\sigma \in S_n$ and $v = (c_1,\dots,c_n) \in \mathbf{R}^n$, set $\pi_\sigma(v) = (c_{\sigma(1)},\dots,c_{\sigma(n)})$.

For example, let $n = 3$, $\sigma = (12)$, $\sigma' = (23)$, and $v = (5,7,9)$. Then

$$\pi_\sigma(\pi_{\sigma'}(v)) = \pi_{(12)}(\pi_{(23)}(5,7,9)) = \pi_{(12)}(5,9,7) = (9,5,7)$$

and

$$\pi_{\sigma\sigma'}(5,7,9) = \pi_{(123)}(5,7,9) = (7,9,5),$$

which do not agree, so sending v to $\pi_\sigma(v)$ in $\mathbf{R}^n$ is *not* a group action of S_n on $\mathbf{R}^n$.

A peculiar thing happens if we calculate $\pi_\sigma(\pi_{\sigma'}(v))$ and $\pi_{\sigma\sigma'}(v)$ in general to see what is going wrong, since it is easy to convince ourselves that we do have a group action:

$$\begin{aligned}\pi_\sigma(\pi_{\sigma'}(c_1,\dots,c_n)) &= \pi_\sigma(c_{\sigma'(1)},\dots,c_{\sigma'(n)})\\ &= (c_{\sigma(\sigma'(1))},\dots,c_{\sigma(\sigma'(n))})\\ &= (c_{(\sigma\sigma')(1)},\dots,c_{(\sigma\sigma'(n)})\\ &= \pi_{\sigma\sigma'}(c_1,\dots,c_n),\end{aligned}$$

which suggests $\pi_\sigma\circ\pi_{\sigma'} = \pi_{\sigma\sigma'}$, and that is not what we saw in the numerical example above. What happened?!? The mistake really is in the general calculation, not the example. Try to find the error before reading further.

The mistake was in the second line, when we computed $\pi_\sigma(c_{\sigma'(1)},\dots,c_{\sigma'(n)})$ by applying σ to the indices $\sigma'(i)$. A vector does not remember the indices of its coordinates after they are permuted: to compute $\pi_{(12)}(\pi_{(23)}(5,7,9)) = \pi_{(12)}(5,9,7)$, the next step treats $(5,9,7)$ as a new vector with coordinates indexed by 1, 2, 3 in that order even though the coordinate order has changed from the original $(5,7,9)$. The computation of $\pi_\sigma(v)$ always needs coordinate indices for v to run from 1 to n in that order. Thus when computing $\pi_{(12)}(\pi_{(23)}(c_1,c_2,c_3)) = \pi_{(12)}(c_1,c_3,c_2)$, in the next step write $(c_1,c_3,c_2) = (d_1,d_2,d_3)$. Then

$$\pi_{(12)}(\pi_{(23)}(c_1,c_2,c_3)) = \pi_{(12)}(c_1,c_3,c_2) = \pi_{(12)}(d_1,d_2,d_3) = (d_2,d_1,d_3) = (c_3,c_1,c_2),$$

which does not agree with

$$\pi_{(12)(23)}(c_1,c_2,c_3) = \pi_{(123)}(c_1,c_2,c_3) = (c_2,c_3,c_1).$$

In general, for σ and σ' in S_n, and $v = (c_1, \dots, c_n)$ in $\mathbf{R}^n$,

$$\begin{aligned}\pi_\sigma(\pi_{\sigma'}(v)) &= \pi_\sigma(c_{\sigma'(1)}, \dots, c_{\sigma'(n)}) \\ &= \pi_\sigma(d_1, \dots, d_n) \text{ where } d_i = c_{\sigma'(i)} \\ &= (d_{\sigma(1)}, \dots, d_{\sigma(n)}) \\ &= (c_{\sigma'(\sigma(1))}, \dots, c_{\sigma'(\sigma(n))}) \\ &= (c_{(\sigma'\sigma)(1)}, \dots, c_{(\sigma'\sigma)(n)}) \\ &= \pi_{\sigma'\sigma}(v),\end{aligned}$$

so $\pi_\sigma \circ \pi_{\sigma'}$ is $\pi_{\sigma'\sigma}$, which is not $\pi_{\sigma\sigma'}$ on $\mathbf{R}^n$ if $\sigma'\sigma \neq \sigma\sigma'$ (*e.g.*, $n \geq 3$, $\sigma = (12)$, $\sigma' = (23)$).

A way to explain why $\pi_\sigma \circ \pi_{\sigma'} = \pi_{\sigma'\sigma}$ without the trick of rewriting coordinates with another letter is to express the formula $\pi_\sigma((c_1, \dots, c_n)) = (c_{\sigma(1)}, \dots, c_{\sigma(n)})$ as $(\pi_\sigma(v))_i = v_{\sigma(i)}$ for $i = 1, \dots, n$ (*e.g.*, when $v = (c_1, \dots, c_n)$ and $i = 1$, $(\pi_\sigma(v))_i$ and $v_{\sigma(i)}$ are both $c_{\sigma(1)}$). Then for all $v \in \mathbf{R}^n$ and $i = 1, \dots, n$,

$$(\pi_\sigma(\pi_{\sigma'}(v)))_i = (\pi_{\sigma'}(v))_{\sigma(i)} = v_{\sigma'(\sigma i)} = v_{(\sigma'\sigma)(i)} = (\pi_{\sigma'\sigma}(v))_i,$$

so $\pi_\sigma \circ \pi_{\sigma'} = \pi_{\sigma'\sigma}$ on $\mathbf{R}^n$.

To have an honest group action here, *redefine* the effect of S_n on $\mathbf{R}^n$ using inverses in S_n: set $\sigma \cdot v = (c_{\sigma^{-1}(1)}, \dots, c_{\sigma^{-1}(n)})$, or equivalently $(\sigma \cdot v)_i = v_{\sigma^{-1}(i)}$ for all i. Then $\sigma \cdot (\sigma' \cdot v) = (\sigma\sigma') \cdot v$ and we have a group action of S_n on $\mathbf{R}^n$, which in fact is essentially the action of S_n from the previous example on homogeneous linear polynomials (see (2.2)). Indeed, if $e_1, \dots, e_n$ is the standard basis of $\mathbf{R}^n$ and $v = \sum_{i=1}^n c_i e_i$ then

$$\sigma \cdot \sum_{i=1}^n c_i e_i = (c_{\sigma^{-1}(1)}, \dots, c_{\sigma^{-1}(n)}) = \sum_{i=1}^n c_{\sigma^{-1}(i)} e_i = \sum_{i=1}^n c_i e_{\sigma(i)},$$

which is how (2.2) looks with T_i in place of e_i. In other words, the action of each $\sigma \in S_n$ on $\mathbf{R}^n$ is $\mathbf{R}$-linear and permutes the basis vectors $\{e_i\}$ (not the coefficients!) in the same way it permutes the indices: $\sigma(e_i) = e_{\sigma(i)}$.

The lesson from these last two examples is that when S_n permutes variables in a polynomial then it acts "directly", but when S_n permutes coordinates in a vector then it has to act using inverses. When S_n acts on variables or coordinates, it acts without inverses in one case and with inverses in the other case, but it's easy to forget which case is which. At least remember that you need to be careful.

Example 2.10. Let G be a group acting on the set X, and S be a set. Write $\mathrm{Map}(X, S)$ for the set of all functions $f\colon X \to S$. It is natural to try defining an action of G on the set $\mathrm{Map}(X, S)$ by the rule

$$(\pi_g f)(x) = f(gx), \tag{2.3}$$

where gx is the action of $g \in G$ on $x \in X$. While $\pi_g f$ is a function $X \to S$, sending each f to $\pi_g f$ *is usually not an action* of G on $\mathrm{Map}(X, S)$ even though it is easy to confuse yourself into thinking it is: for g and h in G, and $x \in X$,

$$\pi_g((\pi_h f)(x)) = \pi_g(f(hx)) = f(g(hx)) = f((gh)x) = (\pi_{gh} f)(x).$$

This holds for all $x \in X$, so $\pi_g(\pi_h f) = \pi_{gh} f$, right? No. The calculation above is bogus since the first expression $\pi_g((\pi_h f)(x))$ is nonsense: $(\pi_h f)(x) = f(hx)$ belongs to S and no action of G has been defined on S, so $\pi_g(f(hx))$ isn't defined. Even if G acts on S, π_g is applied to *functions* $X \to S$, not to elements of S. The mistake was confusing $(\pi_g f)(x)$ in (2.3) with the meaningless expression $\pi_g(f(x))$.

A correct calculation is

$$(\pi_g(\pi_h f))(x) = (\pi_h f)(gx) = f(h(gx)) = f((hg)x) = (\pi_{hg} f)(x).$$

Therefore $\pi_g(\pi_h f) = \pi_{hg} f$. This flipping in the indices is similar to the false action in Example 2.9. To get a group action of G on $\mathrm{Map}(X, S)$ when G already acts on X (from the left), replace g with g^{-1} in (2.3): set

$$(g \cdot f)(x) = f(g^{-1}x).$$

Now

$$(g \cdot (h \cdot f))(x) = (h \cdot f)(g^{-1}x) = f(h^{-1}(g^{-1}x)) = f((gh)^{-1}x) = ((gh) \cdot f)(x),$$

so $g \cdot (h \cdot f) = (gh) \cdot f$. This is a group action of G on $\mathrm{Map}(X, S)$.

If G is S_n, X is $\{1, \ldots, n\}$ with its natural S_n-action, and $S = \mathbf{R}$, then $\mathrm{Map}(X, S) = \mathbf{R}^n$: writing down a vector $v = (c_1, \ldots, c_n)$ amounts to listing the coordinates in order, and a list of coordinates in order is a function $f\colon \{1, 2, \ldots, n\} \to \mathbf{R}$ where $f(i) = c_i$. The definition $(g \cdot f)(i) = f(g^{-1}i)$ amounts to saying $g \cdot (c_1, \ldots, c_n) = (c_{g^{-1}(1)}, \ldots, c_{g^{-1}(n)})$, which is precisely the valid action of S_n on $\mathbf{R}^n$ at the end of Example 2.9.

There are three basic ways we will make an abstract group G act: left multiplication of G on itself, conjugation of G on itself, and left multiplication of G on a coset space G/H. All of these will now be described.

Example 2.11. To make G act on itself by *left multiplication*, we let $X = G$ and $g \cdot x$ (for $g \in G$ and $x \in G$) is the usual product of g and x in G. This example was used already in the proof of Cayley's theorem and in Example 1.5, and the definition of a group action is satisfied by the axioms for multiplication in G, *e.g.*, $g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x$ from associativity in G.

Note that right multiplication of G on itself, given by $r_g(x) = xg$ for g and x in G, is not an action since the order of composition gets reversed: $r_{g_1} \circ r_{g_2} = r_{g_2 g_1}$. But if we set $r_g(x) = xg^{-1}$ then we do get an action. This could be called the action by right-inverse multiplication (nonstandard terminology).

Example 2.12. To make G act on itself by *conjugation*, let $X = G$ and let $g \cdot x = gxg^{-1}$. Here $g \in G$ and $x \in G$. Since $e \cdot x = exe^{-1} = x$ and

$$\begin{aligned} g_1 \cdot (g_2 \cdot x) &= g_1 \cdot (g_2 x g_2^{-1}) \\ &= g_1 (g_2 x g_2^{-1}) g_1^{-1} \\ &= (g_1 g_2) x (g_1 g_2)^{-1} \\ &= (g_1 g_2) \cdot x, \end{aligned}$$

conjugation is a group action.

Example 2.13. For a subgroup $H \subset G$, consider the left coset space $G/H = \{aH : a \in G\}$. (We do *not* care whether or not $H \lhd G$, as we are just thinking about G/H as a set.) Let G act on G/H by *left multiplication*. That is, for $g \in G$ and a left coset aH $(a \in G)$, set

$$g \cdot aH = gaH = \{gy : y \in aH\}.$$

This is an action of G on G/H, since $eaH = aH$ and

$$\begin{aligned} g_1 \cdot (g_2 \cdot aH) &= g_1 \cdot (g_2 aH) \\ &= g_1 g_2 aH \\ &= (g_1 g_2) \cdot aH. \end{aligned}$$

Example 2.11 is the special case when H is trivial.

Example 2.14. Let $G = \mathbf{Z}/(4)$ act on itself ($X = G$) by additions. For instance, addition by 1 has the effect $0 \mapsto 1 \mapsto 2 \mapsto 3 \mapsto 0$. Therefore addition by 1 on $\mathbf{Z}/(4)$ is a 4-cycle (0123). Addition by 2 has the effect $0 \mapsto 2$, $1 \mapsto 3$, $2 \mapsto 0$, and $3 \mapsto 1$. Therefore, as a permutation on $\mathbf{Z}/(4)$, addition by 2 is $(02)(13)$, a product of two 2-cycles. The composition of these two permutations is $(0123)(02)(13) = (0321)$, which is the permutation of G described by addition by 3, and $3 = 1+2$ in $\mathbf{Z}/(4)$. (This is a special case of Example 2.11 using additive notation.)

We return to the action of a group G on itself by left multiplication and by conjugation, and extend these actions to subsets rather than just points.

Example 2.15. When A is a subset of G, and $g \in G$, the subset $gA = \{ga : a \in A\}$ has the same size as A. Therefore G acts by left multiplication on the set of subsets of G, or even on the subsets with a fixed size. Example 2.11 is the special case of one-element subsets of G. Notice that, when $H \subset G$ is a subgroup, gH is usually *not* a subgroup of G, so the left multiplication action of G on its subsets does not convert subgroups into other subgroups.

Example 2.16. As a special case of Example 2.15, let S_4 act on the set of pairs from $\{1, 2, 3, 4\}$ by the rule $\sigma \cdot \{a, b\} = \{\sigma(a), \sigma(b)\}$.

