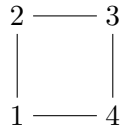
GROUP ACTIONS

KEITH CONRAD

1. Introduction

The symmetric groups S_n, alternating groups A_n, and (for $n \geq 3$) dihedral groups D_n behave, by their very definition, as permutations on certain sets. The groups S_n and A_n both permute the set $\{1, 2, \ldots, n\}$ and D_n can be considered as a group of permutations of a regular n-gon, or even just of its n vertices, since rigid motions of the vertices determine where the rest of the n-gon goes. If we label the vertices of the n-gon in a definite manner by the numbers from 1 to n then we can view D_n as a subgroup of S_n. For instance, the labeling



lets us regard the 90 degree clockwise rotation r in D_4 as (1234) and the reflection s across the x-axis as (12)(34). The rest of the elements of D_4, as permutations of the vertices, are in the table below.

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

If we label the vertices in a different way (*e.g.*, swap the labels 1 and 2), we may get a different subgroup of S_4.

More abstractly, if we are given any set X (not necessarily the set of vertices of a square), then the set $\mathrm{Sym}(X)$ of all permutations of the elements of X is a group under composition, and the subgroup $\mathrm{Alt}(X)$ of even permutations of X, 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.

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

Theorem 1.1 (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, $g_1(g_2 x) = (g_1 g_2)x$ for all $x \in G$), associating g to 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 an abstract group to behave like a group of permutations, as happened in the proof of Cayley's theorem, is a useful tool.

Definition 1.2. An *action* of a group G on a set X is a permutation $\pi_g\colon X \to X$, for each $g \in G$, 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 1.3. Let S_n act on $X = \{1, 2, \ldots, n\}$ in the usual way. Here $\pi_\sigma(i) = \sigma(i)$ in the usual notation.

Example 1.4. Any group G acts on itself ($X = G$) by left multiplication functions.

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 associated to g) 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 any 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.

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

Theorem 1.5. *Let G act on X. If $x \in X$, $g \in G$, and $y = g \cdot x$, then $x = g^{-1} \cdot y$. If $x \neq x'$ then $gx \neq gx'$.*

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 $gx = gx'$ 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 1.6. *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 $f(g_1 g_2) = f(g_1) \circ f(g_2)$. Think about the effect of the permutation $f(g)$ on $x \in X$ as an action: $g \cdot x = f(g)(x)$. This 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 $g \in G$ which 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 which act trivially on X are said to lie in the *kernel* of the action.

We will not often use the interpretation of Theorem 1.6 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 from any group. 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. The group S_n acts on n-variable polynomials $f(T_1, \ldots, T_n)$, by the rule

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

The reader should check $(1) \cdot f = f$ and $\sigma \cdot (\sigma' \cdot f) = (\sigma\sigma') \cdot f$, so (2.1) is a group action. Lagrange's study of this group action (*ca.* 1770) marked the first systematic use of symmetric groups in algebra. Lagrange wanted to understand why nobody up to that time had found an analogue of the quadratic formula for roots of a polynomial of degree greater than four. While Lagrange was not completely successful, he found in this group action that there were some different features in the cases $n \leq 4$ and $n = 5$.

Example 2.2. For $n \geq 3$, the group D_n acts on a regular n-gon as rigid motions. We can also view it as acting just on the n vertices of the regular n-gon. This does not give up any information, since knowing where the vertices go determines the rest of the rigid motion. By restricting the action of D_n from the whole n-gon to the n vertices, and labelling the vertices by $1, 2, \ldots, n$ in some manner, we can make D_n act on $\{1, 2, \ldots, n\}$.

Example 2.3. Let G be the group of Rubik's cube: all sequences of motions on the cube (keeping center facelets in fixed locations). This group acts on two different sets: the 12 edge cubelets and the 8 corner cubelets. Or we could let G act on the set of all 20 non-centerface cubelets together.

Example 2.4. 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.

Example 2.5. 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.