There are 6 pairs:

$$x_1 = \{1,2\}, x_2 = \{1,3\}, x_3 = \{1,4\}, x_4 = \{2,3\}, x_5 = \{2,4\}, x_6 = \{3,4\}.$$

The effect of (12) on these pairs is

$$(12)x_1 = x_1, \quad (12)x_2 = x_4, \quad (12)x_3 = x_5,$$

$$(12)x_4 = x_2, \quad (12)x_5 = x_3, \quad (12)x_6 = x_6.$$

Thus, as a permutation of the set $\{x_1, \dots, x_6\}$, (12) acts like $(x_2x_4)(x_3x_5)$. That is interesting: we have made a transposition in S_4 look like a product of two 2-cycles in S_6. In particular, we have made an odd permutation of $\{1, 2, 3, 4\}$ look like an even permutation on a new set. This is an embedding $S_4 \hookrightarrow A_6$.

Example 2.17. Let G be a group. When $A \subset G$, gAg^{-1} is a subset with the same size as A. Moreover, unlike the left multiplication action of G on its subsets, the conjugation action of G on its subsets transforms subgroups into subgroups: when $H \subset G$ is a subgroup, gHg^{-1} is also a subgroup. For instance, three of the (seven) subgroups of S_4 with size 4 are

$$\{(1), (1234), (13)(24), (1432)\}, \quad \{(1), (2134), (23)(14), (2431)\},$$

$$\{(1), (12)(34), (13)(24), (14)(23)\}.$$

Under conjugation by S_4, the first two subgroups can be transformed into each other, but neither of these subgroups can be conjugated to the third subgroup: the first and second subgroups have an element with order 4 while the third one does not.

While the left multiplication action of G on itself (Example 2.11) turns different group elements into different permutations, the conjugation action of G on itself (Example 2.12) can make different group elements act in the same way: if $g_1 = g_2 z$, where z is in the center of G, then g_1 and g_2 have the same conjugation action on G. Group actions where different elements of the group act differently have a special name:

Definition 2.18. A group action of G on X is called *faithful* (or *effective*) if different elements of G act on X in different ways: when $g_1 \neq g_2$ in G, there is an $x \in X$ such that $g_1 \cdot x \neq g_2 \cdot x$.

Note that when we say g_1 and g_2 act differently, we mean they act differently somewhere, not everywhere. This is consistent with what it means to say two functions are not equal: they take different values somewhere, not everywhere.

Example 2.19. The action of G on itself by left multiplication is faithful: different elements send e to different places.

Example 2.20. The action of G on itself by conjugation is faithful if and only if G has a trivial center, because $g_1 g g_1^{-1} = g_2 g g_2^{-1}$ for all $g \in G$ if and only if $g_2^{-1} g_1$ is in the center of G. When D_4 acts on itself by conjugation, the action is not faithful since r^2 acts trivially (it is in the center), so 1 and r^2 act in the same way.

Example 2.21. When H is a subgroup of G and G acts on G/H by left multiplication (Example 2.13), g_1 and g_2 in G act in the same way on G/H precisely when $g_1 g H = g_2 g H$ for all $g \in G$, which means $g_2^{-1} g_1 \in \bigcap_{g \in G} gHg^{-1}$. So the left multiplication action of G on G/H is faithful if and only if the subgroups gHg^{-1} (as g varies) have trivial intersection.

Example 2.22. The action of $\mathrm{GL}_2(\mathbf{R})$ on $\mathbf{R}^2$ is faithful, since we can recover the columns of a matrix by acting it on $\binom{1}{0}$ and $\binom{0}{1}$.

Viewing group actions as homomorphisms (Theorem 1.7), a faithful action of G on X is an injective homomorphism $G \to \mathrm{Sym}(X)$. Nonfaithful actions are not injective as group homomorphisms, and many important homomorphisms are not injective.

Remark 2.23. What we have been calling a group action could be called a left group action, while a right group action, denoted xg, has the properties $xe = x$ and $(xg_1)g_2 = x(g_1 g_2)$. The exponential notation x^g in place of xg works well here, especially by writing the identity in the group as 1: $x^1 = x$ and $(x^{g_1})^{g_2} = x^{g_1 g_2}$. The distinction between left and right actions is how a product gg' acts: in a left action g' acts first and g acts second, while in a right action g acts first and g' acts second.

Right multiplication of G on itself (or more generally right multiplication of G on the space of right cosets of a subgroup H) is an example of a right action. To take a more concrete example, the action of $\mathrm{GL}_n(\mathbf{R})$ on *row* vectors of length n is most naturally a right action since the product $\mathbf{v}A$ (not $A\mathbf{v}$) makes sense when $\mathbf{v}$ is a row vector and $A \in \mathrm{GL}_n(\mathbf{R})$. The wrong definitions of actions π_g in Examples 2.9 and 2.10, which were wrong because formulas came out backwards ($\pi_g \circ \pi_h = \pi_{hg}$) are legitimate right actions of G.

Many group theorists (unlike most other mathematicians) like to define the conjugate of h by g as $g^{-1}hg$ instead of as ghg^{-1}, and this convention fits well with the right (but not left) conjugation action: setting $h^g = g^{-1}hg$ we have $h^1 = h$ and $(h^{g_1})^{g_2} = h^{g_1 g_2}$.

The difference between left and right actions of a group is largely illusory, since replacing g with g^{-1} in the group turns left actions into right actions and conversely because inversion reverses the order of multiplication in G. We saw this idea at work in Examples 2.9, 2.10, and 2.11. We will not use right actions (except in Example 3.24), so for us "group action" means "left group action."

3. Orbits and Stabilizers

The information encoded in a group action has two basic parts: one part tells us where points go and the other part tells us how points stay put. The following terminology refers to these ideas.

Definition 3.1. Let a group G act on a set X. For each $x \in X$, its *orbit* is

$$\mathrm{Orb}_x = \{g \cdot x : g \in G\} \subset X$$

and its *stabilizer* is

$$\mathrm{Stab}_x = \{g \in G : g \cdot x = x\} \subset G.$$

(The stabilizer of x is often denoted G_x in the literature, where G is the group.) We call x a *fixed point* for the action when $g \cdot x = x$ for every $g \in G$, that is, when $\mathrm{Orb}_x = \{x\}$ (or equivalently, when $\mathrm{Stab}_x = G$).

Writing the definition of orbits and stabilizers in words, the orbit of a point is a *geometric* concept: the set of places where the point can be moved by the group action. The stabilizer of a point is an *algebraic* concept: the set of group elements that fix the point.

We will often refer to the elements of X as *points* and we will refer to the size of an orbit as its *length*. If $X = G$, as in Examples 2.11 and 2.12, then we think about elements of G as permutations when they act on G and as points when they are acted upon.

Example 3.2. When $\mathrm{GL}_2(\mathbf{R})$ acts in the usual way on $\mathbf{R}^2$, the orbit of $\mathbf{0}$ is $\{\mathbf{0}\}$ since $A \cdot \mathbf{0} = \mathbf{0}$ for every A in $\mathrm{GL}_2(\mathbf{R})$. The stabilizer of $\mathbf{0}$ is $\mathrm{GL}_2(\mathbf{R})$.

The orbit of $\binom{1}{0}$ is $\mathbf{R}^2 - \{\mathbf{0}\}$, in other words every nonzero vector can be obtained from $\binom{1}{0}$ by applying a suitable invertible matrix to it. Indeed, if $\binom{a}{b} \neq \mathbf{0}$, then we have $\binom{a}{b} = \left(\begin{smallmatrix} a & 1 \\ b & 0 \end{smallmatrix}\right)\binom{1}{0}$ and $\binom{a}{b} = \left(\begin{smallmatrix} a & 0 \\ b & 1 \end{smallmatrix}\right)\binom{1}{0}$. One of the matrices $\left(\begin{smallmatrix} a & 1 \\ b & 0 \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} a & 0 \\ b & 1 \end{smallmatrix}\right)$ is invertible (since a or b is not zero), so $\binom{a}{b}$ is in the $\mathrm{GL}_2(\mathbf{R})$-orbit of $\binom{1}{0}$. The stabilizer of $\binom{1}{0}$ is $\{\left(\begin{smallmatrix} 1 & x \\ 0 & y \end{smallmatrix}\right) : y \neq 0\} \subset \mathrm{GL}_2(\mathbf{R})$.

Example 3.3. When the group $\mathrm{GL}_2(\mathbf{Z})$ acts in the usual way on $\mathbf{Z}^2$, the orbit of $\mathbf{0}$ is $\{\mathbf{0}\}$ with stabilizer $\mathrm{GL}_2(\mathbf{Z})$. But in contrast to Example 3.2, the orbit of $\binom{1}{0}$ under $\mathrm{GL}_2(\mathbf{Z})$ is not $\mathbf{Z}^2 - \{\mathbf{0}\}$. Indeed, a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{GL}_2(\mathbf{Z})$ sends $\binom{1}{0}$ to $\binom{a}{c}$, which is a vector with relatively prime coordinates since $ad - bc = \pm 1$. (For instance, $\mathrm{GL}_2(\mathbf{Z})$ can't send $\binom{1}{0}$ to $\binom{2}{0}$.) Conversely, each vector $\binom{m}{n}$ in $\mathbf{Z}^2$ with relatively prime coordinates is in the $\mathrm{GL}_2(\mathbf{Z})$-orbit of $\binom{1}{0}$: we can solve $mx + ny = 1$ for some integers x and y, so $\left(\begin{smallmatrix} m & -y \\ n & x \end{smallmatrix}\right)$ is in $\mathrm{GL}_2(\mathbf{Z})$ (its determinant is 1) and $\left(\begin{smallmatrix} m & -y \\ n & x \end{smallmatrix}\right)\binom{1}{0} = \binom{m}{n}$.

Check as an exercise that the orbits in $\mathbf{Z}^2$ under the action of $\mathrm{GL}_2(\mathbf{Z})$ are the vectors whose coordinates have a fixed greatest common divisor. Each orbit contains one vector of the form $\binom{d}{0}$ for $d \geq 0$, and the stabilizer of $\binom{d}{0}$ for $d > 0$ is $\{\left(\begin{smallmatrix} 1 & x \\ 0 & y \end{smallmatrix}\right) : y = \pm 1\} \subset \mathrm{GL}_2(\mathbf{Z})$.

Example 3.4. Identifying $\mathbf{Z}/(2)$ with the subgroup $\{\pm I_n\}$ of $\mathrm{GL}_n(\mathbf{R})$ gives an action of $\mathbf{Z}/(2)$ on $\mathbf{R}^n$, where 0 acts as the identity and 1 acts by negation on $\mathbf{R}^n$. We can restrict this action of $\mathbf{Z}/(2)$ to the unit sphere of $\mathbf{R}^n$, and then it is called the *antipodal* action since its orbits are pairs of opposite points (which are called antipodal points) on the sphere.

Example 3.5. When the Rubik's cube group acts on the non-centerface cubelets of Rubik's cube, there are two orbits: the 8 corner cubelets and the 12 edge cubelets.

Example 3.6. For $n \geq 2$, consider S_n in its natural action on $\{1, 2, \ldots, n\}$. What is the stabilizer of an integer $i \in \{1, 2, \ldots, n\}$? It is the set of permutations of $\{1, 2, \ldots, n\}$ fixing i, which can be thought of as the set of permutations of $\{1, 2, \ldots, n\} - \{i\}$. This is an isomorphic copy of S_{n-1} inside S_n (once we identify $\{1, 2, \ldots, n\} - \{i\}$ in a definite manner with the numbers from 1 to $n - 1$). The stabilizer of each number in $\{1, 2, \ldots, n\}$ for the natural action of S_n on $\{1, 2, \ldots, n\}$ is isomorphic to S_{n-1}.

Example 3.7. For $n \geq 2$, the even permutations of $\{1, 2, \ldots, n\}$ that fix a number k can be identified with the even permutations of $\{1, 2, \ldots, n\} - \{k\}$, so the stabilizer of each point in the natural action of A_n is essentially A_{n-1} up to relabelling.

Remark 3.8. When trying to think about a set as a geometric object, it is helpful to refer to its elements as points, no matter what they might really be. For example, when we think about G/H as a set on which G acts (by left multiplication), it is useful to think about the cosets of H, which are the elements of G/H, as the points in G/H. At the same time, though, a coset is a subset of G. There is a tension between these two interpretations: is a left coset of H a point in G/H or a subset of G? It is both, and it is important to be able to think about a coset in both ways.

All of our applications of group actions to group theory will flow from the relations between orbits, stabilizers, and fixed points, which we now make explicit in our three basic examples of group actions.

Example 3.9. When a group G acts on itself by left multiplication,

- there is one orbit ($g = ge \in \mathrm{Orb}_e$),
- $\mathrm{Stab}_a = \{g : ga = a\} = \{e\}$ is trivial,
- there are no fixed points (if $|G| > 1$).

Example 3.10. When a group G acts on itself by conjugation,

- the orbit of a is $\mathrm{Orb}_a = \{gag^{-1} : g \in G\}$, which is the conjugacy class of a,
- $\mathrm{Stab}_a = \{g : gag^{-1} = a\} = \{g : ga = ag\}$ is the centralizer of a, denoted $Z(a)$,
- a is a fixed point when it commutes with all elements of G, and thus the fixed points of conjugation form the center $Z(G)$.

Example 3.11. When a group G acts on G/H (for a subgroup H) by left multiplication,

- there is one orbit ($gH = g \cdot H \in \mathrm{Orb}_H$),
- $\mathrm{Stab}_{aH} = \{g : gaH = aH\} = \{g : a^{-1}ga \in H\} = aHa^{-1}$,
- there are no fixed points (if $H \neq G$).