Example 2.6. We let S_n act on $\mathbf{R}^n$ by permutations of coordinates: for $\sigma \in S_n$ and $v = (c_1, \ldots, c_n) \in \mathbf{R}^n$, set $\sigma \cdot v = (c_{\sigma(1)}, \cdots, c_{\sigma(n)})$. For $\sigma' \in S_n$, we compute $\sigma' \cdot (\sigma \cdot v)$ by

setting $d_i = c_{\sigma(i)}$, so $\sigma \cdot v = (d_1, \ldots, d_n)$. Then

$$\begin{aligned} \sigma' \cdot (\sigma \cdot v) &= (d_{\sigma'(1)}, \ldots, d_{\sigma'(n)}) \\ &= (c_{\sigma(\sigma'(1))}, \ldots, c_{\sigma(\sigma'(n))}) \\ &= (c_{(\sigma\sigma')(1)}, \ldots, c_{(\sigma\sigma')(n)}) \\ &= (\sigma\sigma') \cdot v, \end{aligned}$$

and the order of multiplication is backwards! So in fact this action of S_n on $\mathbf{R}^n$ is not an action after all. To make things turn out correctly, we should redefine $\sigma \cdot v = (c_{\sigma^{-1}(1)}, \cdots, c_{\sigma^{-1}(n)})$. Then $\sigma \cdot (\sigma' \cdot v) = (\sigma\sigma') \cdot v$ and we have a genuine action of S_n on $\mathbf{R}^n$.

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 2.7. To make G act on itself by *left multiplication*, we let $X = G$ and $g \cdot x$ (for $g \in G$ and $x \in G$) be the usual product of g and x. This example was used already in the proof of Cayley's theorem, and the definition of a group action is satisfied by the axioms for multiplication 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 (non-standard terminology).

Example 2.8. To make G act on itself by *conjugation*, take $X = G$ 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 2.9. 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.) We 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 2.7 is the special case when H is trivial.

Example 2.10. 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 $\mathbf{Z}/(4)$.

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 2.11. 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 2.7 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 2.12. As a special case of Example 2.11, 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 $\{x_1, \ldots, x_6\}$, (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).

Example 2.13. 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 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 2.7) associates different permutations to different group elements, the conjugation action of G on itself (Example 2.8) can make different group elements act in the same way: if $g_1 = g_2z$, 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 2.14. 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 2.15. The action of G on itself by left multiplication is faithful: different elements send e to different places.

Example 2.16. 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 2.17. When H is a subgroup of G and G acts on G/H by left multiplication, two elements g_1 and g_2 act in the same way on G/H precisely when $g_1 gH = g_2 gH$ 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 2.18. 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 $\binom{0}{1}$.

Non-faithful actions are as important as non-injective group homomorphisms (in fact, that is precisely what a non-faithful action is from the viewpoint of Theorem 1.6).

Remark 2.19. 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 $A \in \mathrm{GL}_n(\mathbf{R})$.

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 Example 2.7 and in Example 2.6. We will not use right actions (except in Example 3.18), 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 G act on 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: it is the set of places where the point can be moved by the group action. On the other hand, the stabilizer of a point is an *algebraic* concept: it is the set of group elements which 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 2.7 and 2.8, 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 non-zero 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 non-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, any 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 corner cubelets and the 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 $k \in \{1, 2, \ldots, n\}$? It is the set of permutations of $\{1, 2, \ldots, n\}$ fixing k, which can be identified with the set of permutations of $\{1, 2, \ldots, n\} - \{k\}$. This is just S_{n-1} in disguise (once we identify $\{1, 2, \ldots, n\} - \{k\}$ in a definite manner with the numbers from 1 to $n - 1$). The stabilizer of any 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 $\{1, 2, \ldots, n\}$ which fix a number k can be identified with the even permutations of $\{1, 2, \ldots, n\} - \{k\}$, so the stabilizer of any 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 G acts on itself by left multiplication,

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

Example 3.10. When 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.
- a is a fixed point when it commutes with all elements of G, and thus the fixed points of conjugation form the center of G.

Example 3.11. When G acts on G/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 any fixed points (Example 3.9 with non-trivial 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 a non-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 G acts on X, any subgroup of G also acts on X. Let's look at two examples.