These examples illustrate several facts: an action need not have fixed points (Example 3.9 with nontrivial G), different orbits can have different lengths (Example 3.10 with $G = S_3$), and the points in a common orbit don't have to share the same stabilizer (Example 3.11 if H is not a normal subgroup).

Example 3.12. When G acts on its subgroups by conjugation, $\mathrm{Stab}_H = \{g : gHg^{-1} = H\}$ is the normalizer $\mathrm{N}(H)$ and the fixed points are the normal subgroups of G.

When a group G acts on a set X, each subgroup H of G also acts on X. Let's look at a few examples.

Example 3.13. When H acts on G by left multiplication,

- the orbit of $a \in G$ is $\{ha : h \in H\} = Ha$, a *right* H-coset,
- $\mathrm{Stab}_a = \{h : ha = a\} = \{e\}$ is trivial,
- there are no fixed points (if $|H| > 1$).

Example 3.14. When H acts on G by right-inverse multiplication (see Example 2.11),

- the orbit of $a \in G$ is $\mathrm{Orb}_a = \{ah^{-1} : h \in H\} = aH$, a *left* H-coset,
- $\mathrm{Stab}_a = \{h : ah^{-1} = a\} = \{e\}$ is trivial,
- there are no fixed points (if $|H| > 1$).

Example 3.15. When H acts on G by conjugation,

- the H-orbit of a is $\mathrm{Orb}_a = \{hah^{-1} : h \in H\}$, which has no special name (this is the elements of G that are H-conjugate to a),
- $\mathrm{Stab}_a = \{h \in H : hah^{-1} = a\} = \{h : ha = ah\}$ is the elements of H commuting with a (this is $H \cap Z(a)$, where $Z(a)$ is the centralizer of a in G).
- a is a fixed point when it commutes with all elements of H.

In the summary table below, G is a group and H is a subgroup of G.

Group	Set	Action	Orbit of x	Stabilizer of x
S_n	$\{1, \dots, n\}$	$\sigma \cdot i = \sigma(i)$	$\{1, \dots, n\}$	$\{\sigma : \sigma(x) = x\} \cong S_{n-1}$
G	G	$g \cdot x = gx$	G	$\{e\}$
G	G	$g \cdot x = gxg^{-1}$	Conj. class of x	$\{g : gx = xg\}$
H	G	$h \cdot x = hx$	Hx	$\{e\}$
H	G	$h \cdot x = xh^{-1}$	xH	$\{e\}$
G	G/H	$g \cdot aH = gaH$	G/H	aHa^{-1} $(x = aH)$

Here is the **fundamental theorem** about group actions.

Theorem 3.16. *Let a group G act on a set X.*

a) Different orbits of the action are disjoint and form a partition of X.

b) For each $x \in X$, Stab_x is a subgroup of G and $\mathrm{Stab}_{gx} = g\,\mathrm{Stab}_x\,g^{-1}$ for all $g \in G$.

c) For each $x \in X$, there is a bijection $\mathrm{Orb}_x \to G/\mathrm{Stab}_x$ by $gx \mapsto g\,\mathrm{Stab}_x$. More concretely, $gx = g'x$ if and only if g and g' lie in the same left coset of Stab_x, and different left cosets of Stab_x correspond to different points in Orb_x. In particular, if x and y are in the same orbit then $\{g \in G : gx = y\}$ is a left coset of Stab_x, and

$$|\mathrm{Orb}_x| = [G : \mathrm{Stab}_x].$$

Parts b and c show the role of conjugate subgroups and cosets of a subgroup when working with group actions. The formula in part c that relates the length of an orbit to the index in G of a stabilizer for a point in the orbit, is called the *orbit-stabilizer formula.*

Proof. a) We prove different orbits in a group action are disjoint by proving that two orbits that overlap must coincide.[2] Suppose Orb_x and Orb_y have a common element z:

$$z = g_1 x, \quad z = g_2 y.$$

We want to show $\mathrm{Orb}_x = \mathrm{Orb}_y$. It suffices to show $\mathrm{Orb}_x \subset \mathrm{Orb}_y$, since then we can switch the roles of x and y to get the reverse inclusion.

[2] The argument will be similar to the proof that different left cosets of a subgroup are disjoint: if the cosets overlap they coincide.

For each point $u \in \mathrm{Orb}_x$, write $u = gx$ for some $g \in G$. Since $x = g_1^{-1}z$,

$$u = g(g_1^{-1}z) = (gg_1^{-1})z = (gg_1^{-1})(g_2y) = (gg_1^{-1}g_2)y,$$

which shows us that $u \in \mathrm{Orb}_y$. Therefore $\mathrm{Orb}_x \subset \mathrm{Orb}_y$.

Every element of X is in some orbit (its own orbit), so the orbits partition X into disjoint subsets.

b) To see that Stab_x is a subgroup of G, we have $e \in \mathrm{Stab}_x$ since $ex = x$, and if $g_1, g_2 \in \mathrm{Stab}_x$, then

$$(g_1g_2)x = g_1(g_2x) = g_1x = x,$$

so $g_1g_2 \in \mathrm{Stab}_x$. Thus Stab_x is closed under multiplication. Lastly,

$$gx = x \Longrightarrow g^{-1}(gx) = g^{-1}x \Longrightarrow x = g^{-1}x,$$

so Stab_x is closed under inversion.

To show $\mathrm{Stab}_{gx} = g\,\mathrm{Stab}_x\,g^{-1}$, for all $x \in X$ and $g \in G$, observe that

$$\begin{aligned} h \in \mathrm{Stab}_{gx} &\Longleftrightarrow h \cdot (gx) = gx \\ &\Longleftrightarrow (hg)x = gx \\ &\Longleftrightarrow g^{-1}((hg)x) = g^{-1}(gx) \\ &\Longleftrightarrow (g^{-1}hg)x = x \\ &\Longleftrightarrow g^{-1}hg \in \mathrm{Stab}_x \\ &\Longleftrightarrow h \in g\,\mathrm{Stab}_x\,g^{-1}, \end{aligned}$$

so $\mathrm{Stab}_{gx} = g\,\mathrm{Stab}_x\,g^{-1}$.

c) The condition $gx = g'x$ is equivalent to $x = (g^{-1}g')x$, which means $g^{-1}g' \in \mathrm{Stab}_x$, or $g' \in g\,\mathrm{Stab}_x$. Therefore g and g' have the same effect on x if and only if g and g' lie in the same left coset of Stab_x. (Recall that for all subgroups H of G, $g' \in gH$ if and only if $g'H = gH$.)

Since Orb_x consists of the points gx for varying g, and we showed elements of G have the same effect on x if and only if they lie in the same left coset of Stab_x, we get a bijection between the points in the orbit of x and the left cosets of Stab_x by $gx \mapsto g\,\mathrm{Stab}_x$. (Think carefully about why this is well-defined.) Therefore the cardinality of the orbit of x, which is $|\,\mathrm{Orb}_x\,|$ equals the cardinality of the left cosets of Stab_x in G. □

Remark 3.17. That the orbits of a group action partition the set includes as special cases two basic partition results in group theory: the left (or right) cosets of a subgroup and the conjugacy classes of a group partition the group into disjoint parts. The partition by cosets uses the right-inverse multiplication action (or left multiplication action) of the subgroup on the group (see Examples 2.11, 3.13, and 3.14) and the partition into conjugacy classes use the action of a group on itself by conjugation (see Examples 2.12 and 3.10).

Example 3.18. For $n \geq 2$ and $k \in \{1, \dots, n-1\}$, the group $G = S_n$ acts on the k-element subsets of $\{1, 2, \dots, n\}$ in the usual way: $\sigma(\{i_1, \dots, i_k\}) = \{\sigma(i_1), \dots, \sigma(i_k)\}$. This group action has one orbit since $\{i_1, \dots, i_k\} = \sigma(\{1, \dots, k\})$ where σ is the permutation $\left(\begin{smallmatrix} 1 & 2 & \cdots & k \\ i_1 & i_2 & \cdots & i_k \end{smallmatrix}\right)$.

The number of k-element subsets of $\{1, \dots, n\}$ is $\binom{n}{k}$, by the combinatorial definition of binomial coefficients, so Theorem 3.16(c) implies $\binom{n}{k} = [S_n : \mathrm{Stab}_{\{1,\dots,k\}}]$. What is the stabilizer of $\{1, \dots, k\}$? It is all $\sigma \in S_n$ such that $\{\sigma(1), \dots, \sigma(k)\} = \{1, \dots, k\}$ (equality of sets, not ordered sets or k-tuples), which is the same as saying σ permutes $\{1,\dots,k\}$ and thus also permutes $\{k+1,\dots,n\}$. Therefore $\mathrm{Stab}_{\{1,\dots,k\}} \cong S_k \times S_{n-k}$, so

$$\binom{n}{k} = [S_n : \mathrm{Stab}_{\{1,\dots,k\}}] = \frac{n!}{k!(n-k)!}.$$

This is a derivation of the standard formula for $\binom{n}{k}$ using group actions.

Corollary 3.19. *Let a finite group G act on a set X.*

a) The length of every orbit in X divides the size of G.

b) Points in a common orbit have conjugate stabilizers, and in particular the size of the stabilizer is the same for all points in an orbit.

Proof. a) For $x \in X$, the length of the orbit of x is $[G : \mathrm{Stab}_x]$, which divides $|G|$.

b) If x and y are in the same orbit, write $y = gx$. Then $\mathrm{Stab}_y = \mathrm{Stab}_{gx} = g\,\mathrm{Stab}_x\,g^{-1}$, so the stabilizers of x and y are conjugate subgroups. □

A converse of part b is not generally true: points with conjugate stabilizers need not be in the same orbit. Even points with the *same* stabilizer need not be in the same orbit. For example, if G acts on itself trivially then all points have stabilizer G and all orbits have size 1. For a more interesting example, let A_4 act on itself by conjugation. Then (123) and (132) are in different orbits (they are not conjugate in A_4) but they each have stabilizer $\{(1),(123),(132)\}$. The same feature is true of g and g^{-1} for every 3-cycle $g \in A_4$.

Corollary 3.20. *Let a group G act on a set X, where X is finite. Let the different orbits of X be represented by $x_1,\dots,x_t$. Then*

$$|X| = \sum_{i=1}^{t} |\,\mathrm{Orb}_{x_i}| = \sum_{i=1}^{t} [G : \mathrm{Stab}_{x_i}]. \tag{3.1}$$

Proof. The set X can be written as the union of its orbits, which are mutually disjoint. The orbit-stabilizer formula tells us how large each orbit is. □

Example 3.21. In a finite group G, the size of every conjugacy class divides $|G|$ since conjugacy classes are orbits for the conjugation action of G on itself. For instance, when $G = S_3$ its conjugacy classes are $\{(1)\}$, $\{(123),(132)\}$, and $\{(12),(13),(23)\}$, whose sizes 1, 2, and 3 are all factors of 6: (3.1) here says $6 = 1+2+3$. When $G = S_4$ its conjugacy classes are represented by (1), (1234), (12)(34), (123), and (12) and their conjugacy classes have respective sizes 1, 6, 3, 8, and 6. All are factors of 24 and (3.1) here says $24 = 1+6+3+8+6$.

Example 3.22. Which elements of D_6 commute with the reflection s? This is asking for $\{g \in D_6 : gs = sg\}$. Three such elements are 1, s, and r^3 (since $r^{n/2} \in Z(D_n)$ for even n).

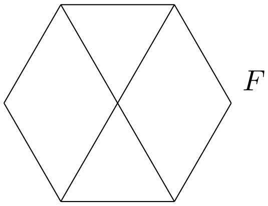
Let's interpret the condition $gs = sg$ as $gsg^{-1} = s$: the task is now computing the stabilizer of s when D_6 acts on itself by conjugation. To compute the stabilizer, let's first compute the orbit: how many different values of gsg^{-1} are there as g runs over D_6? Elements of D_6 are r^k (rotations) and $r^k s$ (reflections, so equal to their inverses). From

$$r^k s r^{-k} = r^{2k}s, \quad (r^k s)s(r^k s)^{-1} = r^k s s r^k s = r^k r^k s = r^{2k}s$$

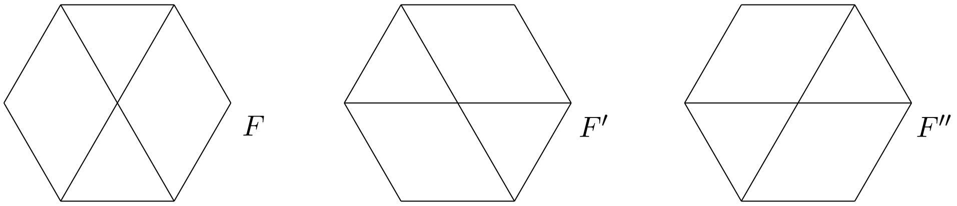
the different gsg^{-1} as g varies in D_6 is $\{r^{\mathrm{even}}s\} = \{s, r^2s, r^4s\}$.

Since the D_6-orbit of s has size 3, the stabilizer of s has index 3 in D_6 and thus its size is $|D_6|/3 = 12/3 = 4$. We already know 1, s, and r^3 are in the stabilizer, so being a group means r^3s is in the stabilizer too. That is a fourth element, and the stabilizer has size 4, so $\{g \in D_6 : gs = sg\} = \{1, s, r^3, r^3s\}$.

Example 3.23. We examine now a geometric example. The figure F below is a hexagon with an X drawn inside of it. Which elements of D_6 preserve this figure when D_6 acts in a natural way on it?



For $g \in D_6$, $g(F) = F$ means $g \in \mathrm{Stab}_F$. To compute Stab_F we first compute the orbit of F: it's easier to figure out all the ways F can change than to figure out all the ways F can stay the same, and these are related by the orbit-stabilizer formula. By rotating and reflecting it is clear that $g(F)$, as g runs over D_6, has only the 3 results below.