Example 3.13. Any subgroup $K \subset G$ acts on G/H by left multiplication. Then

- there could be more than one orbit (not all cosets gH can be reached from H by left multiplication by elements of K unless $KH = G$),
- $\mathrm{Stab}_{aH} = aHa^{-1} \cap K$,
- there are no fixed points (if $H \neq G$).

Example 3.14. Any subgroup $H \subset G$ acts on G by right-inverse multiplication (Example 2.7). Then

- the orbits are the left H-cosets ($\{gh^{-1} : h \in H\} = gH$)
- Stab_a is trivial,
- there are no fixed points (if $\#H > 1$).

Theorem 3.15. *Let G act on X.*

a) Different orbits are disjoint.

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

c) $g \cdot x = g' \cdot x$ if and only if g and g' lie in the same left coset of Stab_x. In particular, the length of the orbit of x is given by

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

The formula in part c, relating 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 which overlap must coincide. Suppose Orb_x and Orb_y have a common element z:

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

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

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

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

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

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

$$(g_1g_2) \cdot x = g_1 \cdot (g_2 \cdot x) = g_1 \cdot x = x,$$

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

$$g \cdot x = x \Longrightarrow x = g^{-1} \cdot x,$$

so Stab_x is closed under inversion.

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

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

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

c) The condition $g \cdot x = g' \cdot x$ is equivalent to $x = (g^{-1}g') \cdot x$, which means $g^{-1}g' \in \mathrm{Stab}_x$, or $g' \in g\,\mathrm{Stab}_x$. Therefore g and g' act in the same way on x if and only if g and g' lie in the same left coset of Stab_x. (Remember that for any subgroup H, $g' \in gH$ if and only if $g'H = gH$.)

Since Orb_x consists of the points $g \cdot x$ for varying g, and elements of G act in the same way on x if and only if they lie in the same left coset of Stab_x, we get a function $G \to \mathrm{Orb}_x$ by $g \mapsto gx$ which is surjective and the inverse image of each point in Orb_x is a left coset of Stab_x. Thus $\#\,\mathrm{Orb}_x$ is the number of left cosets of Stab_x in G, which is the index $[G : \mathrm{Stab}_x]$. □

Remark 3.16. That different orbits of a group action are disjoint includes as special cases two basic disjointness results in group theory: disjointness of conjugacy classes and disjointness of left cosets of a subgroup. The first case uses the action of a group on itself by conjugation, having conjugacy classes as its orbits. The second case uses the right-inverse multiplication action of the subgroup on the group (Example 3.14).

Example 3.17. Let S_3 act on itself by conjugation. Its orbits are its conjugacy classes, which are

$$\{(1)\}, \quad \{(12), (13), (23)\}, \quad \{(123), (132)\}.$$

The conjugacy class of (12) has size 3 and the stabilizer of (12) is its centralizer $\{(1), (12)\}$, which has index 3 in S_3.

Example 3.18. The 2×2 matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbf{R})$ whose columns add up to 1 are a group. This can be checked by a tedious calculation. But it is much simpler to see this if we observe that the column sums are the entries in the vector-matrix product $(1\ 1)\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, so the matrices of interest are those satisfying $(1\ 1)\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = (1\ 1)$. This 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 $\mathbf{v} \cdot A = \mathbf{v}A$, so it is a subgroup of $\mathrm{GL}_2(\mathbf{R})$ since point-stabilizers are always a subgroup. (Theorem 3.15 for right group actions should be formulated and checked by the reader.)

Moreover, because $(0\ 1)\left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right) = (1\ 1)$, $\mathrm{Stab}_{(1\ 1)}$ and $\mathrm{Stab}_{(0\ 1)}$ are conjugate subgroups in $\mathrm{GL}_2(\mathbf{R})$. Since $\mathrm{Stab}_{(0\ 1)} = \{\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{GL}_2(\mathbf{R})\} = \mathrm{Aff}(\mathbf{R})$, a model for our "column-sum-1 group" is its conjugate subgroup $\mathrm{Aff}(\mathbf{R})$. Explicitly,

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

Example 3.19. 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 the set HK like this: $(h, k) \cdot x = hxk^{-1}$. Check this gives a group action (the group is $H \times K$ and the set is HK). There is only one orbit since $e = ee \in HK$ and $hk = (h, k^{-1}) \cdot e$. 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 3.20. We now discuss the original version of Lagrange's theorem in group theory. Here is what he proved: for any polynomial $f(T_1, \ldots, T_n)$ in n variables, the number of different polynomials we can obtain from $f(T_1, \ldots, T_n)$ through all permutations of its variables is a factor of $n!$.