Let r be the 60-degree counterclockwise rotation preserving the hexagonal shape and let s be the reflection across the horizontal line bisecting F. Since F has an orbit of size 3, its stabilizer in D_6 has index 3, so $|\,\mathrm{Stab}_F\,| = |D_6|/3 = 12/3 = 4$. From the 180-degree rotational symmetry of F, $r^3 \in \mathrm{Stab}_F$. Since $s(F) = F$, $s \in \mathrm{Stab}_F$. Since Stab_F is a *subgroup* of D_6, Stab_F also contains r^3s. Thus $\{1, r^3, s, r^3s\} \subset \mathrm{Stab}_F$, and we are done since we know $|\,\mathrm{Stab}_F\,| = 4$: $\boxed{\mathrm{Stab}_F = \{1, r^3, s, r^3s\} = \langle r^3, s\rangle}$.

While F' looks like F, it is not equal to F. What are $\mathrm{Stab}_{F'}$ and $\{g \in D_6 : g(F) = F'\}$? We can compute both as soon as we know just *one* g sending F to F'. Since $F' = r(F)$ we can use $g = r$. Then Theorem 3.16(b) says

$$\mathrm{Stab}_{F'} = \mathrm{Stab}_{r(F)} = r\,\mathrm{Stab}_F\,r^{-1} = r\{1, r^3, s, r^3s\}r^{-1} = \{1, r^3, r^2s, r^5s\}$$

and Theorem 3.16(c) says

$$\{g \in D_6 : g(F) = F'\} = r\,\mathrm{Stab}_F = r\{1, r^3, s, r^3s\} = \{r, r^4, rs, r^4s\}.$$

Similarly, since $F'' = r^{-1}(F'')$,

$$\mathrm{Stab}_{F''} = \mathrm{Stab}_{r^{-1}(F)} = r^{-1}\,\mathrm{Stab}_F(r^{-1})^{-1} = r^{-1}\{1, r^3, s, r^3s\}r = \{1, r^3, r^4s, rs\}$$

and

$$\{g \in D_6 : g(F) = F''\} = r^{-1}\,\mathrm{Stab}_F = r^{-1}\{1, r^3, s, r^3s\} = \{r^5, r^2, r^5s, r^2s\}.$$

Example 3.24. The 2×2 matrices $(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}) \in \mathrm{GL}_2(\mathbf{R})$ whose columns add up to 1 form a subgroup H. This can be checked by a tedious calculation. It can also be seen by observing that the column sums are the entries in the vector-matrix product $(1\ 1)(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix})$, so the matrices in H are those satisfying $(1\ 1)(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}) = (1\ 1)$. So H is the stabilizer of $(1\ 1)$ in the (right!) action of $\mathrm{GL}_2(\mathbf{R})$ on $\mathbf{R}^2$ – viewed as row vectors – by $\mathbf{v} \cdot A = \mathbf{v}A$. Thus H is a subgroup of $\mathrm{GL}_2(\mathbf{R})$ since the stabilizers of a point are always a subgroup. (Theorem 3.16 for right group actions should be formulated and checked by the reader.)

Moreover, because $(0\ 1)(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}) = (1\ 1)$, $\mathrm{Stab}_{(1\ 1)}$ and $\mathrm{Stab}_{(0\ 1)}$ are conjugate subgroups in $\mathrm{GL}_2(\mathbf{R})$. Since $\mathrm{Stab}_{(0\ 1)} = \{(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}) \in \mathrm{GL}_2(\mathbf{R})\} = \mathrm{Aff}(\mathbf{R})$, we have

$$H = \mathrm{Stab}_{(1\ 1)} = \mathrm{Stab}_{(0\ 1)(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix})} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}^{-1} \mathrm{Aff}(\mathbf{R}) \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}.$$

Example 3.25. As a cute application of the orbit-stabilizer formula we explain why $|HK| = |H||K|/|H \cap K|$ for subgroups H and K of a finite group G. Here $HK = \{hk : h \in H, k \in K\}$ is the set of products, which usually is just a subset (not a subgroup) of G. To count the size of HK, let the direct product group $H \times K$ act on G like this: $(h,k) \cdot g = hgk^{-1}$. Check this gives a group action (the group is $H \times K$ and the set is G) and HK is the orbit of e. Therefore the orbit-stabilizer formula tells us

$$|HK| = \frac{|H \times K|}{|\mathrm{Stab}_e|} = \frac{|H||K|}{|\{(h,k) : (h,k) \cdot e = e\}|}.$$

The condition $(h,k) \cdot e = e$ means $hk^{-1} = e$, so $\mathrm{Stab}_e = \{(h,h) : h \in H \cap K\}$. Therefore $|\mathrm{Stab}_e| = |H \cap K|$ and $|HK| = |H||K|/|H \cap K|$.

Example 3.26. We now discuss the original form of Lagrange's theorem in group theory. He proved for each polynomial $f(T_1, \ldots, T_n)$ in n variables that the number of different polynomials we get from $f(T_1, \ldots, T_n)$ by permuting its variables is a factor of $n!$.

For instance, consider the polynomial T_1 and $n = 3$. If we run through all six permutations of $\{T_1, T_2, T_3\}$, and apply each to T_1, we get 3 different results: T_1, T_2, and T_3. The polynomial $T_1T_2^2 + T_2T_3^2 + T_3T_1^2$ has only 2 possibilities under each change of variables: itself and $T_2T_1^2 + T_1T_3^2 + T_3T_2^2$ (check this). The polynomial $T_1 + T_2^2 + T_3^3$ has 6 different possibilities. The number of different polynomials in each case is a factor of $3!$.

To explain Lagrange's general observation, we apply the orbit-stabilizer formula to the group action in Example 2.8. That is the action of S_n on n-variable polynomials by permutations of the variables. For an n-variable polynomial $f(T_1, \ldots, T_n)$, the different polynomials we obtain by permuting its variables are exactly the polynomials in its S_n-orbit. By the orbit-stabilizer formula, the number of different polynomials we get from $f(T_1, \ldots, T_n)$ by permuting its variables is $[S_n : H_f]$, where $H_f = \mathrm{Stab}_f = \{\sigma \in S_n : \sigma \cdot f = f\}$, and this index divides $n!$. Cauchy introduced the term "index" in 1815 for the number of different polynomials we get from a single polynomial by permuting its variables, and its interpretation as $[S_n : H_f]$ is why we use the term index for $[G : H]$ in group theory.

In a group action, the length of an orbit divides $|G|$, but the number of orbits usually does not divide $|G|$. For example, D_4 and Q_8 each have 5 conjugacy classes, and 5 does not divide 8. But there is an interesting relation between the number of orbits and the group action.

Theorem 3.27. *Let a finite group G act on a finite set X with r orbits. Then r is the average number of fixed points of the elements of the group:*

$$r = \frac{1}{|G|} \sum_{g \in G} |\mathrm{Fix}_g(X)|,$$

where $\mathrm{Fix}_g(X) = \{x \in X : gx = x\}$ is the set of elements of X fixed by g.

Don't confuse the set $\mathrm{Fix}_g(X)$ with the fixed points for the action: $\mathrm{Fix}_g(X)$ is only the points fixed by the element g. The set of fixed points for the action of G is the intersection of the sets $\mathrm{Fix}_g(X)$ as g runs over the group.

Proof. We will count $\{(g,x) \in G \times X : gx = x\}$ in two ways.

By counting over g's first we have to add up the number of x's with $gx = x$, so

$$|\{(g,x) \in G \times X : gx = x\}| = \sum_{g \in G} |\operatorname{Fix}_g(X)|.$$

Next we count over the x's and have to add up the number of g's with $gx = x$, *i.e.*, with $g \in \operatorname{Stab}_x$:

$$|\{(g,x) \in G \times X : gx = x\}| = \sum_{x \in X} |\operatorname{Stab}_x|.$$

Equating these two counts gives

$$\sum_{g \in G} |\operatorname{Fix}_g(X)| = \sum_{x \in X} |\operatorname{Stab}_x|.$$

By the orbit-stabilizer formula, $|G|/|\operatorname{Stab}_x| = |\operatorname{Orb}_x|$, so

$$\sum_{g \in G} |\operatorname{Fix}_g(X)| = \sum_{x \in X} \frac{|G|}{|\operatorname{Orb}_x|}.$$

Divide by $|G|$:

$$\frac{1}{|G|}\sum_{g \in G} |\operatorname{Fix}_g(X)| = \sum_{x \in X} \frac{1}{|\operatorname{Orb}_x|}.$$

Let's consider the contribution to the right side from points in a single orbit. If an orbit has n points in it, then the sum over the points in that orbit is a sum of $1/n$ for n terms, and that is equal to 1. Thus the part of the sum over points in an orbit is 1, which makes the sum on the right side equal to the number of orbits, which is r. □

Theorem 3.27 is often called Burnside's lemma, but it is not due to him [6]. He included it in his widely read book on group theory.

Example 3.28. We will use a special case of Theorem 3.27 to prove for all $a \in \mathbf{Z}$ and $m \in \mathbf{Z}^+$ that

$$\sum_{k=1}^{m} a^{(k,m)} \equiv 0 \bmod m. \tag{3.2}$$

When $m = p$ is a prime number, the left side is $(p-1)a + a^p = (a^p - a) + pa$, so (3.2) becomes $a^p \equiv a \bmod p$, which is Fermat's little theorem. Thus (3.2) can be thought of as a generalization of Fermat's little theorem to all moduli that is essentially different from the generalization called Euler's theorem, which says $a^{\varphi(m)} \equiv 1 \bmod m$ if $(a,m) = 1$: (3.2) is true for all $a \in \mathbf{Z}$.

Our setup leading to (3.2) starts with a finite group G and comes from [4]. For a positive integer a, G acts on the set of functions $\operatorname{Map}(G, \{1, 2, \ldots, a\})$ by $(g \cdot f)(h) = f(g^{-1}h)$ for $g, h \in G$. This is a special case of the group action at the end of Example 2.10, where G acts on itself by left multiplication. We want to apply Theorem 3.27 to this action, so we need to understand the fixed points (really, fixed functions) of each $g \in G$. We have $g \cdot f = f$ if and only if $f(g^{-1}h) = f(h)$ for all $h \in G$, which is the same as saying f is constant on every left coset $\langle g \rangle h$ in G. The number of left cosets of $\langle g \rangle$ in G is $[G : \langle g \rangle] = m/\operatorname{ord}(g)$, where $m = |G|$ and $\operatorname{ord}(g)$ is the order of g, so the number of functions fixed by g is $a^{m/\operatorname{ord}(g)}$, since the value of the function on each coset can be chosen arbitrarily in $\{1, \ldots, a\}$. Therefore Theorem 3.27 implies $(1/m)\sum_{g \in G} a^{m/\mathrm{ord}(g)}$ is a positive integer, so

$$\sum_{g \in G} a^{m/\mathrm{ord}(g)} \equiv 0 \bmod m. \tag{3.3}$$

Since (3.3) depends on a only by the value of $a \bmod m$, it holds for all $a \in \mathbf{Z}$, not just $a > 0$.

Taking $G = \mathbf{Z}/(m)$, each $k \in G$ has additive order $m/(k, m)$, so (3.3) becomes

$$\sum_{k=1}^{m} a^{(k,m)} \equiv 0 \bmod m.$$

Next we turn to the idea of two different actions of a group being essentially the same.

Definition 3.29. Two actions of a group G on sets X and Y are called *equivalent* if there is a bijection $f \colon X \to Y$ such that $f(gx) = g(f(x))$ for all $g \in G$ and $x \in X$.

Actions of G on two sets are equivalent when G permutes elements in the same way on the two sets after matching up the sets appropriately. When $f \colon X \to Y$ is an equivalence of group actions on X and Y, $gx = x$ if and only if $g(f(x)) = f(x)$, so the stabilizer subgroups of $x \in X$ and $f(x) \in Y$ are the same.

Example 3.30. Let $\mathbf{R}^\times$ act on a linear subspace $\mathbf{R}v_0 \subset \mathbf{R}^n$ by scaling. This is equivalent to the natural action of $\mathbf{R}^\times$ on $\mathbf{R}$ by scaling: let $f \colon \mathbf{R} \to \mathbf{R}v_0$ by $f(a) = av_0$. Then f is a bijection and $f(ca) = (ca)v_0 = c(av_0) = cf(a)$ for all c in $\mathbf{R}^\times$ and $a \in \mathbf{R}$.

Example 3.31. Let $\mathrm{GL}_2(\mathbf{R})$ act on the set $\mathcal{B}$ of ordered bases (e_1, e_2) of $\mathbf{R}^2$ in the natural way: for $A \in \mathrm{GL}_2(\mathbf{R})$, $A(e_1, e_2) := (Ae_1, Ae_2)$ is another ordered basis of $\mathbf{R}^2$. This action of $\mathrm{GL}_2(\mathbf{R})$ on $\mathcal{B}$ is equivalent to the action of $\mathrm{GL}_2(\mathbf{R})$ on itself by left multiplication. The reason is that the columns of a matrix in $\mathrm{GL}_2(\mathbf{R})$ are a basis of $\mathbf{R}^2$ (the first and second columns are an ordering of basis vectors: the first column is the first basis vector and the second column is the second one) and two square matrices multiply through multiplication on the columns: $A\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = (A\binom{a}{c}\ A\binom{b}{d})$. Letting $f \colon \mathcal{B} \to \mathrm{GL}_2(\mathbf{R})$ by $f(\binom{a}{c}, \binom{b}{d}) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ gives a bijection and $f(A(e_1, e_2)) = A \cdot f(e_1, e_2)$ for all $A \in \mathrm{GL}_2(\mathbf{R})$ and $(e_1, e_2) \in \mathcal{B}$.