For instance, taking $n = 3$, consider the polynomial T_1. If we run through all six permutations of the set $\{T_1, T_2, T_3\}$, and apply each to T_1, we get 3 different results: T_1, T_2, and T_3. The polynomial $(T_1 - T_2)T_3 + (T_2 - T_3)T_1 + (T_3 - T_1)T_2$ has only 2 possibilities under any change of variables: itself or its negative (check this for yourself). The polynomial $T_1 + T_2^2 + T_3^3$ has 6 different possibilities. The number of different polynomials we find 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 2.1. That was 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. Therefore, by the orbit-stabilizer formula, the number of polynomials we get from $f(T_1, \ldots, T_n)$ by permuting its variables is $[S_n : H_f]$, where $H_f = \{\sigma \in S_n : \sigma \cdot f = f\}$. This index is a divisor of $n!$.

The following two corollaries to Theorem 3.15 are reinterpretations of parts of Theorem 3.15, and the proofs are left to the reader.

Corollary 3.21. *Let G act on X, where G is finite.*

a) The length of every orbit 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.

Corollary 3.22. *Let G act on X, where G and X are finite. Let the different orbits of X be represented by $x_1, \ldots, x_t$. Then*

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

Example 3.23. For any finite group G, each conjugacy class in G has size dividing the size of G, since a conjugacy class in G is an orbit in the conjugation action of G on itself, so Corollary 3.21a applies. Moreover, for the conjugation action (3.1) is the class equation.

In a group action, the length of an orbit divides $\#G$, but the number of orbits usually does not divide $\#G$. For example, S_4 has 5 conjugacy classes, and 5 does not divide 24. But there is an interesting relation between the number of orbits and the group action.

Theorem 3.24. *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:*

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

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

Don't confuse $\operatorname{Fix}_g(X)$ with the fixed points for the action: $\operatorname{Fix}_g(X)$ only includes points fixed by the one element g. The set of all fixed points for the action is the intersection of the sets $\operatorname{Fix}_g(X)$ 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 and dividing by $\#G$ 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 m points in it, then the sum over the points in that orbit is a sum of $1/m$ for m 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. □

Corollary 3.25 (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 3.24,

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

Assume all $g \in G$ have at least 1 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 3.26. Using the classification of finite simple groups, it can be shown [1] that g in Corollary 3.25 can be picked to have prime power order. There are examples showing it may not be possible to pick a g with prime order.

Definition 3.27. 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 $f(gx) = gf(x)$ 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 $gf(x) = f(x)$, so the stabilizer subgroups of $x \in X$ and $f(x) \in Y$ are the same.

Example 3.28. 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 3.29. Let S_3 act on the conjugacy class $\{(12), (13), (23)\}$ by conjugation. This action on a 3-element set, which is described in Table 1 below, looks like the usual action of S_3 on $\{1, 2, 3\}$ if we identify (12) with 3, (13) with 2, and (23) with 1. Therefore this conjugation action for S_3 is equivalent to its natural action on $\{1, 2, 3\}$.

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

TABLE 1

Example 3.30. 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: if $A \in \mathrm{GL}_2(\mathbf{R})$ then $A(e_1, e_2) := (Ae_1, Ae_2)$ 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 two columns of a matrix in $\mathrm{GL}_2(\mathbf{R})$ are a basis of $\mathbf{R}^2$ (with the ordering of the columns putting an ordering on the columns as basis vectors: the first column is the first basis vector and the second column is the second basis vector) and two square matrices multiply through multiplication on the columns: $A\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(A\binom{a}{c}\ A\binom{b}{d}\right)$. Letting $f\colon \mathcal{B} \to \mathrm{GL}_2(\mathbf{R})$ by $f\left(\binom{a}{c}, \binom{b}{d}\right) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ 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 3.31. 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: write $K = g_0 H g_0^{-1}$ and let $f\colon G/H \to G/K$ by $f(gH) = g_0 g g_0^{-1} K$. This is well-defined since, for $h \in H$,

$$g_0 g h g_0^{-1} K = g_0 g h g_0^{-1} g_0 H g_0^{-1} = g_0 g H g_0^{-1} = g_0 g g_0^{-1} K.$$

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.

If H and K are non-conjugate 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 3.32. *An action of G with one orbit is equivalent to the left multiplication action of G on a left coset space.*

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

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

A particular case of Theorem 3.32 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 3.33. The action of G on X is called *free* when every point has a trivial stabilizer.

Example 3.34. The left multiplication action of a group on itself is free with one orbit.

Example 3.35. The antipodal action of $\mathbf{Z}/(2)$ on a sphere (where the nontrivial element acts by negation) is a free action. There are uncountably many orbits.

Free actions show up quite often in topology, and Example 3.35 is a typical illustration of that.

Example 3.36. For an integer $n \geq 2$, let X_n be the set of roots of unity of *exact* order n in $\mathbf{C}^\times$, so $\#X_n = \varphi(n)$. (For instance, $X_4 = \{i, -i\}$.) The group $(\mathbf{Z}/(n))^\times$ acts on X_n by $a \cdot \zeta = \zeta^a$. Since every element of X_n is 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), stabilizers are trivial. 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.

It is worth comparing faithful and free actions. An action is faithful (Definition 2.14) when $g_1 \neq g_2 \Rightarrow g_1x \neq g_2x$ for some $x \in X$ (different elements of G act differently at *some* point) while an action is free when $g_1 \neq g_2 \Rightarrow g_1x \neq g_2x$ for all $x \in X$ (different elements of G act differently at *every* point). 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, D_4 is a 2-group. Because all subgroups of a p-group have p-power index, the length of an orbit under an action by a p-group is a multiple of p *unless* the point is a fixed point, when its orbit has length 1. This leads to an important congruence modulo p when a p-group is acting.

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 $x_1, \ldots, x_t$, so Corollary 3.22 leads to

$$\#X = \sum_{i=1}^{t} \#\,\mathrm{Orb}_{x_i}\,. \tag{4.1}$$

Since $\#\,\mathrm{Orb}_{x_i} = [G : \mathrm{Stab}_{x_i}]$ and $\#G$ is a power of p, $\#\,\mathrm{Orb}_{x_i} \equiv 0 \bmod p$ 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

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

□

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 6 acts, we don't get a congruence mod 2 or mod 3. But when a group of size 9 acts, then we get a congruence mod 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, $p|0$) 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 non-zero $v \in (\mathbf{Z}/(p))^n$ 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 non-zero 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 simulatenous 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 involving finite groups (non-constructively) if we can interpret a problem in terms of fixed points. For example, an element of a group is in the center precisely when it is a fixed point for the conjugation action of the group on itself. Thus, if we want to show some class of groups has non-trivial centers then we can try to show there are fixed points other than the identity element for the conjugation action. Interpreting other existence theorems in terms of fixed points can be more subtle than this.

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 non-trivial center and Cauchy's theorem. Our proof that finite p-groups have a non-trivial center is actually the same as the usual proof without group actions, but presented in a more elegant way. We will prove Cauchy's theorem in two ways. Unlike the usual inductive proof that doesn't use group actions, the proofs we will give with group actions treat abelian and non-abelian groups in a uniform manner.

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

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

$$\#G \equiv \#Z(G) \bmod p. \tag{5.1}$$

(This is the class equation for G, reduced modulo p.) Since $\#G$ is a power of p, (5.1) says

$$0 \equiv \#Z(G) \bmod p,$$

so $p|\#Z(G)$. Because $\#Z(G) \geq 1$ (the identity is in $Z(G)$), $Z(G)$ contains at least p elements, so in particular $Z(G) \neq \{e\}$. □

With almost no extra work we can prove a stronger result.