Example 3.32. Let S_3 act on its conjugacy class $\{(12), (13), (23)\}$ by conjugation. This action on a 3-element set, described in the first half of Table 1 below, looks like the usual action of S_3 on $\{1, 2, 3\}$ in the second half of Table 1 if we identify (12) with 3, (13) with 2, and (23) with 1 (in short, identity (ij) with k where $k \not\in \{i, j\}$). Then the action of S_3 on $\{(12), (13), (23)\}$ by conjugation is equivalent to the natural action of S_3 on $\{1, 2, 3\}$.

π	$\pi(12)\pi^{-1}$	$\pi(13)\pi^{-1}$	$\pi(23)\pi^{-1}$	$\pi(3)$	$\pi(2)$	$\pi(1)$
(1)	(12)	(13)	(23)	3	2	1
(12)	(12)	(23)	(13)	3	1	2
(13)	(23)	(13)	(12)	1	2	3
(23)	(13)	(12)	(23)	2	3	1
(123)	(23)	(12)	(13)	1	3	2
(132)	(13)	(23)	(12)	2	1	3

TABLE 1.

Example 3.33. Let H and K be subgroups of G. The group G acts by left multiplication on G/H and G/K. If H and K are conjugate subgroups then these actions are equivalent: fix a representation $K = g_0Hg_0^{-1}$ for some $g_0 \in G$ and let $f\colon G/H \to G/K$ by $f(gH) = gg_0^{-1}K$. This is well-defined (independent of the coset representatives for gH) since, for $h \in H$,

$$f(ghH) = ghg_0^{-1}K = ghg_0^{-1}g_0Hg_0^{-1} = gHg_0^{-1} = gg_0^{-1}K.$$

The reader can check $f(g(g'H)) = gf(g'H)$ for $g \in G$ and $g'H \in G/H$, and f is a bijection. (The mapping f might depend on g_0, but that is not a problem. There can be multiple equivalences between two equivalent group actions, just as there can be multiple isomorphisms between two isomorphic groups.)

If H and K are nonconjugate then the actions of G on G/H and G/K are not equivalent: corresponding points in equivalent actions have the same stabilizer subgroup, but the stabilizer subgroups of left cosets in G/H are conjugate to H and those in G/K are conjugate to K, and none of the former and latter are equal.

The left multiplication action of G on a left coset space G/H has one orbit. It turns out all actions with one orbit are essentially of this form:

Theorem 3.34. *An action of G that has one orbit is equivalent to the left multiplication action of G on some left coset space of G.*

Proof. Suppose that G acts on the set X with one orbit. Fix $x_0 \in X$ and let $H = \mathrm{Stab}_{x_0}$. We will show the action of G on X is equivalent to the left multiplication action of G on G/H.

Every $x \in X$ has the form gx_0 for some $g \in G$, and all elements in a left coset gH have the same effect on x_0: for all $h \in H$, $(gh)(x_0) = g(hx_0) = g(x_0)$. Let $f\colon G/H \to X$ by $f(gH) = gx_0$. This is well-defined, as we just saw. Moreover, $f(g \cdot g'H) = gf(g'H)$ since both sides equal $gg'(x_0)$. We will show f is a bijection.

Since X has one orbit, $X = \{gx_0 : g \in G\} = \{f(gH) : g \in G\}$, so f is onto. If $f(g_1H) = f(g_2H)$ then $g_1x_0 = g_2x_0$, so $g_2^{-1}g_1x_0 = x_0$. Since x_0 has stabilizer H, $g_2^{-1}g_1 \in H$, so $g_1H = g_2H$. Thus f is one-to-one. □

A particular case of Theorem 3.34 says that an action of G is equivalent to the left multiplication action of G on itself if and only if the action has one orbit and the stabilizer subgroups are trivial.

Definition 3.35. The action of G on X is called *free* when every point has a trivial stabilizer.

Example 3.36. The left multiplication action of a group on itself (Example 3.9) is free with one orbit.

Example 3.37. The antipodal action of $\mathbf{Z}/(2)$ on a sphere (Example 3.4) is a free action. There are uncountably many orbits.

Free actions show up often in topology. Example 3.37 is a typical example of that.

Example 3.38. For an integer $n \geq 2$, let X_n be the set of roots of unity of order n in $\mathbf{C}^\times$, so[3] $|X_n| = \varphi(n)$. (For instance, $X_4 = \{i, -i\}$.) The group $(\mathbf{Z}/(n))^\times$ acts on X_n by $a \cdot \zeta = \zeta^a$. (This is well-defined since $a \equiv b \bmod n \Rightarrow \zeta^a = \zeta^b$.) Since every element of X_n is

[3] Having order n is more than just satisfying $z^n = 1$: no smaller power can be 1, so X_n is not all nth roots of unity when $n > 1$.

a power of every other element of X_n using exponents relatively prime to n, this action of $(\mathbf{Z}/(n))^\times$ has a single orbit. Since $\zeta^a = \zeta$ only if $a \equiv 1 \bmod n$ (ζ has order n), all stabilizers are trivial (a free action). Thus $(\mathbf{Z}/(n))^\times$ acting on X_n is equivalent to the multiplication action of $(\mathbf{Z}/(n))^\times$ on itself, except there is no naturally distinguished element of X_n (when $\varphi(n) > 1$, *i.e.*, $n > 2$) while 1 is a distinguished element of $(\mathbf{Z}/(n))^\times$.

It is worth comparing faithful and free actions. An action is faithful (Definition 2.18) when $g_1 \neq g_2$ in $G \Rightarrow g_1x \neq g_2x$ for *some* $x \in X$ (different elements of G act differently at some point in X) while an action is free when $g_1 \neq g_2$ in $G \Rightarrow g_1x \neq g_2x$ for *all* $x \in X$ (different elements of G act differently at every point of X). So all free actions are faithful. Since $g_1x = g_2x$ if and only if $g_2^{-1}g_1x = x$, we can describe faithful and free actions in terms of fixed points: an action is faithful when each $g \neq e$ has $\mathrm{Fix}_g(X) \neq X$ while an action is free when each $g \neq e$ has $\mathrm{Fix}_g(X) = \emptyset$.

4. Actions of p-groups

The action of a group of prime power size has special features. When $|G| = p^k$ for a prime p, we call G a *p-group*. For example, $(\mathbf{Z}/(5))^\times = \{1, 2, 3, 4\}$ and D_4 are 2-groups. The action of a p-group has special features. Because all subgroups of a p-group have p-power index, the length of an orbit under an action by a p-group is divisible by p *unless* the point is a fixed point, when its orbit has length 1. This leads to an important congruence modulo p for actions of a p-group.

Theorem 4.1 (Fixed Point Congruence)**.** *Let G be a finite p-group acting on a finite set X. Then*

$$|X| \equiv |\{\text{fixed points}\}| \bmod p.$$

Proof. Let the different orbits in X be represented by $x_1, \dots, x_t$, so Corollary 3.20 leads to

$$|X| = \sum_{i=1}^{t} |\operatorname{Orb}_{x_i}|. \tag{4.1}$$

Since $|\operatorname{Orb}_{x_i}| = [G : \mathrm{Stab}_{x_i}]$ and $|G|$ is a power of p, $|\operatorname{Orb}_{x_i}| \equiv 0 \bmod p$ unless $\mathrm{Stab}_{x_i} = G$, in which case Orb_{x_i} has length 1, *i.e.*, x_i is a fixed point. Thus when we reduce both sides of (4.1) modulo p, all terms on the right side vanish except for a contribution of 1 for each fixed point. That implies

$$|X| \equiv |\{\text{fixed points}\}| \bmod p. \qquad \square$$

Keep in mind that the congruence in Theorem 4.1 holds only for actions by groups with prime-power size. When a group of size 9 acts we get a congruence mod 3, but when a group of size 6 acts we do not get a congruence mod 2 or 3.

Corollary 4.2. *Let G be a finite p-group acting on a finite set X. If $|X|$ is not divisible by p, then there is at least one fixed point in X. If $|X|$ is divisible by p, then the number of fixed points is a multiple of p (possibly 0).*

Proof. When $|X|$ is not divisible by p, neither is the number of fixed points (by the fixed point congruence), so the number of fixed points can't equal 0 (after all, $p \mid 0$) and thus is ≥ 1. On the other hand, when $|X|$ is divisible by p, then the fixed point congruence shows the number of fixed points is $\equiv 0 \bmod p$, so this number is a multiple of p. □

Example 4.3. Let G be a p-subgroup of $\mathrm{GL}_n(\mathbf{Z}/(p))$, where $n \geq 1$. Then there is a nonzero $v \in (\mathbf{Z}/(p))^n$ such that $gv = v$ for all $g \in G$. Indeed, because G is a group of matrices it naturally acts on the set $V = (\mathbf{Z}/(p))^n$. (The identity matrix is the identity function and $g_1(g_2v) = (g_1g_2)v$ by the rules of matrix-vector multiplication.) Since the set V has size $p^n \equiv 0 \bmod p$, the number of fixed points is divisible by p. The number of fixed points is at least 1, since the zero vector is a fixed point, so the number of fixed points is at least p.

A nonzero fixed point for a group of matrices can be interpreted as a simultaneous eigenvector with eigenvalue 1. These are the only possible simultaneous eigenvectors for G in $(\mathbf{Z}/(p))^n$ since every element of G has p-power order and the only element of p-power order in $(\mathbf{Z}/(p))^\times$ is 1 (so a simultaneous eigenvector for G in $(\mathbf{Z}/(p))^n$ must have eigenvalue 1 for each element of the group).

Theorem 4.1 can be used to prove existence theorems about finite groups (nonconstructively) if we can interpret a problem in terms of fixed points. For example, an element of a group G is in the center precisely when it is a fixed point for the conjugation action of G on itself. So if we want to show a class of groups has nontrivial centers then we can try to show there are fixed points for the conjugation action other than the identity element.

5. New Proofs Using Group Actions

In this section we prove two results using group actions (especially using Theorem 4.1): finite p-groups have a nontrivial center and if $p \mid |G|$ then G has an element of order p.

Theorem 5.1. *Let G be a nontrivial p-group. Then the center of G has size divisible by p. In particular, G has a nontrivial center.*

This theorem is due to Sylow.[4]

Proof. The condition that a lies in the center of G can be written as $a = gag^{-1}$ for all g, so a is a fixed by all conjugations. The main idea of the proof is to consider the action of G on itself ($X = G$) by conjugation and count the fixed points.

We denote the center of G, as usual, by $Z(G)$. Since G is a p-group, and $X = G$ here, the fixed point congruence (Theorem 4.1) implies $|G| \equiv |Z(G)| \bmod p$. Since $|G|$ is a power of p, we get $0 \equiv |Z(G)| \bmod p$, so $p \mid |Z(G)|$. Because $|Z(G)| \geq 1$, from $p \mid |Z(G)|$ we get $|Z(G)| \geq p$, so $Z(G) \neq \{e\}$. □

Corollary 5.2. *For prime p, every group of order p^2 is abelian.*

Proof. Let $|G| = p^2$. Nontrivial elements of G have order p or p^2. If G has an element of order p^2, then G is cyclic, hence abelian. So assume nontrivial elements of G have order p.

By Theorem 5.1 there is $x \neq e$ in $Z(G)$, so x has order p. Let $y \notin \langle x \rangle$. Since $x \in Z(G)$, x and y commute, so all powers x^i and y^j commute. Thus $\{x^iy^j : i, j \in \mathbf{Z}\}$ is an abelian subgroup of G. It is larger than $\langle x \rangle$ since it contains y, so its order is p^2 (the order divides p^2 and is bigger than p). Thus $\{x^iy^j : i, j \in \mathbf{Z}\} = G$, so G is abelian. □

With almost no extra work than the proof of Theorem 5.1, we can prove a stronger result.

Theorem 5.3. *For each nontrivial p-group G, $N \cap Z(G) \neq \{e\}$ for all nontrivial normal subgroups $N \lhd G$. That is, every nontrivial normal subgroup meets the center of G nontrivially.*

[4] See p. 588 of Théorèmes sur les groupes de substitutions, *Mathematische Annalen* **5** (1872), 584–594; URL https://eudml.org/doc/156588. English translation by Robert Wilson, URL http://www.maths.qmul.ac.uk/~raw/pubs_files/Sylow.pdf.

Proof. Argue as in the proof of Theorem 5.1, but let G act on N by conjugation. Since N is a nontrivial p-group, the fixed point congruence (Theorem 4.1) implies $N \cap Z(G)$ has size divisible by p. Thus $N \cap Z(G)$ is nontrivial. □

Theorem 5.4 (Cauchy)**.** *Let G be a finite group and p be a prime factor of $|G|$. Then G has an element of order p.*

Proof. The argument we give is due to James McKay[5]. We are looking for solutions to the equation $g^p = e$ other than $g = e$. It is not obvious in advance that there are such solutions. What we will do is work with a more general equation that has lots of solutions and then recognize solutions to the original equation $g^p = e$ as fixed points under a group action on the solution set of the more general equation.

We will generalize the equation $g^p = e$ to $g_1 g_2 \cdots g_p = e$. This is an equation in p unknowns. If we are given choices for $g_1, \ldots, g_{p-1}$ then g_p is uniquely determined as the inverse of $g_1 g_2 \cdots g_{p-1}$. Therefore the total number of solutions to this equation is $|G|^{p-1}$. By comparison, we have no idea how many solutions there are to $g^p = e$ and we only know one solution, the trivial one that we are not interested in.