Corollary 5.2. *Let G be a non-trivial p-group. For any non-trivial normal subgroup $N \lhd G$, $N \cap Z(G) \neq \{e\}$. That is, every non-trivial normal subgroup meets the center of G non-trivially.*

Proof. We argue as in the proof of Theorem 5.1, but let G act on N by conjugation. Since N is a non-trivial 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 5.3 (Cauchy)**.** *Let G be a finite group and p be a prime factor of $\#G$. Then some element of G has order p.*

Proof. (McKay) We are looking for solutions to the equation $g^p = e$ other than $g = e$. It is not obvious in advance that there are any such solutions. McKay's idea is to work with a more general equation which has many solutions and then recognize solutions to the original equation as fixed points under a group action on the solution set.

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 any 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 which 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

$$(\#G)^{p-1} \equiv \#\{\text{fixed points}\} \bmod p. \tag{5.2}$$

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 (5.2) becomes

$$(\#G)^{p-1} \equiv \#\{g \in G : g^p = e\} \bmod p. \tag{5.3}$$

Up to this point we have not used the condition $p|\#G$. That is, (5.3) is valid for any finite group G and any prime p. This will be useful in Appendix A.

When p divide $\#G$, the left side of (5.3) 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 $g^p = e$. □

Remark 5.4. Letting G be any finite group where $p|\#G$, (5.3) says

$$\#\{g \in G : g^p = e\} \equiv 0 \bmod p. \tag{5.4}$$

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

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

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.

Here is a second group action proof of Cauchy's theorem.

Proof. Let $n = \#G$ and $p|n$. We let the group $\mathbf{Z}/(p) \times G$ act on the set G^p by

$$(i, g) \cdot (g_1, g_2, \ldots, g_p) = (gg_{i+1}, gg_{i+2}, \ldots, gg_{i+p}),$$

where indices are interpreted modulo p. This is a group action.

Let $\Delta = \{(g, g, \ldots, g) : g \in G\}$ be the diagonal in G^p. The action of $\mathbf{Z}/(p) \times G$ on G^p preserves $X = G^p - \Delta$, and we consider the group action on this set, which has size $n^p - n$. Since $\#(\mathbf{Z}/(p) \times G) = pn$, all orbits have length dividing pn. Since $p|n$, pn does not divide $n^p - n$, so some orbit has length less than pn. Let $(g_1, g_2, \ldots, g_p)$ be a point in such an orbit, so this point has non-trivial stabilizer (why?). Let (i, g) be a *non-identity* element in its stabilizer. We will show $g \neq e$ and $g^p = e$. The condition that (i, g) fixes $(g_1, g_2, \ldots, g_p)$ is equivalent to

$$gg_{i+1} = g_1,\ gg_{i+2} = g_2,\ \ldots,\ gg_{i+p} = g_p.$$

Thus

$$\begin{aligned} g_1 &= gg_{i+1} \\ &= g \cdot gg_{2i+1} \quad (\text{since } g_k = gg_{i+k} \text{ for all } k) \\ &= g^2 \cdot g_{2i+1} \\ &= \cdots \\ &= g^r g_{ri+1} \quad \text{for all } r. \end{aligned}$$

If $g = e$ then $i \not\equiv 0 \bmod p$ and $g_{ri+1} = g_1$ for all r. Since $\{ri + 1 \bmod p : r \geq 1\} = \mathbf{Z}/(p)$, every g_k equals g_1, so $(g_1, \ldots, g_p) \in \Delta$, a contradiction. Therefore $g \neq e$. Taking $r = p$, $g_1 = g^p g_{pi+1} = g^p g_1$, so $g^p = e$. □

6. More Applications of Group Actions to Group Theory

In Theorem 1.6 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. *Any nonabelian group of order* 6 *is isomorphic to* S_3.

Proof. Let G be nonabelian with order 6. We will find a set of size 3 which G naturally permutes.