Consider the solution set to the generalized equation:

$$X = \{(g_1, \ldots, g_p) : g_i \in G, g_1 g_2 \cdots g_p = e\}.$$

We noted above that $|X| = |G|^{p-1}$, so this set is big. The nice feature of this solution set is that cyclic shifts of one solution give us more solutions: if $(g_1, g_2, \ldots, g_p) \in X$ then so is $(g_2, \ldots, g_p, g_1)$. Indeed, $g_1 = (g_2 \cdots g_p)^{-1}$ and elements commute with their inverses so $g_2 \cdots g_p g_1 = e$. Successive shifting of coordinates in a solution can be interpreted as a group action of $\mathbf{Z}/(p)$ on X: for $j \in \mathbf{Z}/(p)$, let $j \cdot (g_1, \ldots, g_p) = (g_{1+j}, \ldots, g_{p+j})$, where the subscripts are interpreted modulo p. This shift is a group action. Since the group doing the acting is the p-group $\mathbf{Z}/(p)$, the fixed point congruence (Theorem 4.1) tells us

$$|G|^{p-1} \equiv |\{\text{fixed points}\}| \bmod p. \tag{5.1}$$

What are the points of X fixed by $\mathbf{Z}/(p)$? Cyclic shifts bring every coordinate eventually into the first position, so a fixed point of X is one where all coordinates are equal. Calling the common value g, we have $(g, g, \ldots, g) \in X$ precisely when $g^p = e$. Therefore (5.1) becomes

$$|G|^{p-1} \equiv |\{g \in G : g^p = e\}| \bmod p. \tag{5.2}$$

Up to this point we have not used the condition $p \mid |G|$. That is, (5.2) is valid for all finite groups G and primes p. This will be useful in Appendix A.

Since p divides $|G|$, the left side of (5.2) vanishes modulo p, so the right side is a multiple of p. Thus $|\{g \in G : g^p = e\}| \equiv 0 \bmod p$. Since $|\{g \in G : g^p = e\}| > 0$, there must be some $g \neq e$ with $g^p = e$. □

Cauchy's theorem has other proofs[6] that handle abelian and nonabelian G in different ways. The above proof treats all finite groups in the same way.

Remark 5.5. Letting G be a finite group where $p \mid |G|$, (5.2) says

$$|\{g \in G : g^p = e\}| \equiv 0 \bmod p. \tag{5.3}$$

[5] J. McKay, Another Proof of Cauchy's Theorem, *Amer. Math. Monthly* **66** (1959), 119.

[6] See https://kconrad.math.uconn.edu/blurbs/grouptheory/cauchypf.

Frobenius proved a more general result: when $d \mid |G|$,

$$|\{g \in G : g^d = e\}| \equiv 0 \bmod d.$$

The divisor d need not be a prime. However, the proof is not as direct as the case of a prime divisor, and we don't look at this more closely.

6. More Applications of Group Actions to Group Theory

In Theorem 1.7 we saw how to interpret a group action of G as a homomorphism of G to a symmetric group. We will now put this idea to use.

Theorem 6.1. *Every nonabelian group of order* 6 *is isomorphic to* S_3*.*

Proof. Let G be nonabelian with order 6. We will make G permute a set of size 3.

By Cauchy's theorem, G has elements a of order 2 and b of order 3. If a and b commute, then ab has order 6, so G is cyclic, which is not true. Thus a and b do not commute, so bab^{-1} is not 1 or a. Set $H := \langle a \rangle = \{1, a\}$, which is not a normal subgroup of G since $bab^{-1} \notin H$. There are 3 left H-cosets in G. Let G act on them by left multiplication. This group action is a homomorphism $\ell \colon G \to \mathrm{Sym}(G/H) \cong S_3$. If $g \in \ker(\ell)$ then $gH = H$, so $g \in H$. Thus $\ker(\ell)$ is $\{1\}$ or H. Since $H \ntriangleleft G$, H can't be a kernel, so $\ker(\ell) = \{1\}$: ℓ is injective. Both G and S_3 have order 6, so ℓ is an isomorphism $G \to S_3$. □

Theorem 6.2. *Let G be a finite group and H be a p-subgroup such that $p \mid [G : H]$. Then $p \mid [\mathrm{N}(H) : H]$. In particular, $\mathrm{N}(H) \neq H$.*

We are not assuming here that G is a p-group. The case when G is a p-group as well will show up in Corollary 6.4.

Proof. Let H (not G!) act on G/H by left multiplication. Since H is a p-group, the fixed point congruence Theorem 4.1 tells us

$$[G : H] \equiv |\{\text{fixed points}\}| \bmod p. \tag{6.1}$$

What is a fixed point here? It is a coset gH such that $hgH = gH$ for all $h \in H$. That means $hg \in gH$ for every $h \in H$, which is equivalent to $g^{-1}Hg = H$. This condition means $g \in \mathrm{N}(H)$, so the fixed points are the cosets gH with $g \in \mathrm{N}(H)$. Thus (6.1) says

$$[G : H] \equiv [\mathrm{N}(H) : H] \bmod p.$$

This congruence is valid for all p-subgroups H of a finite group G. When $p \mid [G : H]$, we read off from the congruence that the index $[\mathrm{N}(H) : H]$ can't be 1, so $\mathrm{N}(H) \neq H$. □

Example 6.3. Let $G = A_4$ and $H = \{(1), (12)(34)\}$. Then $2 \mid [G : H]$, so $\mathrm{N}(H) \neq H$. In fact, $\mathrm{N}(H) = \{(1), (12)(34), (13)(24), (14)(23)\}$.

Corollary 6.4. *Let G be a finite p-group. Every subgroup of G with index p is a normal subgroup.*

Proof. We give two proofs. First, let the subgroup be H, so $H \subset \mathrm{N}(H) \subset G$. Since $[G : H] = p$, one of these inclusions is an equality. By Theorem 6.2, $\mathrm{N}(H) \neq H$, so $\mathrm{N}(H) = G$. That means $H \triangleleft G$.

For a second proof, consider the left multiplication action of G on the left coset space G/H. By Theorem 1.7, this action can be viewed as a group homomorphism $\ell \colon G \to \mathrm{Sym}(G/H) \cong S_p$. Let K be the kernel of ℓ, so $K \triangleleft G$. We will show $H = K$. The quotient G/K embeds into S_p, so $[G : K] \mid p!$. Since $[G : K]$ is a power of p, $[G : K] = 1$ or p.

Each $g \in K$ satisfies $gH = H$, so $g \in H$. In other words, $K \subset H$, so $[G : K] > 1$. Thus $[G : K] = p$, so $[H : K] = [G : K]/[G : H] = 1$, *i.e.*, $H = K \lhd G$. □

Corollary 6.5. *Let G be a finite group and p be a prime with $p^n \mid |G|$. Then there is a chain of subgroups*

$$\{e\} = H_0 \subset H_1 \subset \cdots \subset H_n \subset G,$$

where $|H_i| = p^i$.

Proof. We can take $n \geq 1$. Since $p \mid |G|$ there is a subgroup of size p by Cauchy's theorem, so we have H_1. Assuming for some $i < n$ we have a chain of subgroups up to H_i, we will find a subgroup H_{i+1} with size p^{i+1} that contains H_i.

Since $p \mid [G : H_i]$, by Theorem 6.2 $p \mid [\mathrm{N}(H_i) : H_i]$. Since $H_i \lhd \mathrm{N}(H_i)$, we can consider the quotient group $\mathrm{N}(H_i)/H_i$. It has size divisible by p, so by Cauchy's theorem there is a subgroup of size p. The inverse image of this subgroup under the reduction map $\mathrm{N}(H_i) \to \mathrm{N}(H_i)/H_i$ is a group H_{i+1} of size $p|H_i| = p^{i+1}$. □

Theorem 6.6 (C. Jordan)**.** *If a nontrivial finite group acts on a finite set of size greater than 1 and the action has only one orbit then some $g \in G$ has no fixed points.*

Proof. By Theorem 3.27,

$$1 = \frac{1}{|G|}\sum_{g\in G} |\operatorname{Fix}_g(X)| = \frac{1}{|G|}\left(|X| + \sum_{g\neq e} |\operatorname{Fix}_g(X)|\right).$$

Assume all $g \in G$ have at least one fixed point. Then

$$1 \geq \frac{1}{|G|}(|X| + |G| - 1) = 1 + \frac{|X| - 1}{|G|}.$$

Therefore $|X| - 1 \leq 0$, so $|X| = 1$. This is a contradiction. □

Remark 6.7. Using the classification of finite simple groups, it can be shown [2] that g in Theorem 6.6 can be picked to have prime power order. There are examples showing it may not be possible to pick a g with prime order.

Theorem 6.8. *If a group G has a subgroup H with finite index, then* (*i*) *G has a normal subgroup with finite index contained in H and* (*ii*) *finitely many subgroups of G contain H.*

Proof. (i) Let $G \to \mathrm{Sym}(G/H)$ be the left multiplication action and let N be its kernel. Each $n \in N$ satisfies $nxH = xH$ for all $x \in G$, so in particular $nH = H$. Thus $n \in H$, so $N \subset H$. Since G/N embeds into $\mathrm{Sym}(G/H)$, which is finite due to $[G : H]$ being finite, $[G : N]$ is finite.

(ii) The subgroups of G that contain H also contain N and only finitely many subgroups of G contain N since G/N is finite, so only finitely many subgroups of G contain H. □

The following theorem, in a special case, is due to Poincaré [1], [5, p. 410]. Its method of proof is similar to Theorem 6.8(i), so that result is also often attributed to Poincaré.

Theorem 6.9. *If a group G has subgroups H and K with finite index, then G has a normal subgroup with finite index contained in $H \cap K$. In particular, $H \cap K$ has finite index in G.*

Proof. Let $G \to \mathrm{Sym}(G/H \times G/K)$ be the left multiplication action of G on $G/H \times G/K$, where $g(xH, yK) = (gxH, gyK)$. Let N be the kernel. Then the group G/N embeds into the finite group $\mathrm{Sym}(G/H \times G/K)$, so $[G : N]$ is finite.

Each $n \in N$ satisfies $nxH = xH$ and $nyK = yK$ for all $x, y \in G$, so in particular $nH = H$ and $nK = K$. Thus $n \in H \cap K$, so $N \subset H \cap K$. Thus $[G : H \cap K]$ is finite. □

Theorem 6.10. *Let G be a finite group and H a proper subgroup. Then $G \neq \bigcup_{g \in G} gHg^{-1}$. That is, the union of the subgroups conjugate to a proper subgroup do not fill up the whole group.*

Proof. We will give two proofs. The second will use group actions.

Each subgroup gHg^{-1} has the same size, namely $|H|$. How many different conjugate groups gHg^{-1} are there (as g varies)? For $g_1, g_2 \in G$,

$$\begin{aligned} g_1Hg_1^{-1} = g_2Hg_2^{-1} &\iff g_2^{-1}g_1Hg_1^{-1}g_2 = H \\ &\iff g_2^{-1}g_1H(g_2^{-1}g_1)^{-1} = H \\ &\iff g_2^{-1}g_1 \in \mathrm{N}(H) \\ &\iff g_1 \in g_2\,\mathrm{N}(H) \\ &\iff g_1\,\mathrm{N}(H) = g_2\,\mathrm{N}(H). \end{aligned}$$

Therefore the number of different subgroups gHg^{-1} as g varies is $[G : \mathrm{N}(H)]$. These subgroups all contain the identity, so they are not disjoint. Therefore, on account of the overlap at the identity, the size of $\bigcup_{g \in G} gHg^{-1}$ is strictly less than

$$[G : \mathrm{N}(H)]|H| = \frac{|G|}{|\mathrm{N}(H)|}|H| = \frac{|H|}{|\mathrm{N}(H)|}|G| \leq |G|,$$

so the union of all gHg^{-1} is not all of G.

For the second proof, we apply Theorem 6.6 to the action of G on $X = G/H$ by left multiplication. For a 'point' gH in G/H, its stabilizer is gHg^{-1}. By Theorem 6.6, some $a \in G$ has no fixed points, which means $a \notin \bigcup_{g \in G} gHg^{-1}$. □

Remark 6.11. Theorem 6.10 is not always true for infinite groups. For instance, let $G = \mathrm{GL}_2(\mathbf{C})$. Every matrix in G has an eigenvector, so we can conjugate each matrix in G to the form $\left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right)$. Thus $G = \bigcup_{g \in G} gHg^{-1}$, where H is the proper subgroup of upper triangular matrices. Theorem 6.10 *is* true for infinite G when the index $[G : H]$ is finite: by Theorem 6.8, G has a normal subgroup N with finite index contained in H, so if $G = \bigcup_{g \in G} gHg^{-1}$, then reducing both sides modulo N gives us $G/N = \bigcup_{\overline{g} \in G/N} \overline{g}(H/N)\overline{g}^{-1}$, which contradicts Theorem 6.10 since G/N is finite.

Remark 6.12. Here is a deep application of Theorem 6.10 to number theory. Suppose a polynomial $f(X)$ in $\mathbf{Z}[X]$ is irreducible and has a root modulo p for every p. Then $f(X)$ is linear. The proof of this requires Theorem 6.10 and complex analysis.

Remark 6.13. If $H \ntriangleleft G$, then $\bigcup_{g \in G} gHg^{-1}$ need not be a subgroup, but it can be, *e.g.*, if $G = S_4$ and $H = \langle (12) \rangle$, then $\bigcup_{g \in G} gHg^{-1} = \{(1), (12)(34), (13)(24), (14)(23)\}$ is the unique normal subgroup of S_4 with order 4. In general the subgroup of G generated by $\bigcup_{g \in G} gHg^{-1}$ is called the normal closure of H in G.