By Cauchy, G contains an element a of order 2 and b of order 3. Since G is nonabelian, a and b do not commute. Therefore bab^{-1} is neither 1 nor a. Set $H := \langle a \rangle = \{1, a\}$. This is not a normal subgroup of G since $bab^{-1} \notin H$. There are 3 left cosets in G/H. Let G act by left multiplication on G/H. This group action is a homomorphism $\ell\colon G \to \mathrm{Sym}(G/H) \cong S_3$. If g is in the kernel of ℓ then $gH = H$, so $g \in H$. Thus the kernel is either $\{1\}$ or H. Since H is not a normal subgroup, it can't be a kernel, so ℓ has trivial kernel: it is injective. Both G and S_3 have order 6, so ℓ is an isomorphism of G with S_3. □

Theorem 6.2. *Let G be any finite group and H be a p-subgroup such that $p|[G : H]$. Then $p|[\mathrm{N}(H) : H]$. 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 is exactly that $g \in \mathrm{N}(H)$, so the fixed points are the cosets gH with $g \in \mathrm{N}(H)$. Therefore (6.1) says

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

This congruence is valid for any p-subgroup H of a finite group G. When $p|[G : H]$, 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 $2|[G : H]$, 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. Any 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 $H \lhd G$.

For a second proof, consider the left multiplication action of G on the left coset space G/H. By Theorem 1.6, 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. We will show $H = K$. The quotient G/K embeds into S_p, meaning $[G : K]|p!$. Since $[G : K]$ is a power of p, $[G : K] = 1$ or p. At the same time, any $g \in K$ at least 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 $p^n|\#G$. 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 $p|\#G$ 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} which contains H_i.

Since $p|[G : H_i]$, by Theorem 6.2 $p|[\mathrm{N}(H_i) : H_i]$. 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. *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 $g_1, g_2 \in G$,

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

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 Corollary 3.25 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 Corollary 3.25, some $a \in G$ has no fixed points, which means $a \notin \bigcup_{g\in G} gHg^{-1}$. □

Remark 6.7. Theorem 6.6 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 any 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.

Remark 6.8. Here is a deep application of Theorem 6.6 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 6.6 and complex analysis.

Corollary 6.9. *If H is a proper subgroup of the finite group G, there is a conjugacy class in G which 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 6.10. *Let G be a finite group with $\#G > 1$, and p the smallest prime factor of $\#G$. Any subgroup of G with index p is a normal subgroup.*

Group actions don't appear in the statement of Theorem 6.10, but they will play a role in its proof.

Proof. Let H be a subgroup of G with index p, so G/H is a set with size p. We will prove $H \lhd G$ by showing H is the kernel of a homomorphism, and thus is a normal subgroup.

Let G act on G/H by left multiplication, which (by Theorem 1.6) gives a group homomorphism

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

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

Write the kernel of this homomorphism as K, so $K \lhd G$. To say $g \in K$ means ℓ_g is the identity permutation: $g(aH) = aH$ for all cosets aH. One of the cosets is H itself, so in particular $gH = H$, which implies $g \in H$. Therefore $K \subset H$.

By (6.2), since G/K is isomorphic to a subgroup of S_p, $\#(G/K)|p!$ by Lagrange. Since $[G : H] = p$ and $K \subset H \subset G$, we write

$$\#(G/K) = [G : K] = [G : H][H : K] = p[H : K],$$

so the relation $\#(G/K)|p!$ simplifies to

$$[H : K]|(p-1)!.$$

Since $[H : K]$ is a factor of $\#G$, its smallest prime factor is $\geq p$. But this index divides $(p-1)!$, so therefore $[H : K]$ doesn't have any prime factors. That means $[H : K] = 1$, or $H = K$. In particular, H is the kernel of a homomorphism out of G, so $H \lhd G$. □

Some special cases of Theorem 6.10 are worth recording separately.

Corollary 6.11. *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 any subgroup of G with size q is a normal subgroup.

Proof. Parts a and b are immediate consequences of Theorem 6.10. For part c, note that a subgroup with size q is a subgroup with index p. □

Part a can be checked directly, without the reasoning of Theorem 6.10: 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 6.10.) Part c can be checked directly using the explicit list of groups of size pq. In Theorem 6.10, 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 6.12. *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 $\ell\colon G \to \mathrm{Sym}(G/H)$. 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 $\mathrm{Sym}(G/H) \cong S_n$, 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 which is fundamentally different from the one presented here.

This is probably a good place to warn the reader about a false finiteness 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 (5.3) 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 (5.3) 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 (5.3) 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 non-negative 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 $A_0, A_1, \ldots, A_{p-1}$ which 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 any 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). □

References

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