Corollary 6.14. *If H is a proper subgroup of the finite group G, there is a conjugacy class in G that is disjoint from H and its conjugate subgroups.*

Proof. Pick an $x \notin \bigcup_{g \in G} gHg^{-1}$ and use the conjugacy class of x. □

Theorem 6.15. *Let G be a finite group with $|G| > 1$, and p the smallest prime factor of $|G|$. Every subgroup of G with index p is a normal subgroup.*

Corollary 6.4 is a special case of Theorem 6.15. Group actions don't appear in the statement of Theorem 6.15, but they will play a role in its proof. According to [3, pp. 3-4], Theorem 6.15 was conjectured and proved by Ernst Straus when he was a student.

Proof. Let H be a subgroup of G with index p, so G/H is a set with size p. We will prove $H \lhd G$ in two ways using group actions.

Method 1. We will show H is the kernel of a homomorphism out of G, and thus is a normal subgroup of G. The argument will be similar to the second proof of Corollary 6.4.

Let G act on G/H by left multiplication, which (by Theorem 1.7) gives a group homomorphism

$$G \to \mathrm{Sym}(G/H) \cong S_p. \tag{6.2}$$

This homomorphism sends each $g \in G$ to the permutation ℓ_g of G/H, where $\ell_g(aH) = gaH$. We will show this homomorphism has kernel H.

Write the kernel of the homomorphism (6.2) as K, so $K \lhd G$. The group G/K embeds into S_p, so $[G : K] \mid p!$. Since $[G : K]$ divides $|G|$, whose smallest prime factor is p, $(|G|, p!) = p$. Therefore $[G : K]$ is 1 or p. Each $g \in K$ satisfies $gH = H$, so $g \in H$. Thus $K \subset H$, so $[G : K] = [G : H][H : K] = p[H : K]$. Thus $[G : K] = p$ and $[H : K] = 1$, so $H = K \lhd G$.

Method 2.[7] Let H act on G/H by left multiplication, which (by Theorem 1.7) gives a group homomorphism

$$H \to \mathrm{Sym}(G/H) \cong S_p. \tag{6.3}$$

This action of H on a p-element set fixes the coset H, so each orbit has size at most $p-1$. By the orbit-stabilizer formula, an orbit has length dividing $|H|$, which divides $|G|$. The only factor of $|G|$ that's at most $p-1$ is 1 (why?), so all orbits of the H-action in (6.3) have length 1. That means (6.3) is a trivial action: $hgH = gH$ for each $h \in H$ and $g \in G$. Therefore $g^{-1}hgH = H$, so $g^{-1}hg \in H$, which implies (as h varies) $g^{-1}Hg \subset H$, so $g^{-1}Hg = H$ since both sides have the same size. Since this last equation holds for all $g \in G$, $H \lhd G$. □

Some special cases of Theorem 6.15 are worth recording separately.

Corollary 6.16. *Let G be a finite group.*

a) If H is a subgroup with index 2, then $H \lhd G$.

b) If G is a p-group and H is a subgroup with index p, then $H \lhd G$.

c) If $|G| = pq$ where $p < q$ are different primes, then each subgroup of G with size q is a normal subgroup.

Proof. Parts a and b are immediate consequences of Theorem 6.15. For part c, note that a subgroup with size q is a subgroup with index p. This completes the proof. □

Part a can be checked directly, without the reasoning of Theorem 6.15: if $[G : H] = 2$ and $a \notin H$, then the two left cosets of H are H and aH, while the two right cosets of H are H and Ha. Therefore $aH = G - H = Ha$, so $H \lhd G$. Part b was already seen in Corollary 6.4. (In fact, our second proof of Corollary 6.4 used the same idea as the proof of Theorem 6.15.) Part c could also be checked directly with the Sylow theorems, which show

[7] I learned this from the answer by Bar Alon at https://math.stackexchange.com/questions/164244/normal-subgroup-of-prime-index.

a subgroup of order q in G is not just normal but in fact unique. In Theorem 6.15, these disparate results are unified into a single statement.

All of our applications of group actions in this section have been to finite groups. Here is an application to infinite groups.

Theorem 6.17. *A finitely generated group has finitely many subgroups of index n for each integer $n \geq 1$.*

Proof. Let G be a finitely generated group and H be a subgroup with finite index, say n. The left multiplication action of G on G/H is a group homomorphism $\ell \colon G \to \operatorname{Sym}(G/H)$. In this action, the stabilizer of the coset H is H ($gH = H$ if and only if $g \in H$).

Pick an enumeration of the n cosets in G/H so that the coset H corresponds to the number 1. This enumeration gives an isomorphism $\operatorname{Sym}(G/H) \cong S_n$, so we can make G act on the set $\{1, 2, \ldots, n\}$ and the stabilizer of 1 is H. Therefore we have constructed from each subgroup $H \subset G$ of index n an action of G on $\{1, 2, \ldots, n\}$ in which H is the stabilizer of 1. Since H is recoverable from the action, the number of subgroups of G with index n is bounded above by the number of homomorphisms $G \to S_n$. Since G is finitely generated, it has finitely many homomorphisms to the finite group S_n. Therefore G has finitely many subgroups of index n. □

I am not aware of a proof of this theorem that is fundamentally different from the one presented above.

This is probably a good place to warn the reader about a false property of finitely generated groups: a subgroup of a finitely generated group need not be finitely generated! However, every *finite-index* subgroup of a finitely generated group is finitely generated: if the original group has d generators, a subgroup with index n has at most $(d-1)n+1$ generators. This is due to Schreier.

Appendix A. Applications of Group Actions to Number Theory

We apply the fixed point congruence in Theorem 4.1 and its consequence (5.2) to derive three classical congruences modulo p: those of Fermat, Wilson, and Lucas.

Theorem A.1 (Fermat)**.** *If $n \not\equiv 0 \bmod p$, then $n^{p-1} \equiv 1 \bmod p$.*

Proof. It suffices to take $n > 0$, since $(-1)^{p-1} \equiv 1 \bmod p$. (This is obvious for odd p since $p-1$ is even, and for $p = 2$ use $-1 \equiv 1 \bmod 2$.) Apply (5.2) with the additive group $G = \mathbf{Z}/(n)$:

$$n^{p-1} \equiv |\{a \in \mathbf{Z}/(n) : pa \equiv 0 \bmod n\}| \bmod p. \tag{A.1}$$

Since $(p, n) = 1$, the congruence $pa \equiv 0 \bmod n$ is equivalent to $a \equiv 0 \bmod n$, so the right side of (A.1) is 1. □

Theorem A.2 (Wilson)**.** *For a prime p, $(p-1)! \equiv -1 \bmod p$.*

Proof. We consider (5.2) for $G = S_p$:

$$0 \equiv |\{\sigma \in S_p : \sigma^p = (1)\}| \bmod p.$$

An element of S_p has p-th power (1) when it is (1) or a p-cycle. The number of p-cycles is $(p-1)!$, and adding 1 to this gives the total count, so $0 \equiv (p-1)! + 1 \bmod p$. □

Theorem A.3 (Lucas). *Let p be a prime and $n \geq m$ be nonnegative integers. Write them in base p as*

$$n = a_0 + a_1 p + a_2 p^2 + \cdots + a_k p^k, \quad m = b_0 + b_1 p + b_2 p^2 + \cdots + b_k p^k,$$

with $0 \leq a_i, b_i \leq p-1$. Then

$$\binom{n}{m} \equiv \binom{a_0}{b_0}\binom{a_1}{b_1}\cdots\binom{a_k}{b_k} \bmod p.$$

Proof. We will prove the congruence in the following form: when $n \geq m \geq 0$, and $n = pn' + a_0$ and $m = pm' + b_0$, where $0 \leq a_0, b_0 \leq p-1$, we have

$$\binom{n}{m} \equiv \binom{a_0}{b_0}\binom{n'}{m'} \bmod p.$$

The reader should check this implies Lucas' congruence by induction on n.

Decompose $\{1, 2, \ldots, n\}$ into a union of p blocks of n' consecutive integers, from 1 to pn', followed by a final block of length a_0. That is, let

$$A_i = \{in' + 1, in' + 2, \ldots, (i+1)n'\}$$

for $0 \leq i \leq p-1$, so

$$\{1, 2, \ldots, n\} = A_0 \cup A_1 \cup \cdots \cup A_{p-1} \cup \{pn' + 1, \ldots, pn' + a_0\}.$$

For $1 \leq t \leq n'$, let σ_t be the p-cycle

$$\sigma_t = (t, n' + t, 2n' + t, \ldots, (p-1)n' + t).$$

This cycle cyclically permutes the numbers in $A_0, A_1, \ldots, A_{p-1}$ that are $\equiv t \bmod n'$. The σ_t's for different t are disjoint, so they commute. Set $\sigma = \sigma_1\sigma_2\cdots\sigma_{n'}$. Then σ has order p as a permutation of $\{1, 2, \ldots, n\}$ (fixing all numbers above pn').

Let X be the set of m-element subsets of $\{1, 2, \ldots, n\}$, so $|X| = \binom{n}{m}$. Let the group $\langle\sigma\rangle$ act on X. Since σ has order p, Theorem 4.1 tells us

$$|X| \equiv |\{\text{fixed points}\}| \bmod p.$$

The left side is $\binom{n}{m}$. We will show the right side is $\binom{a_0}{b_0}\binom{n'}{m'}$.

When is an m-element subset $M \subset \{1, 2, \ldots, n\}$ fixed by σ? If M contains a number from 1 to pn' then σ-invariance implies M contains a number in the range from 1 to n', *i.e.*, $M \cap A_0 \neq \emptyset$. Let M contain q numbers in A_0. Then M is the union of these numbers and their translates into each of the p sets $A_0, \ldots, A_{p-1}$, along with some set of numbers from $pn' + 1$ to $pn' + a_0$, say ℓ of those. Then $|M| = pq + \ell$. Since M has size $m = pm' + b_0$, we have $b_0 \equiv \ell \bmod p$. Both b_0 and ℓ lie in $[0, p-1]$, so $\ell = b_0$. Thus $q = m'$.

Picking a fixed point in X under σ is thus the same as picking m' numbers from 1 to n' and then picking b_0 numbers from $pn' + 1$ to $pn' + a_0$. Therefore the number of fixed points is $\binom{n'}{m'}\binom{a_0}{b_0}$, even in the case when $a_0 < b_0$ (in which case there are 0 fixed points, consistent with $\binom{a_0}{b_0} = 0$ in this case). □

Appendix B. A group action in physics

In this section we expand on Example 2.7 about the transformations of space and time under which the laws of physics should remain the same,

Our model for spacetime will be $\mathbf{R}^4$. In non-relativistic (classical) physics, points in $\mathbf{R}^4$ are labeled as $(t, \mathbf{x}) = (t, x, y, z)$, where t is time and x, y, z are 3 space coordinates. This is called *Galilean spacetime*. In relativistic physics, $\mathbf{R}^4$ is called *Minkowski spacetime* and its points are written as $(ct, \mathbf{x}) = (ct, x, y, z)$, where c is the speed of light. Since c is a speed and t is time, ct has units of length, just like each component of $\mathbf{x}$. (Physicists who work in relativity choose units that make $c = 1$, so t as time or as length has the same value.) Two differences between non-relativistic and relativistic physics are described in Table 2: in non-relativistic physics, speeds are unlimited and motion in space does not affect time, while in relativistic physics speeds (of physical objects) stay below c and some motions that we'll see below mix time and space coordinates in nontrivial ways. Such mixing is why it's good to make the time coordinate have the same units as the space coordinates by the device of using ct in place of t.

	Allowed speeds	Time/Space coordinates
Non-relativistic	Arbitrary	No mixing
Relativistic	Less than c	Mixing allowed

Table 2. Non-relativistic and relativistic comparisons

The basic transformations of spacetime that don't change physical laws are (i) translations in space and time, (ii) rotations of space, and (iii) traveling at a constant speed in a fixed direction. The transformations in (iii) are called "boosts" and are different in non-relativistic and relativistic physics. Transformations in (i) and (ii) are the same in both settings.

Non-relativistic spacetime transformations

(i) Translations in space and time. These are $(t, \mathbf{x}) \mapsto (t + s, \mathbf{x} + \mathbf{y})$ for a time change (or time shift) by s and space change by $\mathbf{y}$.

(ii) Rotations of space. These are $(t, \mathbf{x}) \mapsto (t, A\mathbf{x})$ where A is a rotation of $\mathbf{R}^3$ fixing the origin. Such rotations form the group $\mathrm{O}(3) = \{A \in \mathrm{M}_3(\mathbf{R}) : AA^\top = I_3\}$.

(iii) Boosts by velocity $\mathbf{v}$ (fixed speed $\|\mathbf{v}\|$ and direction $\widehat{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|$). A boost at speed v along the positive x-axis is $(t, x, y, z) \mapsto (t, tv + x, y, z)$. More generally, a boost $B_\mathbf{v}$ by velocity $\mathbf{v}$ is $B_\mathbf{v}(t, \mathbf{x}) = (t, t\mathbf{v} + \mathbf{x})$. Here $\mathbf{v}$ is an arbitrary (velocity) vector in $\mathbf{R}^3$. The effect of $B_\mathbf{v}$ on $(t, \mathbf{x})$ can be described as a 4×4 matrix transformation, where we write the coordinates of $\mathbf{v}$ as (v_x, v_y, v_z):

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ v_x & 1 & 0 & 0 \\ v_y & 0 & 1 & 0 \\ v_z & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}.$$

The composition of non-relativistic boosts is a non-relativistic boost: $B_\mathbf{v} \circ B_\mathbf{w} = B_{\mathbf{v}+\mathbf{w}}$.

Relativistic spacetime transformations

(i) Translations in space and time. These are $(ct, \mathbf{x}) \mapsto (c(t + s), \mathbf{x} + \mathbf{y})$ for a time change by s and space change by $\mathbf{y}$. This is the same as (i) above, except for using ct in the time coordinate.

(ii) Rotations of space. These are $(ct, \mathbf{x}) \mapsto (ct, A\mathbf{x})$ where $A \in \mathrm{O}(3)$. This matches (ii) above except for using ct in the time coordinate.

(iii) Boosts by velocity $\mathbf{v}$ (fixed speed $\|\mathbf{v}\| < c$ and direction $\widehat{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|$). At speed v along the positive x-axis it is $(ct, x, y, z) \mapsto (c\gamma(t + xv/c^2), \gamma(tv + x), y, z)$, where

$\gamma = 1/\sqrt{1-v^2/c^2}$. More generally, the relativistic boost by velocity $\mathbf{v}$ is

$$(ct,\mathbf{x}) \mapsto \left(\gamma ct + \gamma\frac{\mathbf{v}}{c}\cdot\mathbf{x},\ \gamma t\mathbf{v} + \mathbf{x} + \frac{\gamma-1}{\|\mathbf{v}\|^2}(\mathbf{x}\cdot\mathbf{v})\mathbf{v}\right), \tag{B.1}$$

where $\gamma = 1/\sqrt{1-\|\mathbf{v}\|^2/c^2} = 1/\sqrt{1-\mathbf{v}\cdot\mathbf{v}/c^2}$. The factor γ, which depends on the speed $\|\mathbf{v}\|$, is greater than 1. Its first-order approximation, by the binomial theorem, is $1+\frac{1}{2}\|\mathbf{v}\|^2/c^2$. As a 4×4 matrix transformation, the relativistic boost by $\mathbf{v}$ is

$$\begin{pmatrix} \gamma & \dfrac{\gamma v_x}{c} & \dfrac{\gamma v_y}{c} & \dfrac{\gamma v_z}{c} \\ \dfrac{\gamma v_x}{c} & 1+\dfrac{(\gamma-1)v_x^2}{\|\mathbf{v}\|^2} & \dfrac{(\gamma-1)v_xv_y}{\|\mathbf{v}\|^2} & \dfrac{(\gamma-1)v_xv_z}{\|\mathbf{v}\|^2} \\ \dfrac{\gamma v_y}{c} & \dfrac{(\gamma-1)v_xv_y}{\|\mathbf{v}\|^2} & 1+\dfrac{(\gamma-1)v_y^2}{\|\mathbf{v}\|^2} & \dfrac{(\gamma-1)v_yv_z}{\|\mathbf{v}\|^2} \\ \dfrac{\gamma v_z}{c} & \dfrac{(\gamma-1)v_xv_z}{\|\mathbf{v}\|^2} & \dfrac{(\gamma-1)v_yv_z}{\|\mathbf{v}\|^2} & 1+\dfrac{(\gamma-1)v_z^2}{\|\mathbf{v}\|^2} \end{pmatrix}\begin{pmatrix} ct \\ x \\ y \\ z\end{pmatrix}.$$

As a compressed 2×2 matrix (with lower right entry being a 3×3 matrix), this is

$$\begin{pmatrix} \gamma & \gamma\mathbf{v}^\top/c \\ \gamma\mathbf{v}/c & I_3 + \dfrac{(\gamma-1)\mathbf{v}\mathbf{v}^\top}{\|\mathbf{v}\|^2}\end{pmatrix}\begin{pmatrix} ct \\ \mathbf{x}\end{pmatrix}. \tag{B.2}$$

When $\|\mathbf{v}\|$ is much smaller than c (that is, $\|\mathbf{v}\|/c$ is nearly 0), γ is nearly 1 and this makes the formula in (B.1) approximately $(ct, t\mathbf{v}+\mathbf{x})$. That is the non-relativistic boost by $\mathbf{v}$ if points in Galilean spacetime are labeled as $(ct,\mathbf{x})$, which illustrates how relativistic physics turns into classical physics at speeds much slower than the speed of light. (We are not taking into account here anything involving mass, which has its own effects in relativity even at low speeds.)

While non-relativistic boosts affect only space coordinates, relativistic boosts affect time and space coordinates, and as a result the composition of two relativistic boosts need not be a relativistic boost.

Example B.1. We look at relativistic boosts along the positive x and y-axes, by $(3/5)c\mathbf{e}_1$ and by $(3/5)c\mathbf{e}_2$. Since $\gamma((3/5)c) = 1/\sqrt{1-(3/5)^2} = 5/4$, these two boosts are

$$\begin{pmatrix} 5/4 & 3/4 & 0 & 0 \\ 3/4 & 5/4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1\end{pmatrix} \quad\text{and}\quad \begin{pmatrix} 5/4 & 0 & 3/4 & 0 \\ 0 & 1 & 0 & 0 \\ 3/4 & 0 & 5/4 & 0 \\ 0 & 0 & 0 & 1\end{pmatrix}, \tag{B.3}$$

respectively. Their product in either order is not a relativistic boost: a relativistic boost matrix is symmetric and the product of the matrices in (B.3) in either order is not symmetric.

In both non-relativistic and relativistic physics we can put the transformations of type (i), (ii), and (iii) together into the action of a single group on $\mathbf{R}^4$.

Non-relativistic spacetime transformations

A translation vector $(s,\mathbf{y})\in\mathbf{R}^4$ acts on $\mathbf{R}^4$ in the natural way: $(s,\mathbf{y})(t,\mathbf{x}) := (s+t,\mathbf{y}+\mathbf{x})$. A rotation matrix $A\in \mathrm{O}(3)$ and velocity vector $\mathbf{v}\in\mathbf{R}^3$ act together on $\mathbf{R}^4$ by combining a rotation of space and a boost:

$$(\mathbf{v}, A)(t, \mathbf{x}) = \mathbf{v} \cdot (A \cdot (t, \mathbf{x})) = \mathbf{v} \cdot (t, A\mathbf{x}) = (t, t\mathbf{v} + A\mathbf{x}). \tag{B.4}$$

The group of isometries of $\mathbf{R}^3$ is all functions $\mathbf{x} \mapsto \mathbf{v} + A\mathbf{x}$ for $\mathbf{v} \in \mathbf{R}^3$ and $A \in \mathrm{O}(3)$. Under composition of functions, the group law on isometries is $(\mathbf{v}', A')(\mathbf{v}, A) = (\mathbf{v}' + A'\mathbf{v}, A'A)$ and this makes (B.4) an action on $\mathbf{R}^4$ by the isometry group of $\mathbf{R}^3$ (check!). Such transformations of $\mathbf{R}^4$ are called *Galilean transformations*. Non-relativistic boosts are the special case of (B.4) where $A = I_3$, and that is why non-relativistic boosts are called "rotation-free" Galilean transformations.

The group of isometries of $\mathbf{R}^n$ is denoted $\mathrm{E}(n)$ since distance is fundamental to the geometry of $\mathbf{R}^n$ as Euclidean space. Combining the above actions on $\mathbf{R}^4$ by $\mathbf{R}^4$ and $\mathrm{E}(3)$,

$$(s, \mathbf{y})((\mathbf{v}, A)(t, \mathbf{x})) = (s, \mathbf{y})(t, t\mathbf{v} + A\mathbf{x}) = (s + t, \mathbf{y} + t\mathbf{v} + A\mathbf{x}). \tag{B.5}$$

Since

$$(\mathbf{v}, A)((s, \mathbf{y})(t, \mathbf{x})) = (\mathbf{v}, A)(s + t, \mathbf{y} + \mathbf{x}) = (s + t, (s + t)\mathbf{v} + A\mathbf{y} + A\mathbf{x}),$$

the effects of $\mathbf{R}^4$ and $\mathrm{E}(3)$ on $\mathbf{R}^4$ do not commute.

Make the pair $((s, \mathbf{y}), (\mathbf{v}, A)) \in \mathbf{R}^4 \times \mathrm{E}(3)$ act on $\mathbf{R}^4$ according to (B.5):

$$((s, \mathbf{y}), (\mathbf{v}, A))(t, \mathbf{x}) := (s + t, \mathbf{y} + t\mathbf{v} + A\mathbf{x}). \tag{B.6}$$

The composition of the effects of $((s', \mathbf{y}'), (\mathbf{v}', A'))$ and $((s, \mathbf{y}), (\mathbf{v}, A))$ on $(t, \mathbf{x})$ is

$$\begin{aligned} ((s', \mathbf{y}'), (\mathbf{v}', A'))(((s, \mathbf{y}), (\mathbf{v}, A))(t, \mathbf{x})) &= ((s', \mathbf{y}'), (\mathbf{v}', A'))(s + t, \mathbf{y} + t\mathbf{v} + A\mathbf{x}) \\ &= (s' + (s + t), \mathbf{y}' + (s + t)\mathbf{v}' + A'(\mathbf{y} + t\mathbf{v} + A\mathbf{x})) \\ &= (s' + s + t, (\mathbf{y}' + A'\mathbf{y} + s\mathbf{v}') + t(\mathbf{v}' + A'\mathbf{v}) + (A'A)\mathbf{x}). \end{aligned}$$

Writing the final result as $((?, ?), (?, ?))(t, \mathbf{x})$ to fit (B.6) says we should multiply elements of $\mathbf{R}^4 \times \mathrm{E}(3)$ by the rule

$$((s', \mathbf{y}'), (\mathbf{v}', A'))((s, \mathbf{y}), (\mathbf{v}, A)) = ((s' + s, \mathbf{y}' + A'\mathbf{y} + s\mathbf{v}'), (\mathbf{v}' + A'\mathbf{v}, A'A)). \tag{B.7}$$

Since elements of $\mathrm{E}(3)$ compose by $(\mathbf{v}', A')(\mathbf{v}, A) = (\mathbf{v}' + A'\mathbf{v}, A'A)$, we can rewrite (B.7) as

$$((s', \mathbf{y}'), (\mathbf{v}', A'))((s, \mathbf{y}), (\mathbf{v}, A)) = ((s', \mathbf{y}') + (\mathbf{v}', A')(s, \mathbf{y}), (\mathbf{v}', A')(\mathbf{v}, A)),$$

where $(\mathbf{v}', A')(s, \mathbf{y}) = (s, A'\mathbf{y} + s\mathbf{v}')$, which is how $\mathrm{E}(3)$ acts on $\mathbf{R}^4$ by (B.4). Therefore (B.7) can be described as a semidirect product group $\boxed{\mathbf{R}^4 \rtimes_\varphi \mathrm{E}(3)}$, with φ coming from (B.4). This semidirect product group is the *Galilean group* of Galilean spacetime $\mathbf{R}^4$. The group $\mathbf{R}^4 \rtimes_\varphi \mathrm{E}(3)$ acts on $\mathbf{R}^4$ by (B.6) and the non-relativistic transformations (i), (ii), and (iii) of $\mathbf{R}^4$ are special cases of (B.6) by taking some components in $\mathbf{R}^4 \rtimes_\varphi \mathrm{E}(3)$ to be trivial.

Remark B.2. In (B.6), the origin $(t, \mathbf{x}) = (0, \mathbf{0})$ is fixed if and only if $(s, \mathbf{y}) = (0, \mathbf{0})$, so the elements of the Galilean group fixing the origin are the Galilean transformations. To distinguish the full Galilean group from the subgroup fixing the origin, the full group is called the inhomogeneous Galiean group and the subgroup fixing $\mathbf{0}$ is called the homogeneous Galilean group, by analogy with linear functions $mx + b$ from high school algebra being "inhomogeneous" due to a constant term and the linear functions mx from linear algebra being "homogeneous" due to the constant term being 0.

Relativistic spacetime transformations

The transformations (i), (ii), and (iii) of Minkowski spacetime combine to give the action of a group on $\mathbf{R}^4$ by the same method as in the non-relativistic case. First we make $\mathbf{R}^4$ act on Minkowski spacetime as time and space translations just as in the non-relativistic case: for $(ct, \mathbf{x}) \in \mathbf{R}^4$ and $(cs, \mathbf{y}) \in \mathbf{R}^4$, set $(cs, \mathbf{y})(ct, \mathbf{x}) = (c(s + t), \mathbf{y} + \mathbf{x})$. Each $A \in \mathrm{O}(3)$ acts on $\mathbf{R}^4$ by $A(ct, \mathbf{x}) = (ct, A\mathbf{x})$. Each velocity vector $\mathbf{v} \in \mathbf{R}^3$ acts on $\mathbf{R}^4$ by the relativistic boost (B.1). Putting these three transformations together leads to all the "symmetries" of Minkowski spacetime, which have a complicated composition formula. These transformations and their products form the *Poincaré group*. Its subgroup fixing the origin is built from rotations and boosts (no nonzero translations) and is called the *Lorentz group*.

The Galilean and Poincaré groups both have dimension $4 + 6 = 10$. Analogues on $\mathbf{R}^{n+1}$ (n space coordinates and 1 time coordinate) have dimension $(n + 1)(n + 2)/2$.

References

[1] Georg Essl, Answer to "Source of Poincaré's theorem on subgroups with finite index", History of Science and Mathematics Stack Exchange. April 21, 2025. URL https://hsm.stackexchange.com/questions/18494

[2] B. Fein, W. M. Kantor, M. Schacher, Relative Brauer groups II, *J. Reine Angew. Math.* **328** (1981), 39–57.

[3] B. R. Gelbaum and J. M. H. Olmstead, *Theorems and Counterexamples in Mathematics*, Springer, New York, 1990.

[4] I. M. Isaacs and M. R. Pournaki, "Generalizations of Fermat's Little Theorem Using Group Theory," Amer. Math. Monthly **112** (2005), 734–740.

[5] H. Poincaré, Les fonctions fuchsiennes et l'Arithmétique, *J. Math. Pures Appl.* **3** (1887), 405–464. URL https://www.numdam.org/item/JMPA_1887_4_3__405_0.pdf

[6] Wikipedia, Burnside's lemma, http://en.wikipedia.org/wiki/Burnside%27s_lemma.